LaTeXM1 Note M1 Note

X

December 30, 2025

Github: https://github.com/Baudelaireee/Notebook

Contents

1 9/28 Notes for «Calculus on Manifolds» by Spivak

Problem. Can we derive the explicit formula from an equation with several variables? Here are two examples :

$$x^2 + y^2 - 1 = 0 \tag{1}$$

and

$$e^{xy} + \sin y + x^2 - 2 = 0 \tag{2}$$

In the first example, it is clearly that we can express y as a function of x in a certain interval. But in the second example, it seems that we can draw a picture to visualize the implicit curve.

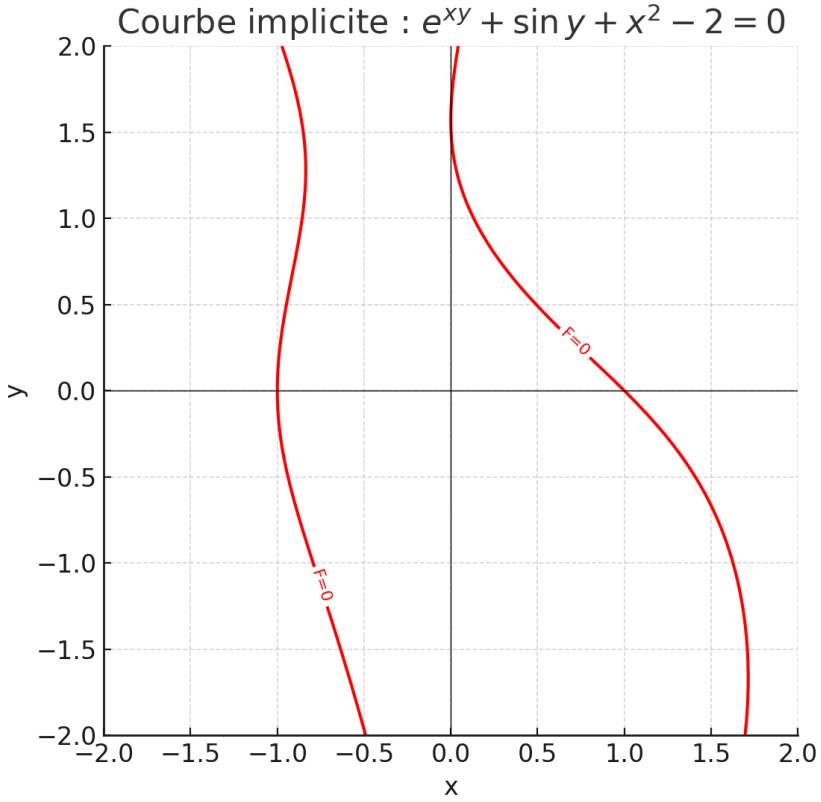


Figure 1: $e^{xy} + \sin y + x^2 - 2 = 0$

The curve here seems very smooth, so for example, at the point $(x_0, 0)$ in the curve we can express y as a function of x in the neighborhood, but it does not mean that we can write the function explicitly.

For thinking about this problem, we assume here we have a multivariable function $F(x, y)$ and we want to solve the equation $F(x, y) = 0$. Like equation (1), we can express it as $F(x, y) = x^2 + y^2 - 1$. Here we assume F is a smooth function, then we can get the total differential of F

$$df = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy$$

If we assume the existence of the implicit function $y = f(x)$, then along the curve

$(G(x) = F(x, f(x)) = 0)$ we can get

$$0 = G'(x) = \frac{\partial F}{\partial x}(x, f(x)) + \frac{\partial F}{\partial y}(x, f(x)) \cdot f'(x)$$

hence we can get the derivative of the function f can be denoted by

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))}$$

let us check some necessary condition here:
1. We try to restrict the function to be vanishing at some point, that means in the voisinage of the point, there exists something like a curve $F(x, y) = 0$ such that we can do the total differential like above, so here we need two conditions: locally, $F(a, b) = 0$ for some point (a, b) and a enough smooth function F.
2. We need to make sure the formula for $f'(x)$ makes sense, so locally $\frac{\partial F}{\partial y}(a, b) \neq 0$.

Actually, under these hypotheses, we can prove the existence of the implicit function $y = f(x)$ in the voisinage of the point (a, b) such that $F(x, f(x)) = 0$, which completes the implicit function theorem in the case of two variables.

Theorem 1.1 (Implicit Function Theorem in $\mathbb{R}^2$)**.** Let F be a C^1 function defined on an open set containing the point (a, b). Assume that $F(a, b) = 0$ and $\frac{\partial F}{\partial y}(a, b) \neq 0$. Then there exists an open interval I containing a, an open interval J containing b, and a unique C^1 function $f : I \to J$ such that for all $x \in I$, $F(x, f(x)) = 0$. Moreover, the derivative of f is given by

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))}$$

Proof. here we just give a proof of the existence of the implicit function f, the propoerties have benn proved as above. The idea here is to construct a locally diffeomorphism and then use the inverse function theorem.

We take the map $\Phi(x, y) = (x, F(x, y))$, then we can compute the Jacobian matrix of Φ at the point (a, b)

$$D\Phi(x, y) = \begin{pmatrix} 1 & 0 \\ \frac{\partial F}{\partial x}(x, y) & \frac{\partial F}{\partial y}(x, y) \end{pmatrix}$$

The condition $\frac{\partial F}{\partial y}(a, b) \neq 0$ implies $\det D\Phi(a, b) \neq 0$ the Jacobian matrix is invertible, hence by the inverse function theorem, there exists an open set W containing (a, b) and an open set W' containing $\Phi(a, b) = (a, 0)$ such that $\Phi : W \to W'$ is a C^1 diffeomorphism.

Finally we can construct the implicit function: the inverse $\Phi^{-1}(x, y) = (x, g(x, y))$ for some C^1 smooth function g in W', so we define the implicit function on the domain $D = \{x \in \mathbb{R} | (x, 0) \in W'\}$ (**RMQ:** here D is the intersection of W' and the x-axis, notice that at least $(a, 0) \in W'$, so D is not empty and can be seen as an open interval conatining a)

$$f(x) = g(x, 0)$$

then

$$\Phi(x, f(x)) = \Phi(x, g(x, 0)) = \Phi \circ \Phi^{-1}(x, 0) = (x, 0)$$

by the definition of Φ we can conlude that $F(x, f(x)) = 0$ □

Let us back to the two examples above, if we apply the implicit function theorem here, we can get a clear ODE for the implicit function $y = f(x)$, that is an equivalent view of the implicit function theorem.

Example 1.1. See the equation (2), we want to find the express of y as a function of x, so see the whole equation as a curve $F(x, y) = e^{xy} + \sin y + x^2 - 2$, then by the implicit function theorem, we can conclude a new realtionship between x and y:

$$y' = f'(x) = -\frac{F_x(x, y)}{F_y(x, y)} = -\frac{ye^{xy} + 2x}{xe^{xy} + \cos y}$$

so finding the explicit formula of y is equivalent to solving above ODE to get a explicit solution. The information here is given by the differentail structure of the curve, and notice that the ODE conatins the original information about the curve. However, sometimes we can find the explicit formula of the implicit function, for example, in equation (1), we can express y as a function of x explicitly:

$$y = \pm\sqrt{1 - x^2}$$

and the corresponding ODE here is

$$y' = -\frac{x}{y}$$

It's a separable ODE, we can solve it easily to get the explicit formula of the $y = f(x)$.

2 CA: Some examples and technics

Example 2.1. We consider two algebraic number fields $\mathbb{Q}(\sqrt{2})$ and $\mathbb{Q}(\sqrt{5})$. and we can study the algebraic intger ring in two fields, then we can find that

$$K = \mathbb{Q}(\sqrt{2}), \qquad \mathcal{O}_K = \mathbb{Z}[\sqrt{2}]$$

and

$$K = \mathbb{Q}(\sqrt{5}), \qquad \mathcal{O}_K = \mathbb{Z}[\frac{1 + \sqrt{5}}{2}]$$

here $X^2 + X + 1$ gives a monic polynomial with integer coefficients such that $\frac{1+\sqrt{5}}{2}$ is a root, that shows that $\frac{1+\sqrt{5}}{2}$ is an algebraic integer, so $\mathbb{Z}[\sqrt{5}] \subsetneq \mathcal{O}_K$.

It's a classic example about ring of algebraic integers, we can say more about quadratic fields

Theorem 2.1. Let d be an integer without square factors, and let $K = \mathbb{Q}(\sqrt{d})$ be a quadratic field. Then the ring of integers $\mathcal{O}_K$ is given by

$$\mathcal{O}_K = \begin{cases} \mathbb{Z}[\sqrt{d}] & \text{if } d \equiv 2, 3 \mod 4, \\ \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right] & \text{if } d \equiv 1 \mod 4. \end{cases}$$

Proof. One proof elementary is to use p-adic valuation, here is the outline:

- Prove the **main lemma**: $a + b\sqrt{d}$ is an algebraic integer if and only if $2a \in \mathbb{Z}$ and $a^2 - db^2 \in \mathbb{Z}$.
- Verify that if $a + b\sqrt{d}$ is an algebraic integer, then $v_p(a) \geq 0$ and $v_p(b) \geq 0$ for any odd prime p. (Notice that here is not sufficent to say $a, b \in \mathbb{Z}$)
- Consider the case of 2-adic valuation we can get the condition $v_2(a) \geq -1$ and $v_2(b) \geq -1$.
- Prove that $v_2(a) = -1$ and $v_2(b) = -1$ if and only if $d \equiv 1 \mod 4$, otherwise $v_2(a) \geq 0$ and $v_2(b) \geq 0$, then we can conclude the result.

□

Notice that it is not so easy to prove the similar result for cubic fields or higher degree fields, the concept of **integral basis** is useful here.

operatio of ideals: a useful techincs to compute the isomorphism.

Proposition 2.2. Let M be a A-module, I be an ideal of A and N be a submodule of M, then we have the isomorphism:

$$I(M/N) \cong (IM + N)/N$$

In particular, when $M = A$ and $N = J$ is a ideal, then we have

$$I(A/J) \cong (I + J)/J$$

Proof. we consider the natural map $\pi : M \to M/N$. We can verify that $I(M/N)$ is a submodule of M/N, then by correspondence theorem, there exists a submodule in M:

$$\begin{aligned} \pi^{-1}(I(M/N)) &= \pi^{-1}\{\sum_{\text{finite}} a_k \cdot (m_k + N) \mid a_k \in I, m_k \in M\} \\ &= \pi^{-1}\{\sum_{\text{finite}} a_k m_k + N \mid a_k \in I, m_k \in M\} \\ &= IM + N \end{aligned}$$

Then again by the image of natural map we can match

$$I(M/N) = (IM + N)/N$$

□

Proposition 2.3. If $I \subset R$ as a ideal, $a \in I$ and $b \notin I$, then

$$a + b \notin I$$

Proof. If $a + b \in I$, then $b = (a + b) - a \in I$, which is a contradiction. □

3 Fundamental group and Covering Space

3.1 Homotopy and π_1 functor

"deforming" is a topological concept, it refers to the process of continously transformation.

Definition 3.1. *Let $f, g : X \to Y$ be two continous maps between topological spaces X, we say that f is **homotopic** to g if there exists a continous map $H : X \times I \to Y$ such that*

$$H(x, 0) = f(x), \quad H(x, 1) = g(x)$$

and we denote it by $f \simeq_H g$.

*Furthermore, if there exists a subset $A \subset X$ such that for any $a \in A$, the homotopy $H(a, -) : I \to Y$ is constant, then we say that f is **homotopic relative to A** to g, and we denote it by $f \simeq_{H,A} g$.*

*In particualr, a map $f : X \to Y$ is (relative) **nullhomotopic** if it is relative homotopic to a constant map.*

One point is that realtive homotopy share the most properties with homotopy, the proof is similar, just checking the condition on A. Here are some basic properties of homotopy:

Proposition 3.1. Let $X, Y, Z \in$ **Top**
(1) (relative) homotopy defines **an equivalence relation** on $\mathcal{C}(X, Y)$.

homotopy behaves well in composition:
(2) If $f, g : X \to Y$ and $h : Y \to Z$, then (relative) homotopy

$$f \simeq_{H,A} g \implies h \circ f \simeq_{F,A} h \circ g$$

with $F(x, t) = h(H(x, t))$.
(3) If $f, g : X \to Y$ and $h : Z \to X$, then (realtive)homotopy

$$f \simeq_{H,A} g \implies f \circ h \simeq_{F,h^{-1}(A)} g \circ h$$

with $F(z, t) = H(h(z), t)$.
(3) If $f, g : X \to Y$ and $f', g' : Y \to Z$, then (relative)homotopy

$$f \simeq_{H,A} g, \quad f' \simeq_{H',A'} g' \implies f' \circ f \simeq_{F,A} g' \circ g$$

with

$$F(x, t) = \begin{cases} f'(H(x, 2t)) & t \in [0, \frac{1}{2}] \\ H'(g(x), 2t - 1) & t \in [\frac{1}{2}, 1] \end{cases}$$

The realtive homology holds when $f(A) \subset A'$ and $g(A) \subset A'$.

Proof. (1) the proof of reflexive and symmetric property is not difficult, but the proof of transitive properties needs a techinc lemma, which is very impotant in algebraic topology.

Lemma (Gluing Lemma)**.**
Let X be a topological spaces with closed subsets A and B such that $X = A \cup B$. If $f : A \to Y$ and $g : B \to Y$ are continous maps such that $f(x) = g(x)$ for all $x \in A \cap B$, then there exists a unique continous map $h : X \to Y$ such that $h|_A = f$ and $h|_B = g$.

with this lemma we can prove that if $f \simeq_F g$ and $g \simeq_H h$, then we can construct a homotopy $H : X \times I \to Y$

$$H(x,t) = \begin{cases} F(x,2t) & t \in [0, \frac{1}{2}] \\ H(x,2t-1) & t \in [\frac{1}{2}, 1] \end{cases}$$

then it finishes the proof. For the realtive case, let $F(x,-)$ and $H(x,-)$ be constant for any $x \in A$, then $F(x,1) = g(x)$ and $H(x,0) = g(x)$ show that $F(x,-) = G(x,-) = H(x,-)$.

(2)(3)(4) are direct verifications. □

Remark 3.1. Category **Top** is too big and rough that we cannot do more, with above properties, we can define the quotient category

$$\mathbf{hTop} \;\Big|\; \begin{array}{l} \text{objects: topological space } X \\ \text{morphisms: homotopy class } [f] \in \mathcal{C}(X,Y)/\sim \end{array}$$

similarly if we consider the realtive homotopy, it also defines a quotient category of pair topological sapce $\mathbf{Top}^2$, where the object is couple (X, A) with $A \subset X$ as a subspace, and the morphism between (X, A) and (Y, B) is a continous map $f : X \to Y$ such that $f(A) \subset B$, i.e. a commutative diagram:

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ {\scriptstyle i_A}\uparrow & & \uparrow{\scriptstyle i_B} \\ A & \xrightarrow[f|_A]{} & B \end{array}$$

Then the correspondence quotient category is denoted by

$$\mathbf{hTop}^2 \;\Big|\; \begin{array}{l} \text{objects: } (X,A) \in \mathbf{Top}^2 \\ \text{morphisms: } [f] \in \mathcal{C}(X,Y)/\sim_{relA} \end{array}$$

In above two categories, the composition of morphism is well-defined by the **composition of homotopy**, i.e. we can define

$$[f] \circ [g] := [f \circ g]$$

for any $g : X \to Y$ and $f : Y \to Z$. And the identity morphism of a object X is $[id_X]$.

In category, we define two objects are isomorphic if there exists two morphisms between them such that their composition is identity morphism, in above two categories, for any two topological spaces X and Y, they are isomorphic if there exists two classes

$[f] \in [X, Y]$ and $[g] \in [Y, X]$ such that

$$[f] \circ [g] = [id_Y] \quad \text{and} \quad [g] \circ [f] = [id_X]$$

which isomorphism is very important in algebraic topology, hence it deserves a name:

Definition 3.2. *Let X and Y be two topological spaces, we say that they are* ***homotopy equivalent*** *if there exists two continous maps $f : X \to Y$ and $g : Y \to X$ such that their compoisitions are homotopic to identity maps:*

$$f \circ g \simeq id_Y \quad \textit{and} \quad g \circ f \simeq id_X$$

The following object of the section is to define a functor from **Top** to other categories, here we see a simple example to **Set** caetgory, the detail can be found in Rotman's book (GTM 119).

Definition 3.3. $\pi_0 : \mathbf{Top} \to \mathbf{Set}$ *is a functor defined as following:*

- On objects: for any topological space X, $\pi_0(X) := X/\sim_c$, where $\sim_c$ is the equivalence relation defined by path-connectness, i.e. $X/\sim_c$ is the set of path-connected components of X.

- On morphisms: for any continous map $f : X \to Y$, we define

$$\pi_0(f) : \pi_0(X) \to \pi_0(Y), \quad [x] \mapsto [f(x)]$$

which maps a component to another compoent under image, whichi is ensured by the continuity.

should pay attention that hoemomorphic condition implies homotopy equivalence, but not vice versa, for example

$$\mathbb{S}^1 \text{ and } \mathbb{R}^2 - \{0\} \text{ are homotopy equivalent but not homeomorphic.}$$

Here π_0 gives a simple way to distinguish the spaces by counting the number of path-connected components, and we can prove the following proposition:

Proposition 3.2. If $f, g : X \to Y$ are homotopic, then $\pi_0(f) = \pi_0(g)$. In particular, If X, Y are homotopy equivalent, then there exists a bijection between $\pi_0(X)$ and $\pi_0(Y)$, i.e. they have same cardinality.

Proof. Let $H : X \times I \to Y$ be a homotopy from f to g, for any $x \in X$, we define a path $\gamma : I \to Y$ by

$$\gamma(t) = H(x, t)$$

then $\gamma(0) = f(x)$ and $\gamma(1) = g(x)$, hence $f(x)$ and $g(x)$ are in the same path-connected component, which implies that $\pi_0(f)([x]) = [f(x)] = [g(x)] = \pi_0(g)([x])$.

For the second part, if X and Y are homotopy equivalent, then there exists two continous maps $f : X \to Y$ and $g : Y \to X$ such that

$$f \circ g \simeq id_Y, \quad g \circ f \simeq id_X$$

then by the first part we have

$$\pi_0(f) \circ \pi_0(g) = \pi_0(f \circ g) = \pi_0(id_Y) = id_{\pi_0(Y)}$$

and similarly

$$\pi_0(g) \circ \pi_0(f) = id_{\pi_0(X)}$$

hence $\pi_0(f)$ and $\pi_0(g)$ are bijections between $\pi_0(X)$ and $\pi_0(Y)$. □

This simple construction shows the following concept:

(1) If a functor $F : \mathbf{Top} \to \mathbf{C}$ satisfies that for any two homotopic maps $f, g : X \to Y$, $F(f) = F(g)$, then it is called **homotpic invariant**, then the functor can induce a well-defined functor on **hTop**. Here π_0 is clearly a homotopic invariant, and we can induce a functor $\bar{\pi}_0 : \mathbf{hTop} \to \mathbf{Set}$ by $\bar{\pi}_0([f]) = \pi_0(f)$.

(2) It often happens that two spaces are not homeomorphic, but they have the similar shape (i.e. homotopy equivalent), usually some topological properties just depends on the the shape of spaces, hence homotopy equivalence will be a good condition to use.

(3) Actually π_0 is so limited that we can not talk more about the detail of the spaces, one reason is that **Set** is also too rough, we can not get more information from it.

From complex analysis and differential geometry, some information can be reflected by loops, i.e. closed paths. for example we consider a such application:

$$w : C(\mathbb{C}^*, 1) \to \mathbb{C}, \quad \gamma \mapsto \int_\gamma \frac{dz}{z}$$

where $C(\mathbb{C}^*, 1)$ denotes the set of closed paths in $\mathbb{C}^*$ starting from 1, with the basic knowledge we can know that all possible values of w is of the form $2\pi in$ for some integer n. Which requires us to classify all path closed paths in $\mathbb{C}^*$, and we can do a very rough partition here without proof, it is very intuitive:

$$C(\mathbb{C}^*, 1) = \sqcup_{k \in \mathbb{Z}} C_k$$

where C_0 denotes the set of closed paths which does not enclose the zero, while C_k for $k \neq 0$ denotes the set of closed paths which enclose the zero k times, moreover we make convention that the positive direction is counter-clockwise, hence which allows us to define a winding number for any closed path in $\mathbb{C}^*$:

$$W : C(\mathbb{C}^*, 1)/\sim \to \mathbb{Z}, \quad [\gamma] \mapsto \frac{1}{2\pi i} \int_\gamma \frac{dz}{z}$$

Furthermore we want to conclude a homomorphism to additive group, because we clearly know that if we concatenate two closed paths γ_1 and γ_2, then the winding number of the new path is the sum of two original winding numbers:

$$W([\gamma_1 * \gamma_2]) = W([\gamma_1]) + W([\gamma_2])$$

hence it gives a group isomorphism. Indeed, the naive patition above is just realtion of homotopy, if we take two homotopy closed path γ_1, γ_2, then there exists a homotopy H such that $\gamma_1 \simeq_{H,\{1\}} \gamma_2$, we can see that

$$W([\gamma_1]) - W([\gamma_2]) = \frac{1}{2\pi i}\int_{\gamma_1-\gamma_2}\frac{dz}{z} = \frac{1}{2\pi i}\int_S \frac{dz}{z} = 0$$

the surface S is enclosed by γ_1 and $-\gamma_2$, and they will not enclose th zero by homotopy, hence Stoke's theorem (or Cauchy's theorem) shows that the integral is zero since it is a closed form on S. It is a inspiration to the fundamental group theory, and it is a really classic and important example, Hence we should pay our attention to the closed paths of a spaces, and define a group via homotopy.

Definition 3.4. *For any two path $f, g : I \to X$ of a topological space with same endpoints, they are **path homotopic** if they are realtive homotopic with respect to $\{0, 1\}$, i.e. there exists a continous map $H : I \times I \to X$ such that*

$$H(s, 0) = f(s), \quad H(s, 1) = g(s), \quad H(0, t) = f(0) = g(0), \quad H(1, t) = f(1) = g(1)$$

for any $s, t \in I$.

As a special case of realtive homotopy, it defines an equivalence reation on $\mathcal{C}(I, X)$, i.e. the set of all paths in X. If g is a path starting from the endpoint of path f, i.e. $f(1) = g(0)$, we can connect it to get a new path, we can formlize it

$$(f * g)(t) := \begin{cases} f(2t), & t \in [0, 1/2] \\ g(2t-1), & t \in [1/2, 1] \end{cases}$$

It is the ideal way to define the algebra on the set of path homotopy classes, and in fact it works not bad:

Proposition 3.3. Let X be a topological space and $x_0, y_0 \in X$, let $C(X; x_0, y_0)$ be the set of all paths from x_0 to y_0 and $\sim$ be the path homotopy relation, then $C(X; x_0, y_0)/\sim$ forms a groupoid under the operation of concatenation of paths:

$$[f] * [g] := [f * g]$$

Proof. It needs to check the definition:
(1) operation is well-defined: if $f \simeq f'$ and $g \simeq g'$, then

$$[f] * [g] = [f * g] =^{?} [f' * g'] = [f'] * [g']$$

(2) associativity: for any three paths f, g, h with proper endpoints, we can verify that $([f] * [g]) * [h] = [f] * ([g] * [h])$ by direct computation.

□

In the system, there does not exists a real identity elemnt, but we can actually find right-identity and left-identity. We define two special paths here:

$$e_{x_0}(t) = x_0, \quad e_{y_0}(t) = y_0$$

for any $t \in I$, then for any path $f \in C(X; x_0, y_0)$, we can verify that

$$[e_{x_0}] * [f] = [f], \quad [f] * [e_{y_0}] = [f]$$

although the two constant paths are exactly not in $C(X; x_0, y_0)$, in particular when $x_0 = y_0$, they are the same path, and it is in $C(X; x_0, x_0)$, hence it will be the identity element of the groupoid, hence we can talk more if we consider about loops(The algebraic reason):

Proposition 3.4. Let X be a topological space and $x_0 \in X$, let $C(X; x_0)$ be the set of all loops based at x_0 and $\sim$ be the path homotopy relation, then $(C(X; x_0)/\sim, *, [e_{x_0}])$ forms a group under the operation of concatenation of paths.

Proof. Based on the previous proposition, we only need to check:
(1) $[e_{x_0}]$ is exactly the identity element.
(2) for any path f, it is natural to reverse the path:

$$\bar{f} : I \to X, \quad t \mapsto f(1-t)$$

then we can verify that its class is the inverse: $[f] * [\bar{f}] = [e_{x_0}]$ □

This group is called the **fundamental group** of X based at x_0, and it is denoted by $\pi_1(X, x_0)$, to consider the functoriality here, we consider a special case of pair topogical category

$\mathbf{Top}_*$	objects: topological space with base point (X, x_0) morphisms: continous map with $f(x_0) = y_0$

Notice that $f, g : (X, x_0) \to (Y, y_0)$ is homotopy if there exists a continous map $H : X \times I \to Y$ such that $f \simeq_{H,\{x_0\}} g$, i.e. relative homotopy with respect to the base point, it is induced by $\mathbf{hTop}^2$!

Definition 3.1. $\pi_1 : \mathbf{Top}_* \to \mathbf{Grp}$ is a covariant functor as following:

-on objects: fundamental group $\pi(X, x_0) = C(X; x_0)/\sim$

-on morphisms: For any continous map $f : (X, x_0) \to (Y, y_0)$, we define

$$\pi_1(f) : \pi_1(X, x_0) \to \pi_1(Y, y_0), \quad [\gamma] \mapsto [f \circ \gamma]$$

usually we denote $\pi_1(f)$ by f_*.

Here are some basic properties of fundamental group functor, it can be left to verify:

1. induce map f_* is well-defined as a group homomorphism.
2. **functoriality**: for any two continous maps $f : (X, x_0) \to (Y, y_0)$ and $g : (Y, y_0) \to (Z, z_0)$, then
$$(g \circ f)_* = g_* \circ f_*$$
3. Let $a \in X$ be a based point, and X_a is the path-connected component of X containing a, then the inclusion map $i : X_a \to X$ induces an isomorphism
$$i_* : \pi_1(X_a, a) \to \pi_1(X, a)$$

4. If $f : I \to X$ is a path connecting two based points x_0 and x_1, then it induces an isomorphism
$$f_{\#} : \pi_1(X, x_0) \to \pi_1(X, x_1), \quad [\gamma] \mapsto [f * \gamma * \bar{f}]$$
In particular, if X is path-connected, then $\pi_1(X, x_0) \cong \pi_1(X, x_1)$ for any two based points $x_0, x_1 \in X$.

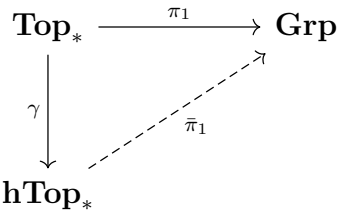
5. product keeps fundamental group: for any two based topological spaces (X, x_0) and (Y, y_0), then
$$\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0)$$

Finally, we conclude the homotopy invariant property of fundamental group functor, it is really important that which allows us to identify the fundamental group of homotopy equivalent spaces.

Theorem 3.5. If $f, g : (X, x_0) \to (Y, y_0)$ are homotopic continous maps, then $f_* = g_*$. In particular, if (X, x_0) and (Y, y_0) are homotopy equivalent, then their fundamental groups are isomorphic.

Proof. Let $f \simeq_H g$, then for any $[\gamma] \in \pi_1(X, x_0)$, we define a homotopy
$$F : I \times I \to Y, \quad (s, t) \mapsto H(\gamma(s), t)$$
It is continous by by composition, and $F(s, 0) = f \circ \gamma(s)$ and $F(s, 1) = g \circ \gamma(s)$, moreover $F(0, t) = H(x_0, t) = y_0$ and $F(1, t) = H(x_0, t) = y_0$, hence it shows that $f \circ \gamma \simeq g \circ \gamma$, i.e. $f_*([\gamma]) = g_*([\gamma])$. For the second part, if (X, x_0) and (Y, y_0) are homotopy equivalent, then there exists two continous maps $f : (X, x_0) \to (Y, y_0)$ and $g : (Y, y_0) \to (X, x_0)$ such that
$$f \circ g \simeq id_{(Y,y_0)}, \quad g \circ f \simeq id_{(X,x_0)}$$
then by the first part and functoriality we have
$$f_* \circ g_* = (f \circ g)_* = (id_{(Y,y_0)})_* = id_{\pi_1(Y,y_0)}$$
and similarly $g_* \circ f_* = id_{\pi_1(X,x_0)}$, QED. □

Remark 3.2. The theorem shows that fundamental group is a homotopy invariant, hence it induces a well-defined functor $\bar{\pi}_1$ such that the following diagram commutes:



i.e there exists a unique factorization of π_1 through $\mathbf{hTop}_*$, that means functor π_1 depends indeed on the homotopy type of spaces, and what we just compress the original category, the construction of fundamental have two steps in general: (1) Firstly we forget the specific topological structure of spaces, and just record the shape of spaces and classify them by homotopy equivalence, which gives the homotopy category $\mathbf{hTop}_*$.

(2) In each homotopy type, we choose a representation to study, and we study them in a certain viewpoint: look at the loops based at a point, actually it will reflect 1-dimensional information of the space, in other words, we forget some other higher-dimensiona information of the space.

3.2 Simple cases

The caculation of fundamental group is not easy in general, and the fundamental group induce a definition of a special coonect space:

Definition 3.5. *A topological space X is **simply connected** if it is path-connected and its fundamental group is trivial, i.e. for any based point $x_0 \in X$, $\pi_1(X, x_0) = \{[e_{x_0}]\}$.*

In complex analysis, simply connected spaces plays a important role and many good properties holds under this condition, and roughly speaking, simply connected spaces in 2-dimension is just a space without holes. It is easy to prove that $\mathbb{R}^n$ is simply connected for any $n \geq 1$, and more generally, we can conlude that:

Proposition 3.6. Any convex subset X of $\mathbb{R}^n$ is simply connected.

Proof. For any two points $x_0, x_1 \in X$, the line segment connecting them is contained in X by convexity, hence X is path-connected. For any loop $\gamma \in C(X; x_0)$, we can define a homotopy

$$H : I \times I \to X, \quad (s,t) \mapsto (1-t)\gamma(s) + tx_0$$

then clearly $H(s,0) = \gamma(s)$ and $H(s,1) = x_0$, moreover $H(0,t) = x_0$ and $H(1,t) = x_0$, hence it shows that $\gamma \simeq_{H,\{0,1\}} e_{x_0}$, which implies that $\pi_1(X, x_0)$ is trivial. □

As we have mentioned before, we construct a group isomorphism by winding integration:

$$C(\mathbb{C}^*, 1)/\sim \cong \mathbb{Z}$$

under the operation of concatenation, and we prove that same homotopy class implies the same winding number, hence realtion of homotopy is just the subrealtion here, and in fact that this realtion is exactly same with the path homotopy realtion, which means that

$$\pi_1(\mathbb{C}^*) = \mathbb{Z}$$

however, it is not easy to prove that same winding number implies path homotopy, by comparison with specific integration and analysis, homotopy is a more abstract concept, some new ideas and tools are needed. Here we admit the result of fundamental group of $\mathbb{C}^*$, and we will see it will be a key to caculate many fundamental groups later.

Claim 3.7.

$$\pi_1(\mathbb{S}^1) \cong \mathbb{Z}$$

The reason is that $\mathbb{C}^*$ is homotopy equivalent to $\mathbb{S}^1$, but more geometrically we realize that the fundamental group depends on the shape of spaces, and $\mathbb{S}^1$ can viewed as a "smallest" model same as $\mathbb{C}^*$, more precisely we define a special case of homotopy equivalent:

Definition 3.6. *Let X be a topological space and $i : A \hookrightarrow X$ be a embedding map.*

*-If there exists a continous map $r : X \to A$ such that $r \circ i = id_A$, then A is called a **retract** of X, and r is called a **retraction**.*

*-Furthermore if $i \circ r \simeq id_X$ gives a homotopy, then A is called a **deformation retract** of X, and r is called a **deformation retraction**.*

By the definition, we can **verify** that deformation retract is a special case of homotopy equivalent, hence it induces a general isomorphism of fundamental groups:

Claim 3.8.

$$\pi_1(\mathbb{S}^n) \cong \pi_1(\mathbb{R}^{n+1} - \{0\})$$

and similarly we can give some other examples:

$$\pi(A, a) \cong \mathbb{Z}$$

where A is the annulus $\{z \in \mathbb{C} : r < |z| < R\}$ for some $0 < r < R$, and $a \in A$. Pay attention that **retract is not necassary deformation retract**, for example the unit circle $\mathbb{S}^1$ is a retract of the disk $D = \{z \in \mathbb{C} : |z| \leq 1\}$ by the retraction, but they are not same homotopy type since $\pi_1(D) = \{0\}$ as a convex set.

We should also notice a particular case of deformation retract: the whole space can be contracted to just one point, i.e. the case of $A = \{\text{pt}\}$, it is indeed possible, for example we consider $\mathbb{R}^n$ and define a deformation retraction

$$r : \mathbb{R}^n \to \{0\}, \quad x \mapsto 0$$

with the inculsion $i : \{0\} \to \mathbb{R}^n$ by $i(0) = 0$, then we can verify

$$r \circ i(0) = 0, \quad i \circ r(x) = 0$$

and we can find a homotopy

$$H : \mathbb{R}^n \times I \to \mathbb{R}^n, \quad (x, t) \mapsto (1 - t)x$$

which satisfies

$$H(x, 0) = x, \quad H(x, 1) = 0$$

hence it shows that $\mathbb{R}^n$ is deformation retract to a point, which implies the fundamental group is just the trival group $\pi_1(\{pt\})$. In general, we can conclude the special spaces:

Definition 3.7. *Let X be a topological space, X is **contractible** if it can be deformation retracted to a point, equivalently**(verify!)**, the identiy map id_X is nullhomotopic.*

Pay attention that

contractible spaces is a special case of simply connected spaces.

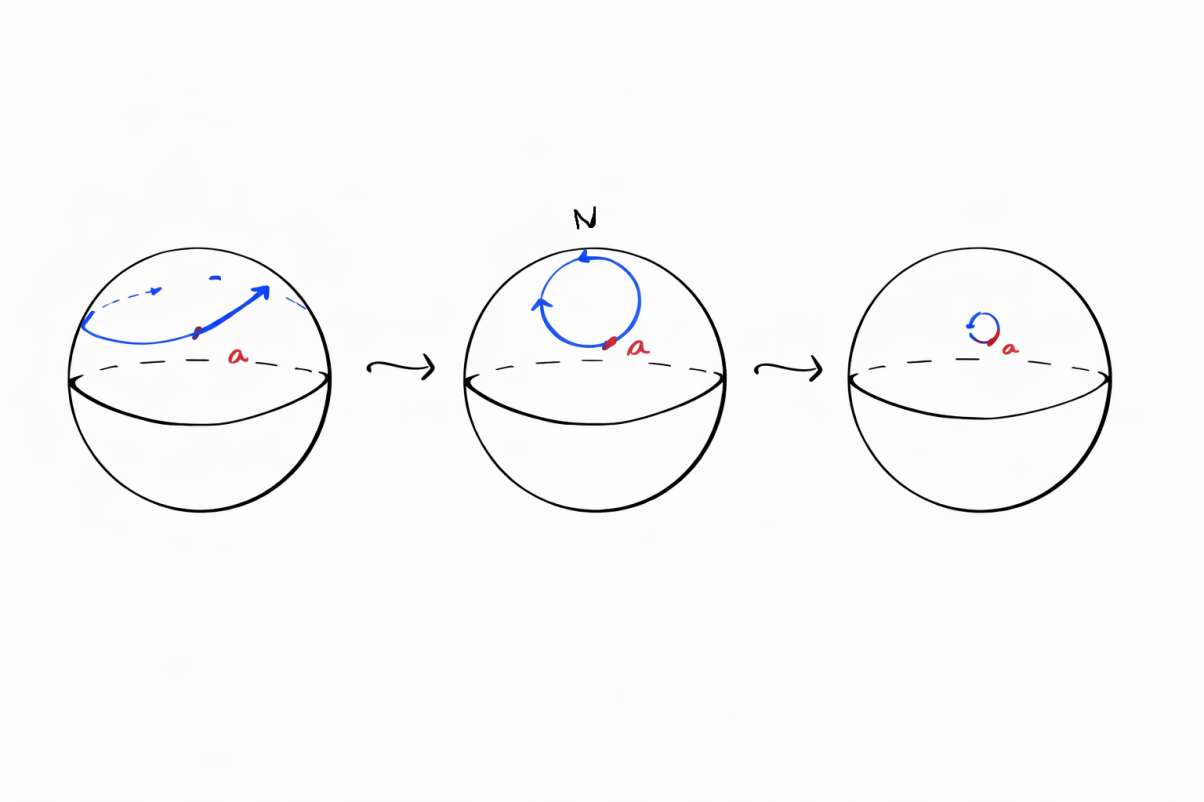


Figure 2: Contract a loop in $\mathbb{S}^2$ to a point by passing the pole

It is not true vice versa, a non-trival example is $\mathbb{S}^2$ whose fundamental group is trival, but it is not contractible. However, the caculation of fundamental group needs some discussion, what we can do here is just to see how a path in sphere can be contracted to a point, draw some pictures we can know the contious transformation of a loop can be done by passing the pole, it gives us a intuition that the fundamental group is trival.

To better intepration the idea, we introduce a special construction of contractible space:

Definition 3.8. *Let X be a topological space, the **cone** of X is defined by*

$$C(X) := (X \times I)/(X \times \{1\})$$

*i.e. we identify all points in $X \times \{1\}$ to a single point, which is called the **cone point**. Similarly, the suspension of X is defined by*

$$S(X) := (X \times I)/(X \times \{0\}, X \times \{1\})$$

*i.e. we identify all points in $X \times \{0\}$ to a single point, and all points in $X \times \{1\}$ to another single point, which are called the **suspension points**.*

with this construction, we can identify $\mathbb{S}^1$ by homeomorphism

$$\mathbb{S}^2 \cong S(\mathbb{S}^1) \cong C(\mathbb{S}^1) \sqcup_{\mathbb{S}^1} C(\mathbb{S}^1)$$

the last construction is to glue two copies along the $\mathbb{S}^1$, hence the question is reduced to consider when the cone and suspension are contractible

Proposition 3.9. For any topological space X,

- The cone $C(X)$ is contractible.

- If X is contractible, then the suspension $S(X)$ is also contractible.

Proof. □

However, we can not talk more about whether $\mathbb{S}^2$ is contractible or not, If $S(X)$ is contractible, we can not conclude that X is contractible or not, and the example is very messy, hence it is an ideal way to handle the problem. Anyway, the contractible property of $\mathbb{S}^n$ is equivalent to the brouwer fixed-point theorem, and someother important results in topology, it reflects the difficulty of the problem without algebraic topology tools.

Claim 3.10. $\mathbb{S}^n$ *is not contractible, in particular* $n \geq 2$

$$\pi_1(\mathbb{S}^n) \cong \{0\}$$

Theorem 3.11. Let $f : \mathbb{S}^1 \to X$ be a continous map, then the following statements are equivalent:

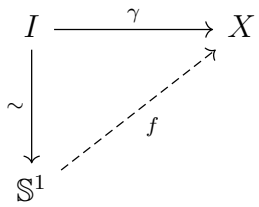
(1) f is nullhomotopic.

(2) f can be extended to a continous map $\tilde{f} : D^2 \to X$ such that $\tilde{f}|_{\mathbb{S}^1} = f$.

(3) The induced homomorphism $f_* : \pi_1(\mathbb{S}^1) \to \pi_1(X)$ is trivial.

Proof. □

Remark 3.3. It is a techinc and useful result, and generally we can think of fundamental group from the view of $\mathbb{S}^1$. For any closed path $\gamma : I \to X$ with $\gamma(0) = \gamma(1)$, we can identify it as a map from $\mathbb{S}^1$ to X by the following diagram



Then It makes sense to consider Hom functor $\mathrm{Hom}_{\mathbf{Top}_*}(\mathbb{S}^1, -)$ which admits object same with $C(X; x_0)$, and use the factorization and Yondea lemma we can get clearly the representation of fundamental group functor:

$$\pi_1(-) \cong \mathrm{Hom}_{\mathbf{hTop}_*}(\mathbb{S}^1, -)$$

It is the category construction of fundamental group functor, here if we use $\mathbb{S}^n$ to replace $\mathbb{S}^1$, then we can get the general homotopy group π_n.

3.3 Van-Kampen Theorem

For a Torus, by classic hoemomorphic we have $T^2 \cong \mathbb{S}^1 \times \mathbb{S}^1$, hence by product property of fundamental group, we have

$$\pi_1(T, a) \cong \pi_1(\mathbb{S}^1, a_1) \times \pi_1(\mathbb{S}^1, a_2) \cong \mathbb{Z} \times \mathbb{Z}$$

for any based point $a = (a_1, a_2) \in T^2$. Now if we remove a point from the torus, then will the fundamental group change? What's the new fundamental group? To see the homotopy type of $T^2 - \{pt\}$, we can consider the definition of torus by quotient

$$T^2 = [0,1]^2/\sim$$

with relation $(0, y) \sim (1, y)$ and $(x, 0) \sim (x, 1)$ for any $x, y \in [0, 1]$, then we can see that removing a point from the torus is just the quoitent space of the square removing a point inside it

$$\begin{array}{ccc} [0,1]^2 - \{p\} & \xrightarrow{\quad r \quad} & \partial[0,1]^2 \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ T^2 - \{\text{pt}\} & \xrightarrow[\bar{r}]{} & \partial[0,1]^2/\sim \end{array}$$

Here r is the deformation retraction from the square removing a point to its boundary, and quotient relation $\sim$ do not affect the deformation by definition, hence the induced map by UPQ-TOP $\bar{r}$ is exactly a deformation retraction, hence we can conclude the homotopy equivalence.

The object we get is actually a wedge sum of two circles, formally

Definition 3.2. Let (X, x_0) and (Y, y_0) be object of $\mathbf{Top}_*$, the **wedge sum** of X and Y is defined by

$$X \vee Y := (X \sqcup Y)/(x_0 \sim y_0)$$

i.e. we glue the two spaces at their base points. and we define the point $[x_0] = [y_0]$ in $X \vee Y$ as the base point

From the view of category, wedge sum is just the coproduct in $\mathbf{Top}_*$, and we (should) know that coproduct in **Grp** is the free product, hence it is natural to guess that the fundamental group of wedge sum is the free product of fundamental groups, and in fact it is true.

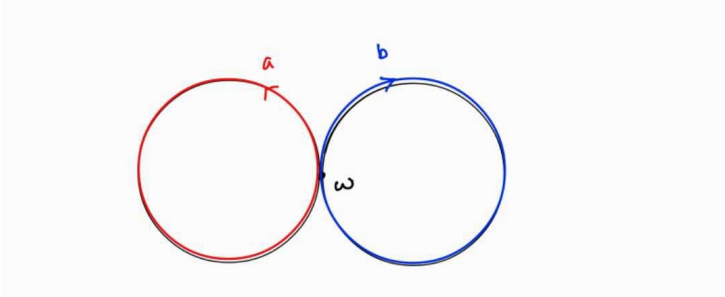


Figure 3: Wedge sum of two circles $\mathbb{S}^1 \vee \mathbb{S}^1$

let A and B be two copies of $\mathbb{S}^1$, and consider the wedge sum with the base point at w, and we suppose that a and b be two paths based at w going around A and B respectively once (see the figure), then clearly $[a]$ and $[b]$ are the generators of $\pi_1(A, w)$ and $\pi_1(B, w)$ respectively, but by conactenation of paths $a * b$ seems not be homotopic to $b * a$ (although it is not so clear to see it), hence the fundamental group seems and be generated by $\langle [a], [b] \rangle$, it is the free product group with rank 2.

Claim 3.12.

$$\pi_1(\mathbb{S}^1 \vee \mathbb{S}^1) \cong \pi_1(\mathbb{S}^1) * \pi_1(\mathbb{S}^1) \cong \mathbb{Z} * \mathbb{Z}$$

In generally, π_1 is not a left-adjoint functor, hence it does not preserve coproudcts, but in some certain case we can see that it does, and in particular preserve push out.

Conversely, we admit the result of fundamental group here, and then we think about the fundamental group of torus in another way, although we have known that $\pi_1(T^2) \cong \mathbb{Z} \times \mathbb{Z}$. we take small open disc U in T^2, and set $D \subsetneq U$ as a smaller closed disc, and then we set $V = T^2 - D$ to be another open set such that $U \cap V$ is homotopy equivalent to $\mathbb{S}^1$:

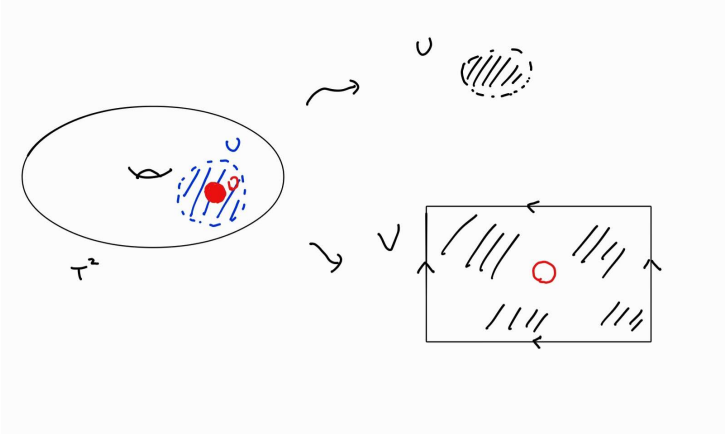


Figure 4: Open cover of T^2 by U and V

V can be seen as a deformation retraction of $T^2 - \{\text{pt}\}$, and U is simplely connected (in view of manifold, it is homeomorphic to $\mathbb{R}^2$), hence we get their fundamental groups by taking a based point $a \in U \cap V$:

$$\pi_1(U, a) \cong \{0\}, \quad \pi_1(V, a) \cong \mathbb{Z} * \mathbb{Z}, \quad \pi_1(U \cap V, a) \cong \mathbb{Z}$$

In the view of topology, T^2 can be viewed as the identification space of U and V along their intersection, i.e. we have the following diagram:

$$\begin{array}{ccc} U \cap V & \xrightarrow{i} & U \\ {\scriptstyle j}\downarrow & & \downarrow{\scriptstyle p} \\ V & \xrightarrow[q]{} & T^2 \end{array}$$

To explictly give the induced maps, we set the generator of $\pi_1(U \cap V, a) \cong \mathbb{Z}$ to be $[\gamma]$, and the generators of $\pi_1(V, a) \cong \mathbb{Z} * \mathbb{Z}$ to be $[\alpha]$ and $[\beta]$, then we can prove that induced

maps of j are $j_*([\gamma]) = [\alpha][\beta][\alpha]^{-1}[\beta]^{-1}$, although it is intuitive to see, but the proof is not trival, we skectch the outline here:

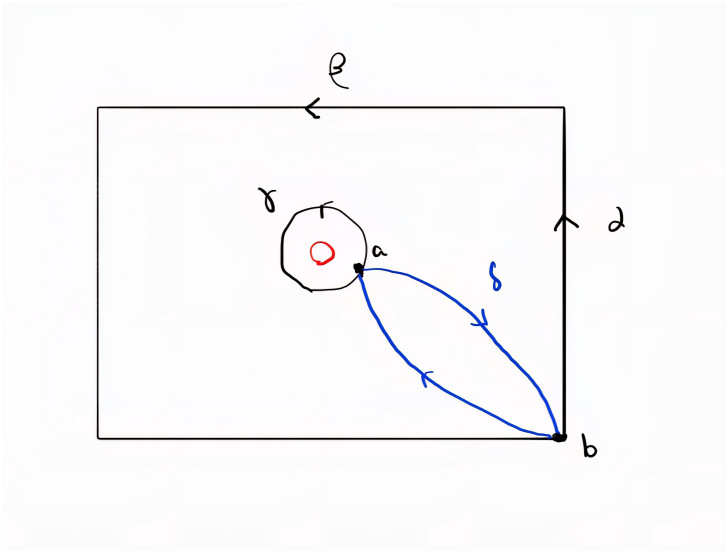


Figure 5: the refered paths

(1) Choosing a point b of the vertex $[0,1]^2$, and let $\alpha\beta\alpha^{-1}\beta^{-1}$ be a path of V based at b, then prove γ and $\alpha\beta\alpha^{-1}\beta^{-1}$ is (free) homotopy in V.

(2) the path connectness ensures to choose a path l in U connecting a and b, and prove that the loop is nullhomotopic in V.

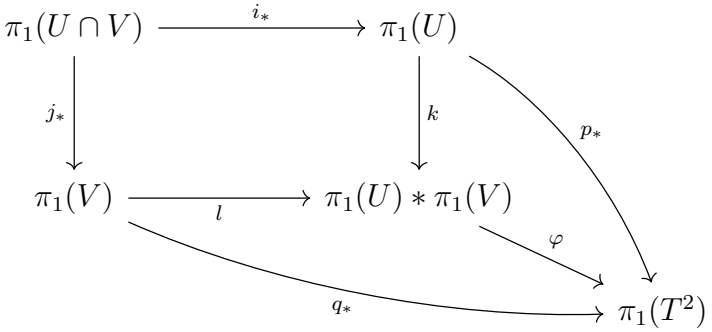
(3) By concatenation, $\delta\alpha\beta\alpha^{-1}\beta^{-1}\bar{\delta}$ gives a loop based at a, and then prove that it is path homotopic to γ in V, and then we can conclude that

$$[\gamma] = [\delta][\alpha\beta\alpha^{-1}\beta^{-1}][\bar{\delta}] = [\alpha\beta\alpha^{-1}\beta^{-1}]$$

By the **universal property of coproduct (free product of groups)**, p_* and q_* induce a unique homomorphism

$$\varphi : \pi_1(V) * \pi_1(U) \to \pi_1(T^2)$$

such that $\varphi \circ k = p_*$ and $\varphi \circ l = q_*$, where k and l are the inclusion maps of $\pi_1(V)$ and $\pi_1(U)$ into the free product respectively, and then the functor π_1 gives the following commutative diagram:



Then we should analysis φ such that we can get the structure of $\pi_1(T^2)$. Indeed, $\pi_1(U \cap V)$ is trival, hence k and p_* just $0 \mapsto 0$, hence we just need to consider j_* and

commute diagram $\varphi \circ l = q_*$. Notice that $\pi_1(V) = \pi_1(U \cap V) = \mathbb{Z} * \mathbb{Z}$, hence $l = id$ and $\varphi = q_*$.

Consider inclusion $q \circ j$, and γ is nullhomotopice in T^2 by embedding, hence $q_* \circ j_*([\gamma]) = [e_a]$, which implies that

$$q_*([\alpha][\beta][\alpha]^{-1}[\beta]^{-1}]) = [e_a]$$

Notice that $q_*([\alpha])$ and $q_*([\beta])$ is not trival and generate $\operatorname{Im} q_*$, then $\operatorname{Im} q_*$ must be abelian. q_* is surjective because any path in T^2 can be countinously deformed to a new path which deoes not pass U, hence by first isomorphism theorem we can conclude that

$$\pi_1(T^2) = \langle q_*(\alpha), q_*(\beta)\rangle / \langle q_*([\alpha][\beta][\alpha]^{-1}[\beta]^{-1}]) = [e_a]\rangle$$

actually it is just $\mathbb{Z}^2$ and can be seen as the pushout of the diagram:

$$\{0\} \longleftarrow \mathbb{Z} \longrightarrow \mathbb{Z} * \mathbb{Z}$$

Theorem 3.13 (Van-Kampen Theorem)**.**
Let X be a topological space, U and V be two open subsets covering X, if U, V and $U \cap V$ is path-connected, then for any $x_0 \in U \cap V$, we have

$$\pi_1(X, x_0) = \pi_1(U, x_0) *_{\pi_1(U\cap V, x_0)} \pi_1(V, x_0)$$

i.e. the natural map

$$\varphi : \pi_1(U, x_0) * \pi_1(V, x_0) \to \pi_1(X, x_0)$$

is surjective with kernel the normal subgroup generated by $i_*([\gamma])j_*([\gamma])^{-1}$, where $[\gamma]$ are the genertors of $\pi_1(U \cap V, x_0)$, and i_* and j_* are the induced map by inclusion

$$i : U \cap V \to U, \quad j : U \cap V \to V$$

Proof. □

Remark 3.4. Consider some speical case of Van-kampen:

1. If $U \cap V$ is just a point, then it is the case of wedge sum, hence

   $$\pi_1(U \vee V, x_0) \cong \pi_1(U, x_0) * \pi_1(V, x_0)$$

   with x_0 being the wedge point.

2. More generally, if $U \cap V$ is simply connected, then the amalgamation product is just the free product, i.e.

   $$\pi_1(X, x_0) \cong \pi_1(U, x_0) * \pi_1(V, x_0)$$

3. If U is simply connected, then the normal subgroup is only generated by $j_*([\gamma])$, hence

   $$\pi_1(X, x_0) \cong \pi_1(V, x_0) / \langle j_*([\gamma])\rangle$$

   In above example, torus T^2 is just the case.

3.4 Covering Spaces

In the previous section, we claim that the fundamental group of $\mathbb{S}^1$ is $\mathbb{Z}$, we will prove it in this section.

Theorem 3.14. $\pi_1(\mathbb{S}^1) \cong \mathbb{Z}$

In this example, what we do is to lift the paths of $\mathbb{S}^1$ to $\mathbb{R}$, and then the paths with the different winding numbers will show the difference and motivates us to conclude the homotopy.

Definition 3.9. *The covering space of a topological space X is a topological space E together with a continnous surjective map $p : E \to X$ satisfying:*
For any $x \in X$, there exists an open neighborhood U of x such that $p^{-1}(U)$ is a disjoint union of open sets in E

$$p^{-1}(U) = \sqcup_{i\in I} U_i$$

and for each U_i, the restriction $p|_{U_i}$ gives a homeomorphism from U_i to U.

Usually we use the notation (E, p) or just a map $p : E \to X$ to denote a covering space, and we call E the **covering space** of X with **covering map** p. For any $x \in X$, the preimage $p^{-1}(\{x\})$ is called the **fiber** of x, and we denote it by F_x.

The covering space have the following topological property:

(a) covering map is open and a local homeomorphism.

(b) If $Y \subset X$ is a subspace, then the preimage $p^{-1}(Y)$ gives a covering space of Y with the covering map $p|_{p^{-1}}(Y)$.

(c) X is Hausdorff if and only if E is also Hausdorff.

(d) If X is path-connected, then all fiber have the same cardianlity, hence we can define the **degree** of covering space as the cardinality of fiber.

(e) If E is compact, (path) connected, then X is also compact, (path) connected.

Similar with $\mathbb{S}^1$, each path in X can be lifted to a path in E, formally in covering space we define a lifting of a map $f : Y \to X$ as a map $\tilde{f} : Y \to E$ such that $p \circ \tilde{f} = f$:

$$\begin{array}{ccc} & & E \\ & \overset{\tilde{f}}{\nearrow} & \downarrow{\scriptstyle p} \\ Y & \xrightarrow[f]{} & X \end{array}$$

Then we can conclude the a general lifting property for covering spaces, they are always be seen as a universal properties.

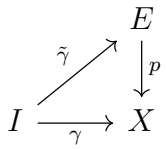
Proposition 3.15 (lifting criterion)**.**
Let $p : E \to X$ be a covering space, Y is a path-connected and locally path-connected topological space, and $p(e_0) = x_0$. Then for any continous map $f : (Y, y_0) \to (X, x_0)$, the **unique lifting** $\tilde{f} : (Y, y_0) \to (E, e_0)$ exists if and only if

$$f_*(\pi_1(Y, y_0)) \subseteq p_*(\pi_1(E, e_0))$$

Proof. □

Remark 3.5. This proposition shows that in the certain case, the lifting map to the covering space is unique up to fibre of the based point, and this criterion covers many case of lifting:

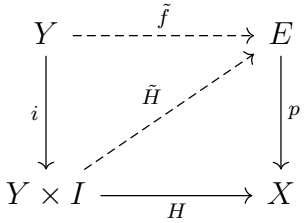
1. In particular, if $Y = [0,1]$ (or generally simply connected space), then the lifting always exists and is unique, and it is the lifting of paths:



   For any path $\gamma : I \to X$ with $\gamma(0) = x_0$, there exists a unique lifting path $\tilde{\gamma} : I \to E$ such that $\tilde{\gamma}(0) = e_0$

2. If $Y = I \times I$, then we can consider the lifting of path-homotopy, For any homotopy $H : I \times I \to X$ with $H(-,0) = f$, if $\tilde{f}$ is a lifting of f, there exists a unique lifting homotopy $\tilde{H} : I \times I \to E$ such that $\tilde{H}(-,0) = \tilde{f}$.

3. If Y is path-connected and locally path-connected, and then we can consider the general lifting of homotopy, For any homotopy $H : Y \times I \to X$ with $H(-,0) = f$, if $\tilde{f}$ is a lifting of f, there exists a unique lifting homotopy $\tilde{H} : Y \times I \to E$ such that $\tilde{H}(-,0) = \tilde{f}$



   The remark 2 is the special case of 3, the proof is not diffcult: set $f(y_0) = x_0$ and fibre $p(e_0) = x_0$, then homotopy $H(y_0, 0) = f(y_0) = x_0$, then there exists a unique lifting $\tilde{H}$ with $\tilde{H}(y_0, 0) = \tilde{f}(y_0) = e_0$, then we need to verify that $\tilde{H(-,0)} = \tilde{f}$ exactly, indeed $\tilde{H(-,0)} : Y \to E$ is a lifting of $H(-,0) = f$, hence by the uniqueness of lifting we can conclude the result.

Corollary 3.16. *Let $p : E \to X$ be a covering space with $p(e_0) = x_0$, and γ and γ' be two paths in X staring at x_0, if they are path homotopic, then their unique liftings at e_0 has the same endpoint and are path homotopic.*

Proof. □

In particular, if γ is a loop based at x_0, then its lifting $\tilde{\gamma}$ must end at the fiber of x_0, it is ensured by $p \circ \tilde{\gamma} = \gamma$. Hence if we consider two homotopic loops based at x_0, then their liftings at e_0 must end at the same fiber point, and hence we can get a natural map

$$\phi_{e_0} : \pi_1(X, x_0) \to F_{x_0}, \quad [\gamma] \mapsto \tilde{\gamma}(1)$$

where $\tilde{\gamma}$ is the unique lifting of γ strating at e_0, and we call it the **lifting correspondece** at e_0. Should pay attention that this map just depends on the point of fiber, and hence if

we fix a homotopy class of loops γ based at x_0, then we can get a natural endomorphism in the sense of **Set**

$$T_{[\gamma]} : F_{x_0} \to F_{x_0}, \quad e_0 \mapsto \phi_{e_0}([\gamma])$$

It motivates us to consider the group action of $\pi_1(X, x_0)$ on the fiber F_{x_0}.

Theorem 3.17. Let $p : E \to X$ be a covering space with $x_0 \in X$, then the right action $\pi_1(X, x_0) \circlearrowleft F_{x_0}$ defined by

$$e_0 \cdot [\gamma] := \phi_{e_0}(\gamma) = \tilde{\gamma}(1)$$

is well-defined, and for any $e_0 \in F_{x_0}$ the orbit of e_0 is exactly the image of lifting correspondece ϕ_{e_0}:

$$\mathrm{Or}(e_0) = \mathrm{Im}(\phi_{e_0})$$

then we have the following conclusions:

- If covering space is path-connected, then the action is transitive.

- If covering space is simply connected, then the action is free and transitive.

Proof. The outline of the proof is here:

(1)By the definition of group action, we just need to verify that for any class $[\gamma]$, endomorphism $T_{[\gamma]}$ is bijective, to prove that we just need to prove

$$T_{[}\gamma]^{-1} = T_{[}\bar{\gamma}]$$

(2) Next, check the definition of orbit, then the proof of transtive is equivalent to the proof that lifting correspondence is surjective.

(3) Finally, the freeness is equivalent to the injectivity of lifting correspondence. □

Lifting correspondence gives a new proof of $\pi_1(\mathbb{S}^1) \cong \mathbb{Z}$:

Proof. $p : \mathbb{R} \to \mathbb{S}^1, \quad t \mapsto e^{2\pi i t}$ is a covering map, and we set the base point $x_0 = 1$, and then we can get the fiber is $F_{x_0} = \mathbb{Z}$. The real line is simplement connected, hence the lifting correspondece at 0 is a bijection

$$\phi : \pi_1(\mathbb{S}^1, 1) \to \mathbb{Z}, \quad [\gamma] \mapsto \tilde{\gamma}(1)$$

and we prove that it is a homomorphism to additive group $\mathbb{Z}$, then we can conclude the isomorphism. Indeed, let γ and δ be two loops based at 1, and their liftings at $0 \in \mathbb{R}$ are $\tilde{\gamma}$ and $\tilde{\delta}$ respectively, and we set $\tilde{\gamma}(1) = m$ and $\tilde{\delta}(1) = n$. Notice the peridoic

$$p(x + n) = x$$

hence we set a new path of $\mathbb{R}$ by $f(t) = m + \tilde{\delta}(t)$, then $f(0) = \tilde{\gamma}(1)$ allows us to get conactenation $\tilde{\gamma} * f$, and then we can verify that

$$\tilde{\gamma} * f(0) = 0, \quad p(\tilde{\gamma} * f) = \gamma * \delta$$

hence $\tilde{\gamma} * f$ is the unique lifting of $\gamma * \delta$ at 0, and then we have

$$\phi([\gamma * \delta]) = \tilde{\gamma} * f(1) = m + n = \phi([\gamma]) + \phi([\delta])$$

□

The fundamental group of covering space has a deep relation with the fundamental group of base space, and in many case we can find that the fundamental group of covering space is easier to comput.

Proposition 3.18. Let $p : E \to X$ be a covering space with $p(e_0) = x_0$, then the covering map p induces an **injective** homomorphism

$$p_* : \pi_1(E, e_0) \to \pi_1(X, x_0), \quad [\gamma] \mapsto [p \circ \gamma]$$

Proof. Let $\delta = p \circ \gamma$ for some loop γ based at e_0, then $\delta(1) = \delta(0) = x_0$ shows that γ is the lifting of δ. And we set $\delta' = p \circ \gamma$, if $[\delta] = [\delta']$ in $\pi_1(X, x_0)$, then there exists a path homotopy $\delta \simeq_H \delta'$, notice that $H(-, 0) = \delta$ has a lifting γ and $H(-, 1) = \delta'$ has a lifting γ', so there exists a unique lifting homotopy $\tilde{H}$ such that $\tilde{H}(-, 0) = \gamma$ and $\tilde{H}(-, 1) = \gamma'$, hence $[\gamma] = [\gamma']$ in $\pi_1(E, e_0)$, which implies the injectivity of p_*. □

we can define a group action

4 Some Results by Algebraic Topology

4.1 The fundamental theorem of algebra

It is a famous theorem in algebra, but the proof can appear in many branches of mathematics via some basic ideas. Here a proof by algebraic topology is given here.

“ Tous les théorèmes fondamentaux de l’algèbre reposent sur la compacité. ”

— Jean-Pierre Serre, *«Cours d’arithmétique»*

Theorem 4.1. A polynomial equation

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

with complex coefficients has at least one root in $\mathbb{C}$.

Proof. The outline of the proof is as follows:
(1) Prove that the continous map $f : \mathbb{S}^1 \to \mathbb{C} - \{0\}$ defined by $f(z) = z^n$ is not nullhomotopic.
(2) Prove the special case of the theorem when $|a_{n-1}| + \cdots + |a_1| + |a_0| < 1$

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

(3) Prove the general case by (2).

Here are the details:
(1) Notice that $\mathbb{S}^1$ can be embedded in $\mathbb{C} - \{0\}$ by the inclusion map i, and similarly we can define a map $g : \mathbb{S}^1 \to \mathbb{S}^1$ by $g(z) = z^n$, then we can write f as the composition

$$f = i \circ g$$

then by the functoriality of fundamental group, we have

$$f_* = i_* \circ g_*$$

notice that i_* is injective since $\mathbb{S}^1$ can be seen as a retract of $\mathbb{C}-\{0\}$, and for the generated loop $[a]$ of $\pi_1(\mathbb{S}^1)$, we can caclulate that $g_*([a]) = [g(a)] = [a^n]$, hence we can explicit write the formula of $g_* : \mathbb{Z} \to \mathbb{Z}$ by $x \mapsto nx$, clearly it is injective.

(2) We prove the special case by contradiction, suppose that the polynomial has no root in $\mathbb{C}$, hence we can define a continous map $p : D \to \mathbb{C} - \{0\}$ by

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0$$

where D is the unit disk in $\mathbb{C}$, similarly if we restrict p to the boundary $\mathbb{S}^1$, i.e. we define $h = p|_{\mathbb{S}^1}$, then by functoriality

$$h_* = p_* \circ j_*$$

where j is the inclusion from $\mathbb{S}^1$ to D. Notice that j_* is the trivial map since $\pi_1(D)$ is trival, hence h_* is also the trivial map, i.e. h is nullhomotopic.

Howerver, in (1) we has shown that f is not nullhomotopic, then a homotopy from f and h can lead to the contradiction, we define $H : \mathbb{S}^1 \times I \to \mathbb{C} - \{0\}$ by

$$H(z,t) = z^n + t(a_{n-1}z^{n-1} + \cdots + a_1 z + a_0)$$

then clearly $H(-,0) = f$ and $H(-,1) = h$.

(3) For the general case, we make a substitute $z = cy$ for some $c > 0$, then

$$y^n + \frac{a_{n-1}}{c}y^{n-1} + \cdots + \frac{a_1}{c^{n-1}}y + \frac{a_0}{c^n} = 0$$

if c is proper, i.e.

$$|\frac{a_{n-1}}{c}| + |\frac{a_{n-2}}{c^2}| + \cdots + |\frac{a_1}{c^{n-1}}| + |\frac{a_0}{c^n}| < 1$$

then we can conclude the existence of root, and we can make a estimate to find that

$$c > \frac{\max\{|a_{n-1}|, \ldots, |a_1|, |a_0|\}}{n}$$

□

4.2 Brouwer fixed point theorem

The Brouwer Fixed Point Theorem is a fundamental result in topology asserting that a continuous self-map of a compact convex set in finite-dimensional Euclidean space must have a fixed point. It captures a purely topological phenomenon: a ball (or simplex) cannot be continuously mapped into itself without leaving at least one point fixed.

Theorem 4.2 (Brouwer)**.** Let D^n be the closed unit disc of $\mathbb{R}^n$, then any continuous map $f : D^n \to D^n$ has at least one fixed point, i.e. there exists $x \in D^n$ such that $f(x) = x$.

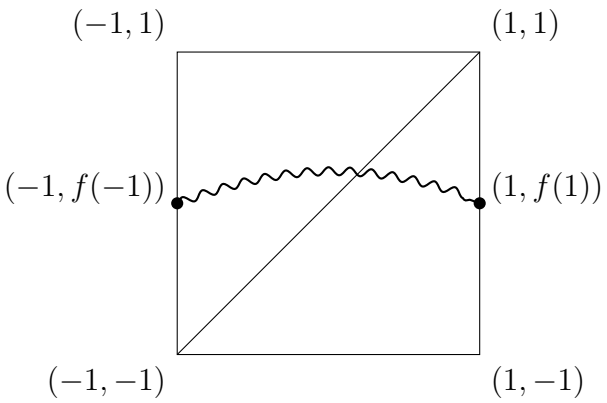
Remark 4.1. The case of $n = 1$ is easy to prove by analytic methods; geometrically $D^1 = [-1, 1]$ implies that we can observe the graph of f. Hence the question is equivalent to finding the intersection of the graph and the line $y = x$. A simple but instructive topological argument is to consider the following two **open** sets:

$$A = \{(x, f(x)) \mid f(x) > x\}, \quad B = \{(x, f(x)) \mid f(x) < x\}$$

If we assume that there is no fixed point, then the graph is the disjoint union of the two open sets

$$G(f) = A \sqcup B$$

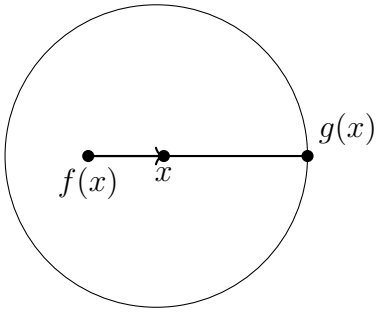
and hence the graph is disconnected, which is a contradiction.



However, the general case is not easy to prove, and the connectness will be a hard condition to use, the dilemma is that we do not have a good property like intermideate value theorem for any n, in another word, the topology of real line is special.

Proof. Here is a proof by homology theory, which is a pure algebraic topology method, we sketch the outline here:
(1) Assume that there exists a continous function f without fixed point when $n \geq 2$, then we can construct a retraction $r : D^n \to \mathbb{S}^{n-1}$ by method of analytic geometry.
(2) We claim that there is no retraction from D^n to $\mathbb{S}^{n-1}$ by homology theory, and prove it.
(3) conclude the result.

Here are the details:
(1) For any $x \in D^n$, we define that $r(x)$ is the intersection of $\mathbb{S}^{n-1}$ and the ray line from $f(x)$ to x:



By vector we can write

$$r(x) = x + \lambda(x - f(x)), \quad \lambda \neq 0 \text{ iff } \|x\| \neq 1$$

by $\|r(x)\| = 1$, we can get a quadratic equation of λ

$$\|x - f(x)\|^2\lambda^2 + 2\langle x, x - f(x)\rangle\lambda + \|x\|^2 - 1 = 0$$

with the positive determination, we can get the exact formula of λ, which is dependent of x continously

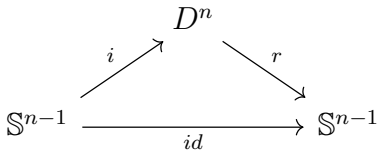
$$\lambda(x) = \frac{-2\langle x, x - f(x)\rangle + 2\sqrt{\langle x, x - f(x)\rangle^2 - \|x - f(x)\|^2(\|x\|^2 - 1)}}{2\|x - f(x)\|^2}$$

in particular, when $x \in \mathbb{S}^1$ i.e. $\|x\| = 1$, we can get

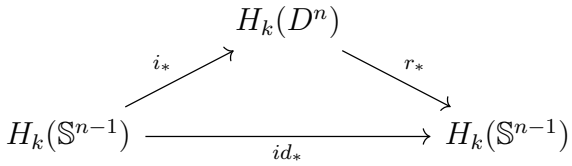
$$\lambda(x) = \frac{-\langle x, f(x)\rangle + \sqrt{\langle x, f(x)\rangle^2}}{\|x - f(x)\|^2} = 0$$

hence $r(x) = x$, which means that r is a retratction from D^n to $\mathbb{S}^{n-1}$.

(2) We prove the claim by homology theory, notice that we have the following commutative diagram:



then by the functoriality of homology, we have the commutative diagram:



notice that $H_{n-1}(D^n) = 0$ and $H_{n-1}(\mathbb{S}^{n-1}) \cong \mathbb{Z}$, so i_* is the trivial map, and $id_* = r_* \circ i_*$ is also the trival map, which leads to a contradiction (functor keeps idenetity map). □

There are other proofs by different methods, but generally the fixed point theorem can be generalized to Banach space by the projective theory.

Theorem 4.3 (Schauder)**.** Let X be a Banach space and let $K \subset X$ be a nonempty, closed, convex set. If $f : K \to K$ is continuous and compact (i.e. $f(K)$ is relatively compact), then f has a fixed point in K.

5 Complex Analysis

5.1 Infinite Product

The form of infinite products is important in complex analysis to study the poles and zeros of functions, here is the definition:

Definition 5.1. Let $(a_n)_{n=1}^{\infty}$ be a sequence of complex numbers. The infinite product $\prod_{n=1}^{\infty} a_n$ is said to converge if there exists $N \in \mathbb{N}$ such that:
(1) $a_n \neq 0$ for all $n \geq N$
(2) the sequence of partial products$(\prod_{k=n}^{N} a_k)_{n \geq N}$ converges to a non-zero limit as $n \to \infty$.
If the limit is zero, we say the product **diverges to zero**. If the limit does not exist, we say the product diverges.

we avoid the case that some terms are zero, because it will make the product zero, which is not interesting; and we notice that if the sequence (a_n) converges to zero, then the product will tend to zero trivally, so we need to avoid this case too. Anyway, the definition here is delicated enough to avoid any trival case.

Proposition 5.1 (The properties of convergence)**.** Let $(a_n)_{n=1}^{\infty}$ be a sequence of complex numbers such that the infinite product $\prod_{n=1}^{\infty} a_n$ converges, then:
(1) $\lim_{n\to\infty} a_n = 1$
(2-LOG) If $a_n \in \mathbb{C} - \mathbb{R}_-$ for sufficent large n, then we have equivalent conidtion for the convergence of the product: the series $\sum_{n=1}^{\infty} \operatorname{Log} a_n$ converges.
(3-CVA) If $\sum_{n\in\mathbb{N}} |a_n| < \infty$, then the product $\prod_{n=1}^{\infty} 1 + a_n$ converges or diverges to zero.

Proof. (1) is immediate form $a_n = (\prod_{k=1}^{n} a_k / \prod_{k=1}^{n-1} a_k)$ for $n \geq 2$.
(2) By the convergence of (1), for sufficent large n, a_n stays near 1, hence we can choose the principal branch such that for sufficent large N

$$\operatorname{Log} \prod_{n\geq N} a_n = \sum_{n \geq N} \operatorname{Log} a_n$$

(3) If $\sum_{n\in\mathbb{N}} |a_n| < \infty$, so a_n converges to zero so we can expend

$$\operatorname{Log}(1 + a_n) = a_n - \frac{a_n^2}{2} + o(a_n^3)$$

Then by (2), the series $\sum_{n\in\mathbb{N}} \operatorname{Log}(1 + a_n)$ converges absolutely, hence the product converges or diverges to zero.

□

With the basic technics, we can study the products of fuctions, that's the core to construct some important functions like gamma function.

Definition 5.1 (Convergence of product of functions)**.**
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of continous function on a open set U
- The product $\prod_{n\in\mathbb{N}} f_n$ converges pointwise to F on U if for each $z \in U$, the product

$\prod_{n\in\mathbb{N}} f_n(z)$ *converges to* $F(z)$.
- The product $\prod_{n\in\mathbb{N}} f_n$ *converges uniformly to* F *on* U *if there exists* $N \in \mathbb{N}$ *such that for all* $n \geq N$, f_n *does not vanish on* U *and the sequence of partial products* $(\prod_{k=n}^{N} f_k)_{n\geq N}$ *converges uniformly to* F *on* U.

Here is the definition following above definition, the zero of the functions in sequence may cause some problems, so there are some difficult in definition. Another clear definition is from **Henri Cartan's book**, here we give as a lemma:

Lemma 5.2. *In above definition, let* $K \subset U$ *as a subset, then the product* $\prod_{n\in\mathbb{N}} f_n$ *converges uniformly on* K *if* f_n *satisfies the following condition:*
(1) $(f_n)_{n\in\mathbb{N}}$ *coverges uniformly to* 1 *on* K.
(2) The series $\sum_{n\in\mathbb{N}} \operatorname{Log} f_n$ *converges normally on* K.

Proof. By (1), for sufficent large N, we have $|f_N - 1| < 1/2$, so for all $n \geq N$, $\operatorname{Log} f_n$ is well defined and f_n does not vanish on K. so for any $z \in K$

$$\operatorname{Log} \prod_{k=n}^{N} |f_k| = \sum_{k=n}^{N} |\operatorname{Log} f_k| \leq \sum_{k=n}^{N} \|\operatorname{Log} f_k\|_K$$

it implies that the sequence of partial products converges uniformlly on K by weierstrass M-test. □

Infinite products is a strong tool to construct function with certain zeros, for example we can constryct a function with zeros at all integers like following:

$$z\prod_{n\geq 1}(1-\frac{z^2}{n^2})$$

It is not difficult to verify that he product converges locally uniformlly to a holomorphic function, and one holomorphic function sharing the propoerties is **sine function**, actually we can prove that they are equally up to a constant factor, which is a famous result called **Euler's sine product formula**. With the following proposition, we can completely prove the formula:

$$\sin(\pi z) = \pi z\prod_{n\geq 1}(1-\frac{z^2}{n^2})$$

The proof can be found in [Stein 2, page 142].

Proposition 5.3. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of holomorphic functions on an open set $\Omega \subset \mathbb{C}$, and suppose that the product $\prod_{n\in\mathbb{N}} f_n$ converges Ω to a function f locally uniformly, then
(1) f is holomorphic on Ω
(2) The zeros and multiplicities of f follows from the sequence:

$$Z(f) = \bigcup_{n\in\mathbb{N}} Z(f_n) \qquad m_f(z) = \sum_{n\in\mathbb{N}} m_{f_n}(z)$$

(3) If f_n does not vanish on Ω for any n, then the series of meromorphic functions $\sum_{n\in\mathbb{N}} f_n'/f_n$ converges locally uniformly on $\Omega - Z(f)$ to f'/f.

Proof. It's the result of locally uniformlly convergence on the inifinite product, notice that the proof of (3) is referred to a idenetity:

$$\frac{(\prod_{k=1}^{n} f_k)'}{\prod_{k=1}^{n} f_k} = \sum_{k=1}^{n} \frac{f_k'}{f_k}$$

which can be proved by induction. Another point is that the proof of (2) needs that the multiplicities of zeros are finite, i.e. for any $a \in \Omega$ we have

$$\#\{n \in \mathbb{N} | f_n(a) = 0\} < \infty$$

Which is ensured by the definition of convergence of infinite product. □

There are two natural question arsing from Euler's formula, one is that if we find another entire function with zeros at all integers, whether the function is same with sine function up to coefficients or anything else? The other is that if there exists a general method to construct entire function with given finite or infinite zeros? The answer is from **Weierstrass's** construction.

Firstly, the zeros of holomorphic is isolated or discrete, so we the given zeros can not have any limit point in $\mathbb{C}$, hence the problem is limited to discrete set of points. **For finite zeros, we can use the fundamental theorem of algebra to construct a polynomial with given zeros**, so the problem is limited to infinite zeros. By Borel-Weierstrass theorem, we can deuduce that the set of zeros must be unbounded.

Another point about zero is that the multiplicity of zeros must be finite. Suppose that f is an non-zero entire function, then by propoerty of analytic function, f can be expended at any point z_0 as a Taylor series

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$$

If z_0 is a zero with inifinite multiplicity, then $f^{(n)}(z_0) = 0$ for all $n \in \mathbb{N}$, hence $f \equiv 0$ in the neighborhood of z_0, which implies $f \equiv 0$ in $\mathbb{C}$ by analytic countiuation, it is absurd. Together with the above statement, if $(a_n)_{n\in\mathbb{N}}$ is the sequence of zeros, then the zeros must satisfy $\lim_{n\to\infty} |a_n| = \infty$.

With the basic analysis of zeros, now we can consrtuct the entire function by given zeros. Notice that we can not simply combine the zeros together by linear terms

$$\prod_{n\in\mathbb{N}} (z - a_n)$$

The product diverges for infinite unbounded zeros, so we can copy the idea of Euler's formula, i.e.

$$\prod_{n\in\mathbb{N}} (1 - \frac{z}{a_n})$$

However, we can not ensure the convergence of the product, for example we take zeros as all positive integers. To ensure the convergence we need to add some extra terms in each factor, that is called **Weierstrass's canonical factors**.

Lemma 5.4. *We define the weierstrass's canonical factor of the degree p as*

$$E_p(z) := \begin{cases} (1-z) & p = 0 \\ (1-z)e^{(z+\frac{z^2}{2}+\cdots+\frac{z^p}{p})} & p \in \mathbb{N} \end{cases}$$

then the factor has the following propoerties:
(1) It is an entire function with a simple zero at $z = 1$ and no other zeros.
(2) It is a function of finite order $p + 1$.
(3) For $|z| \leq r < 1$, we have the estimate

$$|E_p(z) - 1| \leq C_r|z|^{p+1}$$

for some constant $C_r > 0$. That means the larger the p is, the convergence of $E_p(z)$ to 1 is faster.

Proof. (1) and (2) is clear, we prove (3).

$$\begin{aligned} \log(E_p(z)) &= \log(1-z) + z + \frac{z^2}{2} + \cdots + \frac{z^p}{p} \\ &= -\sum_{n=1}^{\infty} \frac{z^n}{n} + (z + \frac{z^2}{2} + \cdots + \frac{z^p}{p}) \\ &= -\sum_{n=p+1}^{\infty} \frac{z^n}{n} = O(|z|^{p+1}) \end{aligned}$$

the expansion here is ensured by $|z| < 1$, hence

$$E_p(z) = \exp(\log E_p(z)) = 1 + O(|z|^{p+1})$$

by expansion of expoential function $e^z = 1 + z + o(z^2)$, so we can conlude the result. □

So we can response the original question by collecting the above ideas:

Theorem 5.5 (Weierstrass's Factorization Theorem)**.**
Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of complex numbers such that $\lim_{n\to\infty} |a_n| = \infty$, and we make convention that $E_n(\frac{z}{0}) := z$, then the infinite product

$$f(z) = \prod_{n=1}^{\infty} E_n(\frac{z}{a_n})$$

converges locally uniformly in $\mathbb{C}$ to an entire function f whose zeros are precisely the points a_n, with multiplicities by counting. Moreover, if g is any other entire function with the same zeros and multiplicities, then there exists an entire function h such that

$$g(z) = f(z)e^{h(z)}$$

for all $z \in \mathbb{C}$.

Another better theorem is given by **Hadamard**, it shows the growth of the entire function will influence the construction of the function by given zeros.

Theorem 5.6 (Hadamard's Factorization Theorem)**.**
Suppose that f is an entire function of finite order ρ with zeros $(a_n)_{n\in\mathbb{N}}$ (counting multiplicities), then there exists a polynomial P of degree at most $k = \lfloor\rho\rfloor$ such that

$$f(z) = e^{P(z)} \prod_{n=1}^{\infty} E_n(\frac{z}{a_n})$$

5.2 Mellin transformation for analytic continuation

The Mellin transformation is a integral transformation which is useful in complex analysis to study the analytic continuation of functions. For example, the gamma function is well-defined for $\Re(s) > 0$ by the integral

$$\Gamma(s) = \int_0^{\infty} e^{-t} t^{s-1} dt$$

Moreover, zeta function is well-defined for $\Re(s) > 1$ by the integral **(Ex15, Chapter6, Stein 2, a easy exercise)**

$$\zeta(s) = \frac{1}{\Gamma(s)} \int_0^{\infty} \frac{t^{s-1}}{e^t - 1} dt$$

and similarly Beta function, a function has a deep relationship with Gamma function, is well-defined for $\Re(a), \Re(b) > 0$ by the integral **(Ex7, chapter 6, Stein 2)**

$$B(a,b) = \int_0^{\infty} \frac{t^{a-1}}{(1+t)^{a+b}} dt$$

Hence we can pay attention to the form of the integration, and then we can a form of integral transformation:

$$\int_0^{\infty} f(t) t^{z-1} dt$$

which is called **Mellin transformation** of the function f, and we usually denote it by $\mathcal{M}[f; z]$.

5.3 Gamma function

As we have mentioned above, the gamma function is defined by two equivalent forms, this section is to study the specific properties of gamma function.

Definition 5.2. Let γ be the Euler constant, $z \in \mathbb{C} - \mathbb{Z}_-$

-infinite product definition:

$$\Gamma(z) = \frac{1}{z e^{\gamma z} \prod_{n\geq 1} (1 + \frac{z}{n}) e^{-z/n}}$$

-integral definition:

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

as the mellin transformation of $f(t) = e^{-t}$.

The Gamma function is a **meromorphic** function on $\mathbb{C}$ with poles at all $\mathbb{Z}_- = \{0, -1, -2, \dots\}$, and its residues are

$$\operatorname{Res}(\Gamma, -n) = \frac{(-1)^n}{n!}, \quad -n \in \mathbb{Z}_-$$

The existence of poles is clear by the infinite product, but the calculation of residues is not clear, it depends on the functional equation to give a simple proof.

Proposition 5.7. For any $z \in \mathbb{C} - \mathbb{Z}_-$

$$\Gamma(z+1) = z\Gamma(z)$$

In particular, for any $n \in \mathbb{N}$, $\Gamma(n+1) = n!$.

Proof. □

Remark 5.1. With the formula, we can verify the residues at poles easily. For any $-n \in \mathbb{Z}_-$ and $\Re(z) > -n - 1$

$$\Gamma(z)(z+n) = \frac{\Gamma(z+n+1)}{z(z+1)\cdots(z+n-1)}$$

by the functional equation, hence calculate the limit of right term at $-n$, we can get the residue at $-n$.

Guass integration is one of the most famous integration

$$I = \int_0^\infty e^{-x^2} dx$$

we can set $x^2 = t$, then subsitute

$$I = \frac{1}{2}\int_0^\infty t^{-1/2} e^{-t} dt = \frac{\Gamma(1/2)}{2}$$

hence the gamma function can give the integration result, which motivates the following functional equation:

Proposition 5.8. For any $z \in \mathbb{C} - \mathbb{Z}_-$

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$$

In particular, $\Gamma(1/2) = \sqrt{\pi}$.

Proof. □

Proposition 5.9 (Legendre's Formula)**.** For any $z \notin \mathbb{Z}_{-1} \cap (\mathbb{Z}_{-1} - \frac{1}{2})$

$$\Gamma(z)\Gamma(z+\frac{1}{2}) = 2^{1-2z}\sqrt{\pi}\Gamma(2z)$$

In particular for any $n \in \mathbb{N}$

$$\Gamma(n+\frac{1}{2}) = \frac{(2n)!}{4^n n!}\sqrt{\pi}$$

Proof. □

In many case, we need to study the asymptotic behavior of gamma function when $z \to \infty$, and gamma function has a deep relation with factoria, a useful equivalent formula help us to do that:

Proposition 5.10 (Euler-Gauss's Formula)**.** For any $z \in \mathbb{C} - \mathbb{Z}_-$

$$\Gamma(z) = \lim_{n\to\infty} \frac{n!n^z}{z(z+1)(z+2)\cdots(z+n)}$$

In particular, we have the asymptotic formula:

$$\Gamma(z+n+1) \sim_{n\to\infty} n!n^z$$

5.4 Zeta function

6 Commutative Algebra

6.1 Modules

Definition 6.1. *Let R be a ring, an R-module M is an abelian group $(M, +)$ together with an operation of R on M:*

$$\cdot : R \times M \to M$$

satisfying the following axioms for all $r, s \in R$ and $m, n \in M$:
(1) $r \cdot (m + n) = r \cdot m + r \cdot n$
(2) $(r + s) \cdot m = r \cdot m + s \cdot m$
(3) $(rs) \cdot m = r \cdot (s \cdot m)$
(4) $1_R \cdot m = m$

- A R-module induces a natural ring homomorphism

  $$\phi : R \to \mathrm{End}(M), \quad r \mapsto \phi_r$$

  with $\phi_r(m) = r \cdot m$. Here axiom (1) ensures ϕ is well-defined (i.e. ϕ_r is an endomorphism of the group), axiom (2)(3)(4) ensures ϕ is a ring homomorphism.

  Conversely, any ring homomorphism $\phi : R \to \mathrm{End}(M)$ induces a R-module structure on M by defining

  $$r \cdot m := \phi_r(m)$$

  Hence we can conclude a correspondence:

  $$\{R\text{-module structures on } M\} \leftrightarrow \{\text{Ring homomorphisms } R \to \mathrm{End}(M)\}$$

- Similarly we can define the morphism between modules, if M, N are two R-modules, then a group homomorphism $f : M \to N$ is a **R-module homomorphism (or R-linear)** if two condition s are satisfied:
  (1) $f(m_1 + m_2) = f(m_1) + f(m_2)$ for all $m_1, m_2 \in M$
  (2) $f(r \cdot m) = r \cdot f(m)$ for all $r \in R$ and $m \in M$

Example 6.1. Here are some important example of modules:
(1) Any vector space can be seen as a k-module, here k is a field.

(2) Any abelian group can be seen as a $\mathbb{Z}$-module.

(3) Any ideal I of a ring R can be seen as a R-module.

(4) Any $k[x]$-module can be seen as a pair (V, T) where V is a k-vector space and $T : V \to V$ is a linear transformation. The module structure is given by

$$(\sum_{i=0}^{n} a_i x^i) \cdot v := \sum_{i=0}^{n} a_i T^i(v)$$

In the another view, it refers a morphism of polynomial rings:

$$\phi : k[x] \to \mathrm{End}(V), \quad x \mapsto T$$

Another important example is algebra, it always appears in the theory of field extension.

Definition 6.2. *Let R and A be two rings with a ring homomorphism $\varphi : R \to A$, then we say (A, φ) is a R-algebra, and we can define a R-module structure on A by*

$$r \cdot a := \varphi(r)a$$

Remark 6.1. In particular, an R-module is not necessary a R-algebra, because the multiplication in A may not be defined in M. For example, an ideal $I \subset R$ is a R-module, but it is not a R-algebra unless $I = R$.
There are some properties about algebras:

(1) If M is a A-module with (A, φ) a R-algebra, then M is also a R-module. It is clearl by

$$R \longrightarrow A \longrightarrow \mathrm{End}(M)$$

(2) If M and N are two A-modules with (A, φ) a R-algebra, then any A-module homomorphism $f : M \to N$ is also a R-module homomorphism, i.e.

$$\mathrm{Hom}_A(M, N) \subset \mathrm{Hom}_R(M, N)$$

In particular, we can take equality if φ is surjective.

Similar with vector spaces, we can define the submodule annd quotient module as following:

Definition 6.3. *Let M be a R-module, a subset $N \subset M$.*

- *N is called a submodule of M if N is also a R-module with the induced operations from M. Equivalently, N satisfies the following conditions:*
  (1) N is an additive subgroup of M.
  (2) For any $r \in R$ and any $n \in N$, we have $r \cdot n \in N$.
- *Let N be a submodule of M, then there exists a natural quotient group homomorphism $\pi : M \to M/N$, if we add the condition that π is a R-module homomorphism, then we can naturally get the multiplication on M/N by*

  $$r \cdot \bar{m} := r \cdot \pi(m) = \pi(rm) = r\bar{m}$$

  Then M/N is called the quotient module of M by N.
- *the R-module homomorphism $\pi : M \to M/N$ is called the quotient map or **canonical projection**. It induced a correspondence:*

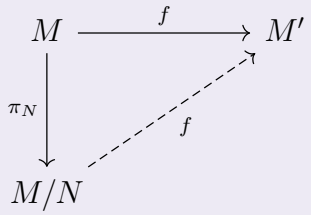
  $$\{\textit{submodules of } M \textit{ containing } N\} \leftrightarrow \{\textit{submodules of } M/N\}$$

Now we conlude the isomorphism theorems for modules, the results are similar with groups and vector spaces, so we just give the statements here without proof.

Theorem 6.1 (UPQ-Module)**.**
Let $f : M \to M'$ be a R-module homomorphism, and N be a submodule of M such that $N \subset \ker f$, then there exists a unique R-module homomorphism

$$\bar{f} : M/N \to f$$

such that the following diagram commutes:



The results of the universal properties are three important isomoprhism for modules:

Theorem 6.2 (Isomorphism Theorems for Modules)**.**
Let M be a R-module, and let N, P be two submodules of M, then:
(1) (First Isomorphism Theorem) If $f : M \to N$ is a R-module homomorphism, then we have the isomorphism:

$$M/\ker f \cong f(M)$$

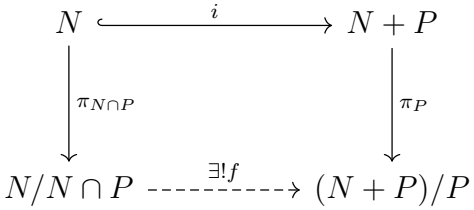
(2) (Second Isomorphism Theorem) We have the isomorphism:

$$(N + P)/P \cong N/(N \cap P)$$

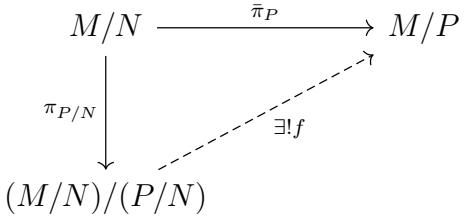
(3) (Third Isomorphism Theorem) If $P \subset N$, then we have the isomorphism:

$$(M/P)/(N/P) \cong M/N$$

Proof. Proof for (1) is directly from UPQ-Module when $N = \ker f$. For (2), we consider diagram:



Here we just need to verify that $\ker(\pi_P \circ i) = N \cap P$ and $pi_P \circ i$ is surjective, then by UPQ-Module we can get the isomorphism. For (3), we consider diagram:



□

where $\bar{\pi}_P$ is induced by the natural quotient map $\pi_P : M \to M/P$. We just need to verify that $\ker \bar{\pi}_P = P/N$ and $\bar{\pi}_P$ is surjective, then by (1) we can get the isomorphism.

Like vector spaces, we also hope to find some elemnts to generate the whole module, here is some vacabulary:

Definition 6.4. *Let M be a R-module, let $(m_i)_{i\in I}$ be a family of elements in M, then:*

*(1) The family is **a generate set** of M if any element of M can be written as a **finite** linear combination of elements in the family, i.e. for any $m \in M$, there exists a finite subset $J \subset I$ and $c_j \in R$, $j \in J$, such that*

$$m = \sum_{j\in J} c_j m_j$$

*(2) The family is **linearly independent** if for any finite subset $J \subset I$ and $c_j \in R$, $j \in J$, the condition $\sum_{j\in J} c_j m_j = 0$ implies that $c_j = 0$ for all $j \in J$.*

*(3) The family is a **basis** of M if it is a generate set and linearly independent.*

*(4) The module M is called **finitely generated** if there exists a finite generate set of M.*

*(5) The module M is called **free** if it has a basis.*

Remark 6.2. Remember that "a free module is a lucky accident", not all modules are free, and even finitely generatd modules are not necessarily free. Here are some examples for clarification:

(a) Any vector space is a free module, but the proof is not trival, it is something about **Zorn's lemma (Axiom of choice.)**. In particular, we have implication

$$\{ \begin{array}{c}\text{finite generated}\\ k\text{-module}\end{array} \} = \{\text{finite dimensional vector space}\} \implies \text{free}$$

(b) An example that a finitely generated module is not free: let $R = \mathbb{Z}$, $M = \mathbb{Z}/2\mathbb{Z}$ as a $\mathbb{Z}$-module, then M is finitely generated by $\{1 + 2\mathbb{Z}\}$, but it is not free since it is not a base by

$$2 \cdot \bar{1} = \bar{2} = \bar{0}$$

here 2 is non-zero.

(c) In a R-module M, we define the **torsion** to be the element $m \in M$ such that

$$r \cdot m = 0 \quad \text{for some } r \in R - \{0\}$$

The torsion elements will obstruct the module to be free, in example (2) we can see that $\bar{1}$ is a torsion element.

(d) An submodule of a finitely generated module is not necessarily finitely generated, for example, let $R = k[x_1, x_2, \cdots]$ be the polynomial ring of infinite variables over a field k, then R is a finitely generated module over itself, but the ideal $I = \langle x_1, x_2, \cdots \rangle$ is not finitely generated.

Some basic properties about finitely generated modules is useful, because we always need to claim whether a module is finitely generated or not.

Proposition 6.3. Let M be a R-module, and $N \subset M$ be a submodule:
(1) If M is finitely generated, then quotient module M/N is also finitely generated.

(2) If N and M/N are finitely generated, then M is also finitely generated.

(3) If M_i is finitely generated R-modules for all $i = 1, 2, \cdots, n$, then the product module $\prod_{i=1}^{n} M_i$ is also finitely generated.

(4) Let M be A-module, and (A, ϕ) be a R-algebra, then if M is finitely generated as a R-module, then M is also finitely generated as a A-module.

Proof. (1) Let $(m_1, m_2, \cdots, m_n)$ be a finite generate set of M, then we claim that $(\bar{m_1}, \bar{m_2}, \cdots, \bar{m_n})$ is a finite generate set of M/N. For any $\bar{m} \in M/N$, there exists $c_i \in R$ such that

$$m = \sum_{i=1}^{n} c_i m_i$$

hence

$$\bar{m} = \sum_{i=1}^{n} c_i \bar{m_i}$$

it implies the result.
(2) Let $(n_1, n_2, \cdots, n_k)$ be a finite generate set of N, and let $(\bar{m_1}, \bar{m_2}, \cdots, \bar{m_l})$ be a finite generate set of M/N. We claim that $(n_1, n_2, \cdots, n_k, m_1, m_2, \cdots, m_l)$ is a finite generate set of M. For any $m \in M$, there exists $c_j \in R$ such that

$$\bar{m} = \sum_{j=1}^{l} c_j \bar{m_j}$$

hence

$$m - \sum_{j=1}^{l} c_j m_j \in N$$

so there exists $d_i \in R$ such that

$$m - \sum_{j=1}^{l} c_j m_j = \sum_{i=1}^{k} d_i n_i$$

it implies that

$$m = \sum_{i=1}^{k} d_i n_i + \sum_{j=1}^{l} c_j m_j$$

hence the result.

(3) Let $(m_{i1}, m_{i2}, \cdots, m_{in_i})$ be a finite generate set of M_i for each $i = 1, 2, \cdots, r$, then the finite set (order is $n_1 + \cdots + n_r$)

$$\{(0, \cdots, 0, m_{ij}, 0, \cdots, 0) \mid i = 1, 2, \cdots, r; j = 1, 2, \cdots, n_i\}$$

is a generate set of $\prod_{i=1}^{r} M_i$.

(4) Let $(m_1, m_2, \cdots, m_n)$ be a finite generate set of M as a R-module, then for any $m \in M$, there exists $c_i \in R$ such that

$$m = \sum_{i=1}^{n} c_i \cdot m_i = \sum_{i=1}^{n} \phi(c_i) m_i$$

□

The statement (2) can be generalized to a exact sequence as following:

Corollary 6.4. *Let $0 \to N \to M \to P \to 0$ be an exact sequence of R-modules, then*

$$\{N, P\} \textit{ finitely generated} \implies M \textit{ finitely generated}$$

However, we can not say more about the implication because the submodule of a finitely generated module is not necessarily finitely generated. We should make some restriction on the module such that some good properties can be ensured, that is called **Noetherian module**.

Definition 6.5. *A R-module M is called a **Noetherian module** if it satisfies the following equivalent conditions:*
(1) Any submodule of M is finitely generated.
(2) Any ascending chain of submodules of M

$$N_1 \subset N_2 \subset N_3 \subset \cdots$$

stabilizes, i.e. there exists $k \in \mathbb{N}$ such that for all $n \geq k$, $N_n = N_k$.
(3) Any non-empty set of submodules of M has a maximal element with respect to inclusion.

The equivalent of the conditions need a proof here for clarification.

Proof. □

Immediate with the definition, we can get some properties of Noetherian modules like proposition 6.3:

- Let $0 \to N \to M \to P \to 0$ be an exact sequence of R-modules, then

  $$\{N, P\} \text{ Noetherian} \iff M \text{ Noetherian}$$

  in particular, let N be a submodule of M, then

  $$M \text{ Noetherian} \iff \{N, M/N\} \text{ Noetherian}$$

- Let M_i be Noetherian R-modules for all $i = 1, 2, \cdots, n$, then the product module $\prod_{i=1}^{n} M_i$ is also Noetherian.

- Let M be a A-module, and (A, ϕ) be a R-algebra, then M is Noetherian as a A-module if it is Noetherian as a R-module.

Noetherian condition is very important in commutative algebra, because it ensures that many bad situations will not happen. For example, in a locally ring (R, m), we only have one chain of ideals

$$\cdots I_2 \subset I_1 \subset m$$

so which ensures the maximal ideal m is finitely generated, and other ideals will be generated by part of the generators of m.

Definition 6.1. Let R be a commutative ring, R is called a **Noetherian ring** if it is Noetherian as a R-module, i.e. it satisfies one of the following equivalent conditions:
(1) Any ideal of R is finitely generated.
(2) A.C.C condition on ideals.
(3) Any non-empty set of ideals has a maximal element with respect to inclusion.

Remark 6.3. Many rings we meet are Noetherian rings:

(a) Any field k is a Noetherian ring, because the only ideals are $\{0\}$ and k.

(b) Any principal ideal domain (PID) is a Noetherian ring, because any ideal is generated by a single element.

(c) The extension of $\mathbb{Z}$ is a Noetherian ring, for example $\mathbb{Z}[\sqrt{2}]$ and $\mathbb{Z}[i]$. The result can be generalized to the following proposition.

Proposition 6.5. Let R be a Noetherian ring, and M is a finitely generated R-algebra, then M is also a Noetherian ring.

Proof. Let $\{m_1, \ldots, m_n\}$ be a set of generators of M as a R-module, then define $R^n := \prod_{i=1}^n R$ to be the product R-module. It is clear that R^n is a Noetherian module since R is a Noetherian module as a R-module, then we can define a natural surjection of R-modules by

$$\varphi : R^n \to M, \quad (a_1, ..., a_n) \mapsto \sum_{i=1}^{n} a_i m_i$$

It is actually a R-module homomorphism, so by the first isomorphism can conclude that $A^n / \ker \varphi \cong M$, then the quotient module of the Noetherian module R^n is also Noetherian, hence M is a Noetherian ring. □

The idea of the proof is important, In the proof we try to construct a module isomorphic to M, so we define a natural object R^n and a natural morphism φ to M, we can find that the propoerty of R^n is so nice that we can transfer the property to M by the isomorphism theorem, i.e. R^n is a free object.

The statement without the finiteness condition is not correct, for example we consider rational number $\mathbb{Q}$, it is not finitely generated as a $\mathbb{Z}$-module, and it is not a noetheriam ring since we can find a chain of ideals:

$$\mathbb{Z} \subset \frac{1}{2}\mathbb{Z} \subset \cdots \subset \frac{1}{2^n}\mathbb{Z} \subset \cdots$$

it will not stop, so $\mathbb{Q}$ is not a Noetherian ring.

6.2 direct sum and product, free module

With the basic knowledge of modules, we can define the the category of modules

$$Mod_R \left| \begin{array}{l} \text{objects: } R\text{-modules} \\ \text{morphisms: } R\text{-module homomorphisms} \end{array} \right.$$

We define the product of modules as following:

Definition 6.6. *Let $(M_i)_{i\in I}$ be a family of R-modules, then the* ***product module*** *is defined by*

$$\prod_{i\in I} M_i := \{(m_i)_{i\in I} \mid m_i \in M_i\}$$

with the operations defined by

$$(m_i)_{i\in I} + (n_i)_{i\in I} := (m_i + n_i)_{i\in I}$$

and

$$r \cdot (m_i)_{i\in I} := (r \cdot m_i)_{i\in I}$$

for any $r \in R$.

There exists a natural projection (a R-module homomorphism)

$$\pi : \prod_{i\in I} M_i \to M_i, \quad \pi((m_i)_{i\in I}) = m_j \text{ for some fixed } j \in I$$

We can conclude that the product of modules satisfies the universal property of product in category $\mathbf{Mod_R}$.

Proposition 6.6. Let $f : N \to \prod_{i\in I} M_i$ be a R-module homomorphism, then there exists a unique R-module homomorphism $\tilde{f} : N \to M_i$ such that $\pi_i \circ f = \tilde{f}$ for all $i \in I$. Equivalently, we have a natural isomoprhism

$$\mathrm{Hom}_R(N, \prod_{i\in I} M_i) \cong \prod_{i\in I} \mathrm{Hom}_R(N, M_i)$$

Proof. □

Moreover, we can define the dual sturcture of product, that is the direct sum of modules.

Definition 6.7. *Let $(M_i)_{i\in I}$ be a family of R-modules, then the* ***direct sum module*** *is defined by*

$$\bigoplus_{i\in I} M_i := \{(m_i)_{i\in I} \mid m_i \in M_i, \text{ and } m_i \neq 0 \text{ for finitely many } i\}$$

with the operations defined by

$$(m_i)_{i\in I} + (n_i)_{i\in I} := (m_i + n_i)_{i\in I}$$

and

$$r \cdot (m_i)_{i\in I} := (r \cdot m_i)_{i\in I}$$

for any $r \in R$.

Remark 6.4.

(1) $\bigoplus_{i\in I} M_i$ is a submodule of $\prod_{i\in I} M_i$. In particular, if the index set is finite, then they are same.

(2) There exists a projection (a R-module homomorphism)

$$\pi : \bigoplus_{i\in I} M_i \to M_i, \quad \pi((m_i)_{i\in I}) = m_j \text{ for some fixed } j \in I$$

However, the projection is not natural in general, because it does not satisfy the universal property of product. For example we consider $I = \mathbb{N}$ and $M_i = M = \mathbb{Z}$, if we decide the application in component by $f_i : \mathbb{Z} \to \mathbb{Z}, \quad z \mapsto z$, then by universal property we can get a application f such that $\pi_i \circ f = f_i$ for all i. But f is not well-defined in direct sum since the image of $1 \in \mathbb{Z}$ is $(1, 1, 1, \cdots)$ which is not in $\bigoplus_{i\in I} M_i$.

(3) There exists a natural injection (a R-module homomorphism)

$$\iota : M_j \to \bigoplus_{i\in I} M_i, \quad i(m) = (0, \cdots, 0, m, 0, \cdots) \text{ for some fixed } j \in I$$

We can conclude that the direct sum of modules satisfies the universal property of **coproduct** in category $\mathbf{Mod_R}$.

Proposition 6.7. Let $f_i : M_i \to N$ be a R-module homomorphism for all $i \in I$, then there exists a unique R-module homomorphism $\tilde{f} : \bigoplus_{i\in I} M_i \to N$ such that $\tilde{f} \circ \iota_i = f_i$ for all $i \in I$. Equivalently, we have a natural isomoprhism

$$\mathrm{Hom}_R(\bigoplus_{i\in I} M_i, N) \cong \prod_{i\in I} \mathrm{Hom}_R(M_i, N)$$

Proof. We can define $\tilde{f}$ by

$$\tilde{f}((m_i)_{i\in I}) = \sum_{i\in I} f_i(m_i)$$

for all $(m_i)_{i\in I} \in \bigoplus_{i\in I} M_i$. It is easy to see that $\tilde{f}$ is well-defined (direct sum supports on finitely many non-zero components) and R-linear. Moreover, we have

$$\tilde{f} \circ \iota_i(m) = \tilde{f}((0, \cdots, 0, m, 0, \cdots)) = f_i(m)$$

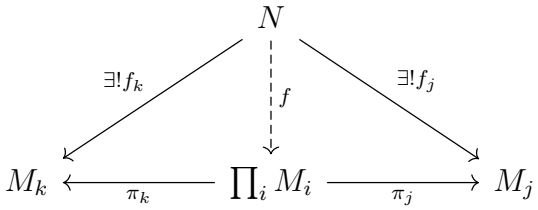
for all $m \in M_i$, which shows $\tilde{f}$ is a desired homomorphism. For the uniqueness, if there exists another solution f, then $(\tilde{f} - f) \circ \iota_i = 0$ for each $i \in I$, hence for any

$(m_i)_{i\in I} \in \bigoplus_{i\in I} M_i$, we have

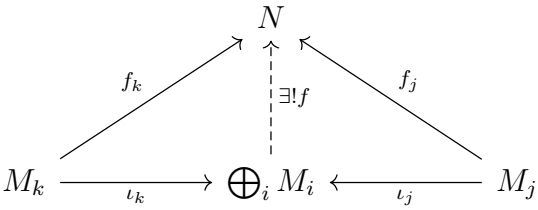
$$(\tilde{f}-f)((m_i)_{i\in I}) = (\tilde{f}-f)(\sum_{i\in I}\iota_i(m_i)) = \sum_{i\in I}(\tilde{f}-f)\circ\iota_i(m_i) = 0$$

it implies $\tilde{f} = f$. □

The properties of direct sum and product can be concludes by the diagram as following:
–product:



and
–direct sum (coproduct):



Sum of modules

Let M be a R-module, and let $E \subset M$ be a subset, then we can define the module generated by E as following:

$$M(E) := \{a_1e_1 + \cdots + a_ne_n \mid n \in \mathbb{N}, a_i \in R, e_i \in M\}$$

in particular, if $E = \{a\}$, then we denote $Ra := M(E)$ to be the **cyclic module** generated by a. Hence $M(E)$ is actually the finite sum of linear combinations of elements in E, so we can define the sum of modules as following:

Definition 6.8. *Let $(M_i)_{i\in I}$ be a family of submodules of a R-module M, then the **sum** of modules is defined as the generated module by the union of the submodules $(M_i)_{i\in I}$, i.e.*

$$\sum_{i\in I} M_i := \{\sum_{i\in J} m_i \mid J \subset I \text{ finite}, m_i \in M_i\}$$

Naturally, we consider the inclusion map $\phi : M_i \hookrightarrow M_i$ for each $i \in I$, then by the universal properties of direct sum, there exists a unique R-module homomorphism

$$\Phi : \bigoplus_{i\in I} M_i \to M, \quad (m_i)_{i\in I} \mapsto \sum_{i\in I} m_i$$

Then we can conclude that the image of Φ is exactly the sum of modules $\sum_{i\in I} M_i$:

$$\sum_{i\in I} M_i = \operatorname{Im}\Phi$$

Definition 6.2. A sum of modules $\sum_{i\in I} M_i$ is called a **(internal) direct sum**, if the homomorphism $\Phi : \bigoplus_{i\in I} M_i \to M$ above is injective. In this case, an isomoprhism is induced:

$$\sum_{i\in I} M_i \cong \bigoplus_{i\in I} M_i$$

Remark 6.5. Equivalently, the family of submodules $(M_i)_{i\in I}$ is a direct sum if for any finite subset $J \subset I$, the following condition holds:

$$\sum_{i\in J} m_i = 0 \implies m_i = 0 \text{ for all } i \in J, m_i \in M_i$$

In particular, if I is **finite**, then the sum $\sum_{i=1}^{n} M_i$ is a direct sum if and only if

$$M_i \cap \sum_{j\neq i} M_j = \{0\} \quad \text{for all } i = 1, 2, \cdots, n$$

The definition of external direct sum is actually the direct sum defined at first, i.e. let $(M_i)_{i\in I}$ be a family of R-modules, then the direct sum $\bigoplus_{i\in I} M_i$ is the **(external) direct sum** of the submodules. If we choose exactly the subomodules of the same modules, then the external direct sum will be reduced to the internal direct sum.

A question raised here is that: for a module M, if we take a submodule N, can we find another submodule P such that $M = N \oplus P$ such that M has a good decomposition? The answer is not always true, and we define that the submodule N is **direct summand** if such P exists.

Example 6.2.
(1) $\mathbb{Z}$ is a module over itself, and let $2\mathbb{Z}$ be the submodule of $\mathbb{Z}$, then we can find that $2\mathbb{Z}$ is not direct summand. (Reason: any submodule of $\mathbb{Z}$ is of the form $n\mathbb{Z}$ for some $n \in \mathbb{N}$, then $2\mathbb{Z} \cap n\mathbb{Z}$=$\{0\}$ iff $n = 0$.)

(2) Any subspace of a **vector space** is direct summand, because any vector space has a basis. If V is a vector space with a basis $\{v_i\}_{i\in I}$, and let $W \subset V$ be a subspace with $W = \text{Span}(\{v_i\}_{i\in J})$ with $J \subset I$, then we can define the complement subspace by $U = \text{Span}(\{v_i\}_{i\in I-J})$.

(3) Simarly with (2), any submodule of a **free module** is direct summand.

An condition can be given here to determine whether a submodule is direct summand or not.

Proposition 6.8. Let M be a R-module, and let $N \subset M$ be a submodule, then the following conditions are equivalent:
(1) N is direct summand of M
(2) M/N is isomorphic to a submodule of M.

(3) There exists a homomorphism (**section**) $s : M/N \to M$ such that $\pi \circ s = \mathrm{id}_{M/N}$, where $\pi : M \to M/N$ is the canonical projection.
(4) There exists a homomorphism (**retraction**) $\rho : M \to N$ such that $\rho(x) = x$ for all $x \in N$.

Proof. (1) $\implies$ (2): Let P be a submodule such that $M = N \oplus P$, then by the second isomorphism theorem we have

$$M/N = (P + M)/N \cong P/(P \cap N) \cong P$$

since direct sum ensures that $P \cap N = \{0\}$.

(2) $\implies$ (3): Let $P \subset M$ be a submodule such that $M/N \cong P$, then the canonical projection $\pi : M \to M/N$ restricts to an isomorphism $\pi|_P : P \to M/N$, hence we can define the section $s : M/N \to M$ by $s^{-1} = (\pi|_P)^{-1}$, and immediately we have $\pi \circ s = \mathrm{id}_{M/N}$.

(3) $\implies$ (1): Let $s : M/N \to M$ be a section, then we can define a homomorphism $\rho : M \to M$ by $\rho = \mathrm{id}_M - s \circ \pi$, then we will show that it is a retraction onto N. For any $m \in M$, we have

$$\pi \circ \rho(m) = \pi(m) - \pi \circ s \circ \pi(m) = \pi(m) - \pi(m) = \bar{0}$$

which implies that $\mathrm{Im}\,\rho \subset N$. Moreover, for any $x \in N$, we have $\rho(x) = n - s \circ \pi(x) = n - s(\bar{0}) = n$ hence we finish the proof.

(4) $\implies$ (1): Let $\rho : M \to N$ be a retraction, then we will prove that $M = N \oplus \ker \rho$. For any $m \in M$, we have

$$m = m - \rho(m) + \rho(m)$$

where m-$\rho(m) \in \ker \rho$ and $\rho(m) \in N$, hence $M = N + \ker \rho$. Moreover, we can caclulate the intersection to prove the dierectness:

$$N \oplus \ker \rho = \{x \in N | \rho(x) = 0\} = \{0\}$$

□

Remark 6.6. It is a motivation to consider the decomposition of the exact sequence. Let

$$0 \longrightarrow N \xrightarrow{f} M \xrightarrow{g} P \longrightarrow 0$$

be a short exact sequence of R-modules, then the following conditions are equivalent:
(1) There exists an isomorphism $M \cong N \oplus P$.
(2) There exists a section $s : P \to M$ such that $g \circ s = \mathrm{id}_P$.
(3) There exists a retraction $r : M \to N$ such that $r \circ f = \mathrm{id}_N$.
If one of the condition holds, then the s.q.e is called **split**. The proof is similar with the previous proposition, we just need to notice that f is injective and g is surjective in a s.e.q.

Free module

For the convience of notation, we define the product of copies of a commutative ring A as following:

$$A^I := \prod_{i\in I} A \quad \text{and} \quad A^{(I)} := \bigoplus_{i\in I} A$$

for any index set I. In particular, if $I = \{1, 2, \cdots, n\}$ is finite, then $A^I = A^{(I)} = A^n$.

Remark 6.7. Recall that a module is free if it has a basis, i.e. a linearly independent generate set.
(1) $A^{(I)}$ is a free A-module with the standard basis as following:

$$e_k := (\delta_{k,i})_{i\in I} = (\dots, 0, \underset{k\text{-th}}{1}, 0, \dots)$$

for all $(a_i)_{i\in I} \in A^{(I)}$, we have

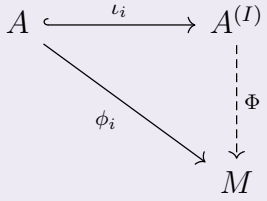
$$(a_i)_{i\in I} = \sum_{i\in I} a_i e_i = \sum_{i\in J} a_i e_i$$

where J is the finite subset of I such that $a_i \neq 0$, so the family is a set of generators. Moreover, if $\sum_{i\in J} a_i e_i = 0$, then $a_i = 0$ for all $i \in J$, so the family is linearly independent.

(2) A^I is not a free A-module if I is infinite. For example, let $I = \mathbb{N}$, then consider the element $(1, 1, 1, \cdots) \in A^I$, it can not be expressed as a finite combination of the standard basis elements above. However, the proof is not trival here, we will do it later.

As we have constructed, $A^{(I)}$ is a **standard model** as a free module, it refers to the following universal property:

Proposition 6.9 (UP-Free module)**.**
Let M be a free A-module with a family of elements $\{m_i\}_{i\in I}$, there exists a unique morphism of A-modules $\Phi : A^{(I)} \to M$ such that $\Phi(e_i) = m_i$ for all $i \in I$.



Equivalently, it induces a natural isomoprhism by $\Phi \mapsto (\phi(e_i))_{i\in I}$

$$\mathrm{Hom}_A(A^{(I)}, M) \cong \prod_{i\in I} M$$

Proof. The proof is similar to the properties of linear map **(the choice of images of**

basis determines a unique linear map). For any $(a_i)_{i\in I} \in A^{(I)}$, we can define

$$\Phi((a_i)_{i\in I}) = \sum_{i\in I} a_i m_i$$

It is easy to verify that Φ is well-defined (direct sum supports on finitely many non-zero components) and A-linear. For the uniqueness, if there exists another morphism $\Psi : A^{(I)} \to M$ such that $\Psi(e_i) = m_i$ for all $i \in I$, then for any $(a_i)_{i\in I} \in A^{(I)}$, we have

$$\Psi((a_i)_{i\in I}) = \Psi(\sum_{i\in I} a_i e_i) = \sum_{i\in I} a_i \Psi(e_i) = \sum_{i\in I} a_i m_i = \Phi((a_i)_{i\in I})$$

hence $\Psi = \Phi$. Hence we can conclude that the two A-linear maps are the same if and only if they have the same images of the basis elements, which implies the natural isomorphism. □

Hence we can generalize the statement of the generating set and basis by the language of morphisms.

Corollary 6.10. *Let M be an A-module, and let I be an index set, then*
*(1) M is **generated** by $\{m_i\}_{i\in I}$ if and only if the morphism $\Phi : A^{(I)} \to M$ defined by $\Phi(e_i) = m_i$ for all $i \in I$ is surjective.*
*(2) $\{m_i\}_{i\in I}$ is **linearly independent** set of M if and only if the morphism $\Phi : A^{(I)} \to M$ defined by $\Phi(e_i) = m_i$ for all $i \in I$ is an injective.*
*(3) $\{m_i\}_{i\in I}$ is a **basis** of M if and only if the morphism $\Phi : A^{(I)} \to M$ defined by $\Phi(e_i) = m_i$ for all $i \in I$ is an isomorphism.*
*(4) M is a **finitely generated** A-module if and only if there exists a surjective morphism $\Phi : A^n \to M$ for some $n \in \mathbb{N}$.*

If M is a free A-module and finitely generated, then we can define "dimension" of M by proving the following statement:

Lemma 6.9. *Let M be a free A-module, and let $\{m_1, \cdots, m_n\}$ and $\{n_1, \cdots, n_k\}$ be two bases of M, then $n = k$.*

Proof. □

6.3 Polynomial and Series

As we know that a polynomial can be defined as following:

$$\mathbb{Z}[X] = \{a_n x^n + \cdots + a_1 x + a_0 \mid a_i \in \mathbb{Z}\}$$

It is clearly a free $\mathbb{Z}$-module with basis $\{x^n \mid n \in \mathbb{N}\}$, that means $\mathbb{Z}^{(\mathbb{N})} \cong \mathbb{Z}[X]$ as modules. Moreover, we notice that the polynomial is furthermore a ring, so it is a $\mathbb{Z}$-algebra, hence the operations should be extended here to generalize the definition of polynomial ring.

Definition 6.3. Let $A^{(\mathbb{N})}$ be a free $A-module$ with basis $(e_i)_{i\in\mathbb{N}} = (\delta_j, i)_{i\in\mathbb{N}}$, then we can define a multiplication on it by make conservation:

$$e_n \cdot e_m = e_{n+m} \quad \text{for all } n, m \in \mathbb{N}$$

and develop it for any elements

$$(a_i)_{i\in\mathbb{N}} \cdot (b_i)_{i\in\mathbb{N}} = (c_i)_{i\in\mathbb{N}} \quad \text{with} \quad c_i = \sum_{l+k=i} a_l b_k$$

Then $A^{(\mathbb{N})}$ is a A-algebra with ring homomorphism

$$\varphi : A \to A^{(\mathbb{N})}, \quad a \mapsto ae_0 = (a, 0, 0, \ldots)$$

Remark 6.8.

(1) It is easy to verify that the multiplication is well-defined, i.e. the result is still in $A^{(\mathbb{N})}$ since only finitely many a_i and b_i are non-zero.
(2) The multiplicative identity is e_0, since for any $(a_i)_{i\in\mathbb{N}} \in A^{(\mathbb{N})}$, we have

$$e_0 \cdot \sum_{i\in\mathbb{N}} a_i e_i = \sum_{i\in\mathbb{N}} a_i (e_0 \cdot e_i) = \sum_{i\in\mathbb{N}} a_i e_i = (a_i)_{i\in\mathbb{N}}$$

(3) The multiplication is associative, commutative and distributive with respect to addition, which can be verified by direct calculation.
with the work of (1)(2)(3), we can conclude that $A^{(\mathbb{N})}$ is a commutative ring with unity, hence it is a A-algebra.

Then we can construct the polynomial ring as following:

Definition 6.4. Let $A^{(\mathbb{N})}$ be the A-algebra defined above, we identify it as the **polynomial ring** $A[X]$ over A by making the following notation:
- basis: $X^n :\equiv e_n$ for all $n \in \mathbb{N}$ with $1_A = X^0 :\equiv e_0$.
- elements: $f(X) = \sum_{i\in\mathbb{N}} a_i X^i :\equiv (a_i)_{i\in\mathbb{N}}$ for all $(a_i)_{i\in\mathbb{N}} \in A^{(\mathbb{N})}$.
- addition: $\sum_n a_n X^n + \sum_n b_n X^n = \sum_n (a_n + b_n) X^n$.
- multiplication: $(\sum_n a_n X^n) \cdot (\sum_n b_n X^n) = \sum_n c_n X^n$ with $c_n = \sum_{l+k=n} a_l b_k$.

Or equivalently, if we define the polynomial ring as what we usually do, then we can get the isomorphism $A[X] \cong A^{(\mathbb{N})}$ by sending X^n to e_n for all $n \in \mathbb{N}$.

More generally, we can define the polynomial ring with multiple variables, it is similarlt with what we have done just now, we need to extend the the multiplication of free-module on any reasonable index set.

Definition 6.10. *Let $(N, +, 0)$ be an associative monoid with identity, we can define a multiplication on the free A-module $A^{(N)}$ with basis $(e_n)_{n\in N} = (\delta_{k,n})_{n\in N}$ by*

$$e_u \cdot e_v = e_{u+v} \quad \textit{for all } u, v \in N$$

and develop it for any elements

$$(a_n)_{n\in N} \cdot (b_n)_{n\in N} = (c_n)_{n\in N} \quad \textit{with} \quad c_n = \sum_{u+v=n} a_u b_v$$

Then $A^{(N)}$ is a A-algebra with ring homomorphism

$$\varphi : A \to A^{(N)}, \quad a \mapsto ae_0$$

and we call it the ***polynomial ring*** *over A with index monoid N, denoted by $A[N]$.*

In particular, if we take $N = \mathbb{N}$ we can get the polynomial ring in one variable $A[X]$. If we take $N = \mathbb{N}^n$ with the component-wise addition, then we can get the polynomial ring in n variables $A[X_1, \cdots, X_n]$ by identifying the basis element $e_{(k_1,\cdots,k_n)}$ with $X_1^{k_1}X_2^{k_2}\cdots X_n^{k_n}$ for all $(k_1, \cdots, k_n) \in \mathbb{N}^n$, then we can get the algebra of polynomial as following:
- elements:

$$f(X_1, ..., X_n) = \sum_{(k_1,...,k_n)\in\mathbb{N}^n} a_{(k_1,...,k_n)} X_1^{k_1}X_2^{k_2}\cdots X_n^{k_n} = \sum_{\kappa\in\mathbb{N}} a_\kappa X^\kappa$$

where only finitely many a_κ are non-zero.

- addition:

$$(\sum_{\kappa\in\mathbb{N}} a_\kappa X^\kappa) + (\sum_{\kappa\in\mathbb{N}} b_\kappa X^\kappa) = \sum_{\kappa\in\mathbb{N}} (a_\kappa + b_\kappa) X^\kappa$$

- multiplication:

$$(\sum_{\kappa\in\mathbb{N}} a_\kappa X^\kappa) \cdot (\sum_{\kappa\in\mathbb{N}} b_\kappa X^\kappa) = \sum_{\kappa\in\mathbb{N}} c_\kappa X^\kappa \quad \text{with } c_\kappa = \sum_{\lambda+\mu=\kappa} a_\lambda b_\mu$$

where the addition $\lambda + \mu$ is defined by component-wise addition in $\mathbb{N}^n$. The verification here is omitted for brevity, the process is similar to the one-variable case.

Polynomial is a natural object in algebra, in a certainn vacabulary, we say that polynomial ring $A[X]$ is a free commutative A-algebra, it refers to the following universal properties:

Proposition 6.11 (UP-Poly)**.**
For any A-algebra B with elements $b_1, ..., b_n$, there exists a unique morphism of A-algebras $\Phi : A[X_1, ..., X_n] \to B$ such that $\Phi(X_i) = b_i$ for all $i = 1, 2, ..., n$.
Or equivalently, it induces a natural bijection by $\Phi \mapsto (\Phi(X_1), ..., \Phi(X_n))$

$$\mathrm{Hom}_{A\text{-alg}}(A[X_1, ..., X_n], B) \cong B^n$$

The inverse is given by evaluation map $(b_1, \ldots, b_n) \mapsto \mathrm{ev}_{(b_1,...,b_n)}$, where $\mathrm{ev}_{(b_1,...,b_n)} : A[X_1, ..., X_n] \to B$ is defined by $ev_{(b_1,...,b_n)(f)} = f(b_1, \ldots, b_n)$.

Proof. For any $f(X_1, ..., X_n) = \sum_{\kappa\in\mathbb{N}^n} a_\kappa X^\kappa \in A[X_1, ..., X_n]$, we can define

$$\Phi(f) = \sum_{\kappa\in\mathbb{N}^n} a_\kappa b^\kappa \in B$$

where $b^\kappa = b_1^{k_1}b_2^{k_2}\cdots b_n^{k_n}$ for $\kappa = (k_1, ..., k_n) \in \mathbb{N}^n$. It is easy to verify that Φ is well-defined (only finitely many a_κ are non-zero) and a morphism of A-algebras. Moreover, we have

$$\Phi(X_i) = \Phi(e_{(0,...,1,...,0)}) = b_i$$

for all $i = 1, 2, ..., n$, which shows that Φ is a desired morphism. For the uniqueness, if there exists another solution Ψ, then for any $f(X_1, ..., X_n) = \sum_{\kappa\in\mathbb{N}^n} a_\kappa X^\kappa \in A[X_1, ..., X_n]$, we have

$$\Psi(f) = \Psi(\sum_{\kappa\in\mathbb{N}^n} a_\kappa X^\kappa) = \sum_{\kappa\in\mathbb{N}^n} a_\kappa \Psi(X^\kappa) = \sum_{\kappa\in\mathbb{N}^n} a_\kappa b^\kappa = \Phi(f)$$

hence we finish the proof. □

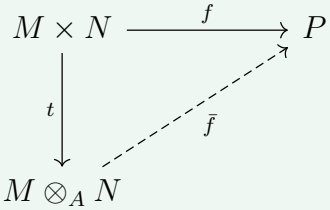
6.4 Projective, Injective Modules

6.5 Tensor product

Definition 6.5. Let M and N be two A-modules, the **tensor product** of M and N over A is an A-module $M \otimes_A N$ together with a bilinear map

$$t : M \times N \to M \otimes_A N, \quad (m, n) \mapsto m \otimes n$$

such that the following Unniversal properties (UPQ-TENSOR) holds:



for any A-module P and any bilinear map $f : M \times N \to P$, there exists a unique A-linear $\bar{f} : M \otimes_A N \to P$ such that $f = \bar{f} \circ t$, i.e. $\bar{f}(a \otimes b) = f(a, b)$.

Remark 6.9. The definition is concluded by some work to give a good construction (1) The existence of tensor product can be constructed by taking the free module $A^{(M\times N)}$ and its submodule R generated by the following elements:

$$\begin{cases} e_{m+m',n} - e_{m,n} - e_{m',n} \\ e_{m,n+n'} - e_{m,n} - e_{m,n'} \\ e_{am,n} - ae_{m,n}, \\ e_{m,an} - ae_{m,n} \end{cases}$$

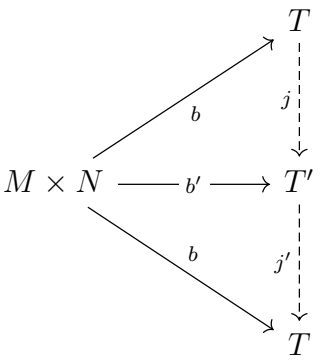
where $m, m' \in M, n, n' \in N, a \in A$, and $(e_{m,n})_{(m,n)\in M\times N}$ is the standard basis of $A^{(M\times N)}$. Then we can define

$$M \otimes_A N := A^{(M\times N)}/R$$

with $m \otimes n := \overline{e_{m,n}}$. then some verification needs here to show that the construction satisfies the universal property.

(2) The tensor product is unique up to isomorphism, i.e. if (T, t) and (T', t') are two tensor products of M and N, then the universal properties induce two morphism

$j : T \to T'$ and $j' : T' \to T$ with the following diagram:



then $j \circ j' : T \to T$ a morpshim satisfying $b = (j \circ j') \circ b$, by the uniqueness of universal property we have $j \circ j' = \mathrm{id}_T$. Similarly, we can get $j' \circ j = \mathrm{id}_{T'}$, hence $T \cong T'$.

(3) The concret construction of tensor product is not important, the nice universal property allows us to treat a multilinear map as a linear map, usually we call $M \otimes_A N$ as «tensor», and $m \otimes n \in M \otimes_A N$ as «tensor element».

Some basic algebraic properties of tensor product can be concluded as following:

Lemma 6.11. *The letter we use are A-modules, then we have the following isomorphisms of A-modules:*

(1) $A \otimes_A M \cong M$ *with the isomoprhism* $a \otimes m \mapsto am$.

(2) $M \otimes_A N \cong N \otimes_A M$ *with the isomorphism* $m \otimes n \mapsto n \otimes m$.

(3) $(M \otimes_A N) \otimes_A P \cong M \otimes_A (N \otimes_A P)$ *with the isomoprhism* $(m \otimes n) \otimes p \mapsto m \otimes (n \otimes p)$.

(4) $M \otimes_A (N \oplus P) \cong (M \otimes_A N) \oplus (M \otimes_A P)$ *with the isomoprhism* $m \otimes (n, p) \mapsto (m \otimes n, m \otimes p)$

Remark 6.10. Pay attention to the abuse of the equality, sometimes we write $=$ instead of $\cong$ to simplify the notation.
(1) the first result can be generalized to $A^n \otimes_A M \cong M^n$.
(2) the last result can be generalized to

$$M \otimes_A (\bigotimes_{i \in I} M_i) \cong \bigotimes_{i \in I} (M \otimes_A M_i)$$

where I is any index set, with the isomorphism we can conclude that tensor product can preserve the free module structure by isomoprhism

$$A^{(I)} \otimes_A M \cong M^{(I)}$$

which covers the first remark.

(3) In general, by above result we can conclude in category: $\mathbf{Mod_A}$ with tensor product $\otimes_A$ forms a monoidal category.

Definition. A monoidal category contains following data:
(1) a category $\mathcal{C}$.
(2) a bifunctor $\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$.
(3) an unit object $I \in \mathcal{C}$ having two natural isomorphisms.

$$I \otimes X \cong X \cong X \otimes I$$

(4) a natural isomorphism (associator)

$$X \otimes Y \otimes Z \cong X \otimes (Y \otimes Z)$$

(5) These ismoprhism satisfy the Maclane axiom and triangle axiom.

We can generalize the universal property to construct the correspondence between multilinear maps and linear maps from the tensor product as following:

Proposition 6.12 (UPQ-MULTI-TENSOR)**.**
Let $M_1, \ldots, M_k$ be A-modules, then for any A-module P, there exists a natural isomorphism

$$\mathrm{Hom}_A(M_1 \otimes_A \cdots \otimes_A M_k, P) \cong \mathrm{Mult}_A(M_1 \times \cdots \times M_k, P)$$

with the bijection $f \mapsto f(- \otimes - \cdots - \otimes -)$

The tensor product of two module is a larger module containg both information of two object in parallel way, hence the morphism of two tensor product is similarly a "product" of morphisms of two modules, which is the tensor product of morphisms:

Proposition 6.13. Let $f : M \to M'$ and $g : N \to N'$ be two morphisms of A-modules, then there exists a unique morphism of A-modules

$$f \otimes g : M \otimes_A N \to M' \otimes_A N', \quad m \otimes n \mapsto f(m) \otimes g(n)$$

6.6 Ideal

Ideal is something like normal subgroups in group theory, it is important to reflect the structure and relationship of rings. Here is the definition:

Definition 6.6. Let R be a commutative ring, a subset $I \subset R$ is called a ideal of R if the following conditions are satisfied:
(1) I is an additive subgroup of R, that is for any $a, b \in I$, we have $a - b \in I$.
(2) For any $r \in R$ and any $a \in I$, we have $ra \in I$ and $ar \in I$.

Remark 6.11. Here is something to add as the properties, proof is not diffcult:

(1) Ideal $I \subset R$ can be seen as an R-submodule of R, roughly speaking, it can be seen as a "vector subspace". (The union of ideals is not necassary an ideal)

(2) An ideal generated by a subset $S \subset R$ is defined by

$$\langle S \rangle := \bigcap_{I \text{ ideal}, S \subset I} I$$

it is the smallest ideal containing S, and we can verify that

$$\langle S \rangle = \{ \sum_{\text{finite}} r_i s_i \mid r_i \in R, s_i \in S \}$$

In particular, a **principal ideal** is an ideal of the form $(a) = \langle \{a\} \rangle$, i.e a ideal generated by a single element (or it is a cyclic R-module).

(3) A useful statement: If a unit $r \in I$, then $I = R$.

(4) The operation of ideal: $IJ \subset I \cap J \subset I \subset \langle I \cup J \rangle$

6.7 Noetherian ring

Theorem 6.14 (Hilbert's Basis Theorem)**.** If A is a Noetherian ring, then the polynomial ring $A[X]$ is also Noetherian.

6.8 UFD

In this section we will study the commutative ring with nice factorization properties, which allows us to define arithemtic sturcture as what we have done in $\mathbb{Z}$.

Definition 6.12. *An integral domain A is called* ***unique factorization domain (UFD)*** *if it satisfies the following conditions:*

(Ex): A is a factorization ring: any non-unit element $a \in A - \{0\}$ can be written as a finite product of irreducible elements.

(Un): The factorization is unique up to order and unit factors, i.e. if

$$x = p_1 p_2 \dots p_r = p'_1 p'_2 \dots p'_{r'}$$

Then $r = r'$ and there exists a permutation $\sigma \in S_r$ such that p_i and $p'_{\sigma(i)}$ are ***associate*** *(equal up to a unit) for all $i = 1, 2, \cdots, r$.*

Remark 6.12. Here is some example of UFD:

(1) $\mathbb{Z}$ is a UFD by the fundamental theorem of arithmetic.

(2) $\mathbb{Z}[\sqrt{-5}]$ is not a UFD since

$$6 = (1 - \sqrt{-5}) \times (1 + \sqrt{-5}) = 2 \times 3$$

where $2, 3, 1 \pm \sqrt{-5}$ are irreducible elements. pay attention that many algebraic number rings are not UFDs.

(3) The elementary number theory is based on the division with remainder, with that we can develop the fundamental theorem of arithmetic in $\mathbb{Z}$, similarly we can define the division with remainder in the polynomial ring over a field $K[X]$, which allows us to conclude that $K[X]$ is a UFD. This is an informal proof, we will talk about it later in Euclidean domains.

(4) An interesting example is $\mathcal{O}(\mathbb{C})$, the ring of all entire function on $\mathbb{C}$, in complex analysis we have Euler's formula

$$\sin(\pi z) = \pi z \prod_{n \geq 1} (1 - \frac{z^2}{n^2})$$

where $\sin(\pi z)$ is not a unit but it can write as an infinite product of irreducible elements, hence $\mathcal{O}(\mathbb{C})$ is not a UFD.

Review the example (2) above, we notice here is a difference betweenn prime elemennts annd irreducible elements:

$$2 \mid 6 \nRightarrow 2 \mid 1 + \sqrt{-5}$$

Hence 2 is not a prime elements. However, in $\mathbb{Z}$ all irreducible elements are essentially prime elements, which leads to the proof of FTA according to the properties of prime numbers (Euclid's lemma). This motivates us to redefine UFD in another way:

Proposition 6.15. If A is a FD (factorization domain), then the following conditions are equivalent:
(1) A is a UFD.
(2) Euclid's lemma holds in A: irreducible elements are prime elements.
(3) Gauss's lemma holds in A: $(a \mid bc \quad \wedge \quad gcd(a, b) = 1) \implies (a \mid c)$

Valuation

Valuation is a useful too to study the local properties of commutative rings, in $\mathbb{Z}$ it relects the divisbility of integers by prime numbers, which is a fundamental tool in number theory, we define it on UFD, but the definition can be generalized to FD although some good properties may not hold.

Definition 6.13. *Let A be a UFD, and let $p \in A$ be an irreducible (prime) element, then we define the p-adic valuation a function $v_p : A - \{0\} \to \mathbb{N}$ by*

$$v_p(a) = \max\{n : p^n \mid a\}$$

and we make convention that $v_p(0) = +\infty$.

> *Remark* 6.13. The following statement can be seen as a lemma for the definition, we suppose that A is just a FD here, and p is a prime element:
>
> (1) For any $a \in A - \{0\}$, the set $E = \{n : p^n \mid a\} \subset \mathbb{N}$ is finite, esure the existence of the maximum such that $v_p(a)$ is well-defined.
>
> (2) If $v_p(a) = n$ the maximum of E above, then there exists a unique element $b \in A - \{0\}$ such that $a = p^n b$ and $p \nmid b$.

The valuation has some nice properties as following, pay attention that the proof is independent of the UFD property:

Proposition 6.14. Let v_p be a p-adic valuation on a UFD A, then for any $a, b \in A$, the following conditions hold:
(1) $v_p(ab) = v_p(a) + v_p(b)$
(2) $v_p(a + b) \geq \min\{v_p(a), v_p(b)\}$, with equality if $v_p(a) \neq v_p(b)$

with the help of valuation, we can give some characterizations of elements in UFD, sometimes it rewrites the definitions of some concepts we meet in elementary number theory.

Proposition 6.15. Let A be a UFD

(1) an element $a \in A$ is a **unit** if and only if $v_p(a) = 0$ for all irreducible elements $p \in A$.

(2) an element $a \in A$ is an **irreducible** element if and only if there exists an irreducible element $p \in A$ such that $v_p(a) = 1$ and $v_q(a) = 0$ for all irreducible elements $q \neq p$.

(3) two elements $a, b \in A$ are **associate** if and only if $v_p(a) = v_p(b)$ for all irreducible elements $p \in A$.

(4) two elements $a, b \in A$, the **divisibility** $a \mid b$ if and only if $v_p(a) \leq v_p(b)$ for all irreducible elements $p \in A$.

(5) an element $a \in A$ with factorization $a = p_1 p_2 \dots p_r$ into irreducible elements, then $v_p(a)$ is the number of factors p_i that are associate to p.

(6) Let P be the set of representatives of the associate classes of irreducible elements in A, then for any $a \in A - \{0\}$ there exists a unique unit $u \in A$ such that

$$a = u \prod_{p \in P} p^{v_p(a)}$$

(7) The **greatest common divisor** of two elements $a, b \in A$ is given by

$$\gcd(a, b) = \prod_{p \in P} p^{\min(v_p(a), v_p(b))}$$

(8) The **least common multiple** of two elements $a, b \in A$ is given by

$$\operatorname{lcm}(a, b) = \prod_{p \in P} p^{\max(v_p(a), v_p(b))}$$

Proof. □

Remark 6.14. In a noetherian intger or a FD, none of the above properties may hold true, that is the reason why we restrict our discussion about valuation techinic on UFD. For example we consider the ring $\mathbb{Z}[\sqrt{-5}]$ again, then some counterexample will happenn:

- take $a = 2(1 + \sqrt{-5})$ and $b = 6$, then (4) fails since $v_2(a) = 1$ but $v_2(b) = 2$, however $a \nmid b$.

- take $6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$, then (5) fails since $v_2(6) = 1$ but there is only one factor 2 in the factorization of 6.

- take $a = 6$, then (6) fails so that (7) and (8) also fail, since

$$\prod_{p \in P} p^{v_p(a)} = 2 \times 3 \times (1 + \sqrt{-5}) \times (1 - \sqrt{-5}) = 36 \neq 6$$

As the consequence of the valuation tool, we prove the transfert theorem of UFD, the traditional proof is based on the study of primitive polynomial, and it is a bit complicated.

Theorem 6.16 (Guass's Theorem)**.**
If A is a UFD, then the polynomial ring $A[X]$ is also a UFD.

Proof. □

6.9 Localization

Localization is a technique to extend a commutative ring by inverting some elements, such that we can study the local properties of the original ring. For example, we want to view 2 as a unit in some algebraic ring, it is natural to consider $\mathbb{Z}[\frac{1}{2}]$ such that $2^{-1} = \frac{1}{2}$.

Definition 6.16. *Let A be a ring, a subset $S \subset A - \{0\}$ is called a* ***multiplicative subset*** *if $1 \in S$ and for any $s, t \in S$, we have $st \in S$.*

with the multiplicative subset, we choose the elements we want to invert, then we can construct the localization of the ring as following:

Lemma 6.17. *Let A be a ring and S be a multiplicative subset of A.*

- We define a equivalent realtion $\sim$ on $A \times S$ by

$$(a, s) \sim (a', s') \iff \exists t \in S, t(as' - a's) = 0$$

it induces a quotient set denoted by $S^{-1}A = (A \times S)/\sim$.

(1-addition) we can define an addition on $S^{-1}A$ by

$$\frac{a}{s} + \frac{a'}{s'} := \frac{as' + a's}{ss'}$$

with a additive identity $\frac{0}{1}$.

(2-multiplication) we can define a multiplication on $S^{-1}A$ by

$$\frac{a}{s} \cdot \frac{a'}{s'} := \frac{aa'}{ss'}$$

with a multiplicative identity $\frac{1}{1}$.

Then $S^{-1}A$ is a ring with the operations above, called the ***localization*** *of A at S.*

Proof. □

Remark 6.15. There exists a natural ring homomorphism onto any localization of A:

$$\ell : A \to S^{-1}A, \quad a \mapsto \frac{a}{1}$$

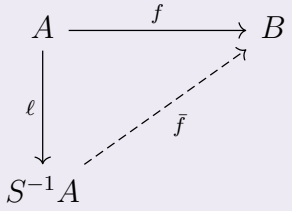
which shows that any localization of A is a A-algebra. And we can verify the kernel of ℓ is given by

$$\ker \ell = \{a \in A \mid \exists s \in S, sa = 0\}$$

it is actually some zero divisors of A "killed" by some elements in S. Hence if A is an **integral domain**, then ℓ is injective, and $\frac{a}{s} = 0$ if and only if $a = 0$.

As a A-algebra object, the localization satisfies the universal properties a little like "analytic continuation" in complex analysis, it refers to extend the morphism to a larger initial object:

Proposition 6.18 (UP-Localization)**.**
Let $S^{-1}A$ be the localization of a ring A, then for any A-algebra (B, f) such that $f(S) \subset B^\times$, then there exist a unique morphism of A-algebras $\bar{f} : S^{-1}A \to B$ such that the $\bar{f} \circ \ell = f$.



Proof. □

Remark 6.16. There are some consequennces of the universal property above:

(1) For a A-module M, if we consider $B = \mathrm{End}(M)$ and f is the structure morphism of M, then UP induces a natural morphsim $\bar{f}$ if any element in S acts as an automorphism on M, i.e. for any $s \in S$, the map $f(s) : M \to M$ is bijective. Hence M can be viewed as a $S^{-1}A$-module by defining

$$\frac{a}{s} \cdot m := f(a) \circ f(s)^{-1}(m)$$

for any $\frac{a}{s} \in S^{-1}A$ and $m \in M$.

(2) Conversely, any $S^{-1}A$-module can be trivally viewed as a A-module since $S^{-1}A$ is a A-algebra. Hence a correspondence about morphisms of module can be concluded here if the condition in (1) hols:

$$\mathrm{Hom}_A(M, N) = \mathrm{Hom}_{S^{-1}A}(M, N)$$

(3) If we set $B = S^{-1}A$ itself, then the UP implies that the uniqueness of the endomorphism of localization, i.e.

$$\mathrm{End}_{A\text{-alg}}(S^{-1}A) = \{\mathrm{id}\}$$

which shows that the localization in a cononcial construction.

Another important respects of localization is about correspondence of ideals, which allows us to identify the ideals in the localization, it is the base of the study of local rings.

Let $S^{-1}A$ be the localization of a ring A at a multiplicative subset S, then there exists two maps

$$\begin{aligned} \{I \subset A \text{ ideal} \mid I \cap S = \emptyset\} &\longleftrightarrow \{J \subset S^{-1}A \text{ ideal}\} \\ I &\longmapsto S^{-1}I := \{\frac{a}{s} \mid a \in I, s \in S\} \\ \ell^{-1}(J) &\longleftarrow J \end{aligned}$$

Proof. □

However, the correspondence above may not be a bijection, so it is not nice to do some identificaion:

$$\ell^{-1}(S^{-1}I) = \{a \in A \mid \frac{a}{1} \in S^{-1}I\} = \{a \in A \mid \exists t \in S, ta \in I\}$$

If the ideal here is prime, then we can get a better result that $\ell^{-1}(S^{-1}I) = I$, hence we can conlude a good corresponence as following:

Proposition 6.19. Let $S^{-1}A$ be the localization of A, and define

$$D(S) := \{p \in \mathrm{Spec}(A) \mid p \cap S = \emptyset\}$$

then there exists a bijection

$$\mathrm{Spec}(S^{-1}A) \to D(S), \quad \mathfrak{p} \mapsto \ell^{-1}(\mathfrak{p})$$

Remark 6.17.

(1) $D(S)$ is not empty if $S^{-1}A \neq 0$ (which map happen if $o \in S$, but we aviod it by definition), the reason is that any ring has at least one maximal ideal by Zorn's lemma, hence by bijection we can at least find one element in $D(S)$.

(2) The statement (1) can be concluded to be a techinic lemma to prove the existence of prime ideals in any ring, which is a fundamental result in commutative algebra:

Lemma. Let S be a multiplicative subset of a ring A, if $I \subsetneq A$ is an ideal such that $I \cap S = \emptyset$, then there exists a prime ideal $p \subset A$ such that $I \subset p$ and $p \cap S = \emptyset$.

The quotient of localization is another basic question as a object of ring. Here we make some remarks to avoid talk too much in the proof of main propoition below, and it gives a clear motivation to the construction:

Let I be a ideal of A such that $I \cap S = \emptyset$ disjoint, then by above correspondence we can know that $S^{-1}I$ is actually an ideal of localization $S^{-1}A$, so it makes sense to define the quotient ring $S^{-1}A/S^{-1}I$.

On the other hand, we consider natural morphism $\pi : A \to A/I$, then $\bar{S} = \pi(S)$ is a multiplicative subset (verify) of A/I if $I \cap S = \emptyset$, then we can do localization on A/I at $\bar{S}$ to get the ring $\bar{S}^{-1}(A/I)$.

So a immediate question is that the two objects above are isomorphic or not? It is essentially that the quoitent of localization is again a localization, and it is the localization of the quotient of initial ring.

Proposition 6.20 (UPQ-Localization)**.** Let $S^{-1}A$ be a localization of a ring A, and I is a ideal of A such that $I \cap S = \emptyset$, then there exists a natural isomorphism of rings

$$\bar{S}^{-1}(A/I) \cong S^{-1}A/S^{-1}I$$

Proof. □

The localization can be extended to modules in a natural way, which allows us to study the local properties of modules over commutative rings.

Definition 6.17. *Let M be a A-module, with $S^{-1}A$ the localization of A, then we define the **localization** of M at S to be a $S^{-1}A$-module*

$$S^{-1}M := M \times S/\sim$$

with the equivalence

$$(m,s) \sim (m',s') \iff \exists t \in S, \quad t \cdot (s'm - sm') = 0_M$$

and similarly we denote the class of (m,s) by $\frac{m}{s}$.

the process of localization is actullay a type of extension of scalar when we see M as a subset of $S^{-1}M$, hence we can apply the tensor product here to give isomorphic construction of localization of modules:

Proposition 6.21 (localization-module)**.**
The inclusion $M \hookrightarrow S^{-1}M$ induces a natural isomorphism of $S^{-1}A$-modules

$$S^{-1}M \cong S^{-1}A \otimes_A M$$

Proof. □

The structure shows the essential difference of localization of modules and rings here, the localization of ring is process of extension, it adds some new elements to the ring such that it can be the inverse of some elements; while the localization of modules is a process of extension of scalar, or by the proposition above it is a tensor product, which means that

localization of modules is not necessarily larger than the initial module!

here is some examples:

Lemma 6.22. *Let M be a A-module, and S be a multiplicative subset of A, then the following conditions are equivalent:*
(1) $S^{-1}M = 0$
(2) for any $m \in M$, there exists $s \in S$ such that $s \cdot m = 0_M$

Proof. (1) $\implies$ (2): for any $m \in M$, we have $\frac{m}{1} = 0$ in $S^{-1}M$, by the definition of equivalence, there exists $s \in S$ such that $s \cdot (1 \cdot m - 1 \cdot 0_M) = 0_M$, hence $s \cdot m = 0_M$.

(2) $\implies$ (1): for any $\frac{m}{s} \in S^{-1}M$, by the condition there exists $t \in S$ such that $t \cdot m = 0_M$, then by the definition of equivalence we have

$$\frac{m}{s} = \frac{t \cdot m}{t \cdot s} = \frac{0_M}{t \cdot s} = 0$$

hence $S^{-1}M = 0$. □

(1) If G is an finite abelian group, then any element has finite order, for example $g \in G$ with order n, then $n \cdot g = g^n = 0$ which implies that

$$\mathbb{Q} \otimes_{\mathbb{Z}} G = S^{-1}G = 0$$

which is a example that the localization of module is even vanishing.

(2) In the previous example, the choice $\mathbb{Q}$ is a little large such that we can not get any interesting result, here we condier $G = \mathbb{Z}/15\mathbb{Z}$ and we condier the localization $\mathbb{Z}_{(3)}$, i.e. take multiplicative subset $S = \mathbb{Z} - 3\mathbb{Z}$, then

$$\mathbb{Z}_{(3)} \otimes_{\mathbb{Z}} G \cong S^{-1}\mathbb{Z}/3\mathbb{Z} \oplus S^{-1}\mathbb{Z}/5\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z}$$

It is easy to the result by the lemma above. The result can be generalized to any finite abelian group G

$$\mathbb{Z}_{(p)} \otimes_{\mathbb{Z}} G \cong G_p$$

where G_p is the p-primary component of G, which is an example that the localization module is smaller.

(3) A A-module M induce a map $\psi : A \to \mathrm{End}(M)$ by

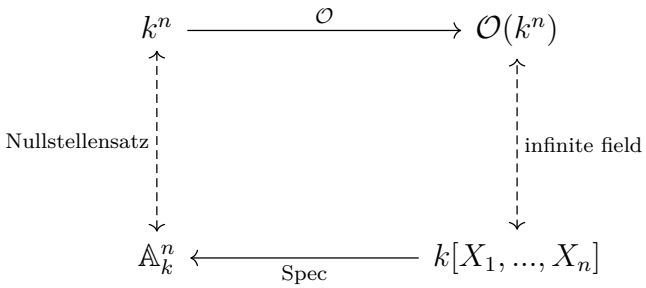
$$\psi_a : (m \mapsto a \cdot m) \in \mathrm{End}(M)$$

then the localization $S^{-1}M = M$ if and only if ψ_s is bijective for any $s \in S$, a simple example is that

$$\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} = \mathbb{Q}$$

7 Affine Algebraic Geometry

The object of this section is to give a basic **geometry-algebra correspondence** between affine algebraic varieties and a polynomial ring over a field, which can be concluded by the following graph:



Firstly we should clarify the isomorphism of polynomial functions and the formal polynomial in certain condition, which allows us to freely use to two objects without confusion.

Definition 7.1. *A formal polynomial is a element of the polynomial ring $k[X_1, \dots, X_n]$ over a field k, while a polynomial function is a function with a formal polynomial expression, i.e. a map*

$$f : k^n \to k, \quad (x_1, ..., x_n) \mapsto \sum_{i=0}^{n} a_i x_i$$

and we denote the set of all polynomial functions on k^n by $\mathcal{O}(k^n)$.

It is easy to verify that $\mathcal{O}(k^n)$ is a k-algebra type fini, hence by the universal property of polynomial, there exists a unique surjective k-linear map

$$\mathrm{ev} : k[X_1, \dots, X_n] \to \mathcal{O}(k^n), \quad X_i \mapsto \pi_i : (x_1, ..., x_n) \mapsto x_i$$

In particular, when k is an **infinite field (verify and find counterexample)**, the map ev is also injective, hence it is an isomorphism of k-algebras, which allows us to abuse the notation of polynomial functions and polynomials.

A classic question in old algebraic geometry is to study the solution of the polynomial equations over some field k, for example the famous Fermat's last theorem is to study the integer solutions of the equation

$$x^n + y^n = z^n$$

if we see it as a polynoimial equation over the field $\mathbb{Q}$. generally, if we have a set of polynoimial equations

$$\begin{cases} f_1(x_1, \dots, x_n) = 0 \\ f_2(x_1, \dots, x_n) = 0 \\ \vdots \\ f_m(x_1, \dots, x_n) = 0 \end{cases} \tag{3}$$

then it induces a subset of k^n, called algebraic subsets:

Definition 7.2. *Let k be a field, and I is a ideal (subset) of the polynomial ring $k[X_1, \ldots, X_n]$, then we define the* ***algebraic subset*** *of k^n defined by I to be*

$$V(I) := \{(a_1, \ldots, a_n) \in k^n \mid f(a_1, \ldots, a_n) = 0, \forall f \in I\}$$

it is also called the ***variety*** *defined by I.*

Remark 7.1. we should remark that the following facts**(verify!)**:

$$V(S) = V((S)), \quad \text{(S) is the ideal generated by S}$$

then the common zeros of $f_1, \ldots, f_m$ is just the algebraic subsets defined by the ideal $(f_1, \ldots, f_m)$.

Moreover, some algebraic facts should be known**(verify)**:

(1) $V(1) = \emptyset$ and $V(0) = k^n$ as two trivial cases.

(2) $V(I) \cup V(J) = V(IJ) = V(I \cap J)$ for any two ideals.

(3) $\cap_k V(I_k) = V(\sum_k I_k)$ for any index.

Conversly, for any subset $V \subset k^n$, we can try to find the largest sets of polynomials vanishing on V, which leads to the definition of **vanishing ideal**:

$$I(V) = \{f \in k[X_1, ..., X_n] \mid f(a) = 0, \forall a \in V\}$$

equivalently, conisder the ideal of polynomial functions, it makes sense to conisder the restriction:

$$I(V) = \{f \in \mathcal{O}(k^n) \mid f|_V = 0\}$$

Remark 7.2. Here are some facts:

(1) $I(k^n) = \{0\}$ and $I(\emptyset) = k[X_1, ..., X_n]$ as two trival case.

(2) vanishing ideal $I(V)$ is a radical ideal, **should know**

$$\text{maximal ideal} \subset \text{prime ideal} \subset \text{radical ideal}$$

(3) **antitone Galois connection:**

(4) $V = V(I(V))$ for any subset $V \subset k^n$:

It is clear that $V \subset V(I(V))$, for avoid confusion of notation in traditional proof, we define a set $\mathcal{J} = \{I \subset K[X_1, \ldots, X_n] \text{ ideal } \mid f|_V = 0, \forall f \in I\}$, then the set is a partial order set under inclusion, it is easy to verify that $I(V)$ is a maximal element in it, i.e. each chain in $\mathcal{J}$ will stop at $V(I(V))$:

$$I_1 \subset I_2 \subset \cdots \subset I(V)$$

then immediately apply $V(-)$, we can get a decrasing chain of algebraic subsets:

$$V \subset V(I(V)) \subset \cdots \subset V(I_2) \subset V(I_1)$$

$V(I(V))$ must be the minimal element in the chain, hence $V(I(V)) = V$.

(5) $I \subset \sqrt{I} \subset I(V(I))$ for any ideal $I \subset k[X_1, ..., X_n]$:

For example we take $I = (X^2)$ in $k[X]$, it is clearly that the common zeros is 0, but the vanising ideal of $\{0\}$ is (X), strictly larger than (X^2).

Here we need a strong condition of field such that we can get a correpsondence betwween algebraic subsets and vanishing ideals, which motivates the Hilbert's zero theorem (Nullstellensatz), we review the original statement here:

Theorem 7.1 (1893)**.** Let $f_1, ..., f_m$ be the polynomials of $\mathcal{O}(\mathbb{C}^n)$, and V is the set of common zeros of $f_1, ..., f_m$ in $\mathbb{C}^n$, and it is non-empty, then for any other polynomial $g \in \mathcal{O}(k^n)$

$$g(x) = 0, \forall x \in V \iff g^k \in (f_1, \ldots, f_m) \text{ for some } k \in \mathbb{N}$$

We will give a general proof later, but firstly we can see how it makes sense. In above we have shown that $\sqrt{I} \subset I(V(I))$, the other side of inclusion is to ask: when we take a polynomial g vanishing on $V(I)$, i.e. the common zeros of some polynomials, what can we say about g? Clearly, $I = (f_1, \ldots, f_m)$ since we know

$$V(\{f_1, \ldots, f_m\}) = V((f_1, \ldots, f_m))$$

then the theorem gives the answer when $k = \mathbb{C}$, it shows that $g \in \sqrt{I}$, hence we have

$$I(V(I)) = \sqrt{I} = I?$$

the first equality holds by the theorem, but the second equality holds under the condition that I is a radical ideal, which shows that the zeros of polynomials has a closed relationship with the radical ideals in certain condition, i.e. algebraic closed fields. Here we give the general statement of zero theorem:

Theorem 7.2 (weak Nullstellensatz)**.** Let k be a field, and M be a maximal ideal of the polynomial ring $k[X_1, \ldots, X_n]$, then the residue field $k[X_1, \ldots, X_n]/M$ is a k-algebra of type fini.
In paticular if k is algebraically closed, then the residue field is isomorphic to k itself, with M of the form

$$M = (X - a_1, ..., X - a_n), \quad (a_1, ..., a_n) \in k^n$$

Proof. □

With the weak nullstellensatz, we can give the correspondence we expect:

Corollary 7.3. *If k is algebraically closed, and then*
(1) For any ideals $I, J \subset K[X_1, \ldots, X_n]$, $V(I) = V(J) \iff \sqrt{I} = \sqrt{J}$.
(2) application $I \mapsto V(I)$ and $V \mapsto I(V)$ induces a bijection

$$\{\textit{radical ideal of } K[X_1, \ldots, X_n]\} \longleftrightarrow \{\textit{algebraic variety of } k^n\}$$

and

$$\{\textit{maximal ideal of } K[X_1, \ldots, X_n]\} \longleftrightarrow \{\textit{one-point variety of} k^n\}$$

Proof. (1) $\Leftarrow$ is clear. Conversely, a radical ideal I can be written as the intersection of some maximal ideals**(verify)**,

$$\sqrt{I} = \bigcap_{I \subset \mathfrak{m}} \mathfrak{m}$$

then by weak nullstellensatz $V(\mathfrak{m}) = \{a\}$ for some point of k^n, hence $I \subset \mathfrak{m}$ if and only if $a \in V(I)$, i.e.

$$\sqrt{I} = \bigcup_{a \in V(I)} \mathfrak{m}_a$$

where notice that $a = (a_1, \ldots, a_n) \mapsto (X - a_1, \ldots, X - a_n) = \mathfrak{m}_a$ gives a bijection by nullstellensatz. Hence $V(I) = V(J)$ gives the same radical ideal.

(2) For any algebraic subset $V \subset k^n$, by definition $I(V)$ is a radical ideal, and $V(I(V)) = V$ holds without any condition. Conversely, if $I = K[X_1, \ldots, X_n]$, then $V(I) = \emptyset$ and $I(V(I)) = I(\emptyset) = I$. If $I \subsetneq K[X_1, \ldots, X_n]$, then there exits at least one maximal ideal $\mathfrak{m}$ such that $I \subset \mathfrak{m}$ by Zorn's lemma, hence $V(I) \neq \emptyset$, hence we can apply (1) to $V(I(V(I))) = V(I)$, then $\sqrt{I(V(I))} = \sqrt{I}$, both are radical ideals, hence we can get the result. □

Remark 7.3. Actually, we can say more about the original zero theorem, If $I = (f_1, \ldots, f_m)$ forms a proper ideal, then there exists there exists a maximal ideal $\mathfrak{m}$ containing it, similarly techincs as above shows that there exists a common zeros of $f_1, \ldots, f_m$ in k^n if k is algebraically closed. Hence the original zero theorem holds without assumption of non-empty common zeros , generally it is called the **strong Nullstellensatz**:

$$V(I) = \emptyset \iff I = A$$

Indeed it reflects the fact about zeros: if a set of polynomials has common zeros if and only if the polynomials forms a proper ideal, i.e. there does not exists some coefficients $c_i \in k$ such that

$$\sum_i c_i f_i = 1$$

A immediate question is that how to describe the prime ideals in terms of algebraic variety, it is more meaningful when consider the Zariski topology, here we suppose that

$I(V)$ is a prime ideal of $K[X_1, \ldots, X_n]$, then for any $f, g \in K[X_1, ..., X_n]$,

$$\begin{aligned} fg \in I(V) &\iff f \in I(V) \text{ or } g \in I(V) \\ &\iff V \subset V(f) \text{ or } V \subset V(g) \\ &\implies V \subset V(f) \cup V(g) \end{aligned}$$

If we want complete above implication, some other condition is needed here.

Definition 7.3. *An algebraic subset $V \subset k^n$ is called **irreducible** if for any algebraic subsets $V_1, V_2 \subset k^n$ such that*

$$V = V_1 \cup V_2$$

then $V = V_1$ or $V = V_2$.

The condition is a little like connectness in topology, hence it is more natural to consider a irreducible space in Zariski topology, here we just finish the correspondence via the condition:

Proposition 7.4. Let k be an algebraically closed field, then there exists a bijection

$$\{\text{prime ideal of } K[X_1, \ldots, X_n]\} \longleftrightarrow \{\text{irreducible algebraic variety of } k^n\}$$

Proof. Let I be a prime ideal of $K[X_1, \ldots, X_n]$, then we need to show that $V(I)$ is irreducible. If $V(J)$ and $V(L)$ are two algebraic subsets such that

$$V(I) = V(J) \cup V(L) = V(JL)$$

then by the correspondence of radical ideals, we have $\sqrt{I} = \sqrt{JL}$, I is prime so it is radical, hence $I = JL$, then by the definition of prime ideal, $J \subset I$ or $L \subset I$, hence $V(I) \subset V(J)$ or $V(I) \subset V(L)$, which shows that $V(I)$ is irreducible.

Let V be an irreducible algebraic subset of k^n, then we need to show that $I(V)$ is a prime ideal. For any $f, g \in K[X_1, \ldots, X_n]$ such that

$$fg \in I(V) \iff V \subset V(fg) = V(f) \cup V(g)$$

then by irreducibility of V, we have $V \subset V(f)$ or $V \subset V(g)$, hence $f \in I(V)$ or $g \in I(V)$, which shows that $I(V)$ is prime. □

So far we have built the basic correspondence between variety and ideals, to say something geomtry, we need to consider the topology of the affine space k^n under the condition of algebraically closed, Review the properties of algebraic varieties in remark 7.1, we can see that the algebraic varieties naturally defined closed sets of k^n, hence it induces a topology on k^n, and we call it the **Zariski topology**. Now we will conclude some data of the topological properties:

(a) It is diffcult to describe the open sets in Zariski topology exactly, but we can define a basis of open sets as following:

$$D(f) := k^n \setminus V(f) = \{x \in k^n \mid f(x) \neq 0\}$$

for any polynomials $f \in K[X_1, ..., X_n]$, it is called the principal open subset defined by f, and **the set of all principal open subsets forms a basis of Zariski topology(verify)**. It is easy to verify some properties of principal open subsets:
(1) $D(0) = \emptyset$ and $D(1) = k^n$.
(2) $D(f) \cap D(g) = D(fg)$ for any polynomials f, g.
(3) $D(f) = \emptyset \iff f$ is nilpotent.

(b) When n=1, the Zariski topology on k is just the cofinite topology: for any finite subset $F = \{c_1, ..., c_n\} \subset k$, we can find that $f(X) = (X - c_1) \cdots (X - c_n)$ vanishes exactly on F, hence $V(f) = F$, which shows that any finite subset is closed, and there does not exist a polynomial vanishing at inifinitely many points in $k[X]$, hence all possible closed sets are

$$\mathcal{F} = \{k, \emptyset, \text{finite subsets}\}$$

which is exactly the cofinite topology, it is a natural example to understand the Zariski topology. **Pay attention** that when $n \geq 2$, the Zariski topology is not the cofinite topology, a example is that

$$V(X_1) = \{(0, a_2, \ldots, a_n) \mid a_i \in k\}$$

it is an infinite closed subset of k^n, hence

$$\text{cofinite topology} < \text{Zariski topology} < \text{Euclidean topology}$$

when "$<$" means "is coarser than".

(c) **The Zariski topology is a T_1-space**, i.e. any single point is closed. Indeed, and point $a \in k^n$ can be seen uniquely as a maximal ideal $\mathfrak{m}_a = (X_1 - a_1, ..., X_n - a_n)$ by nullstellensatz. However, **the Zariski topology is not T_2 (Hausdorff)**, consider any two points $a, b \in k^n$, and two non-empty set principal open subsets $D(f), D(g)$ containing them respectively, then

$$D(f) \cap D(g) = \emptyset \iff D(fg) = \emptyset$$

which implies that fg is nilpotent, then f or g is nilpotent, which contradicts the assumption that $D(f), D(g)$ are non-empty. Hence any two non-empty open subsets of Zariski topology always intersect, which shows that it is not Hausdorff.

(d) **The Zariski topology is quasi-compact**, i.e. any open cover has a finite subcover(easy to verify).

(e) **The Zariski topology is connected**: Suppose that $k^n = U \cup V$ are two non-empty disjoint open subsets, then the open basis allows to choose non-empty open sets $D(f)$ and $D(g)$ such that $D(f) \cap D(g)$, then disjoint condition shows that $D(fg) = \emptyset$, hence fg is nilpotent, then $D(f) = \emptyset$ or $D(g) = \emptyset$, which leads to contradiction.

We can do more with the help of nullstellensatz, we condier a non-emprty varitey $V \subset k^n$, we can consider the quotient ring $K[X_1, \ldots, X_n]/I(V)$, it means we kill all polynomial functions vanishing on V, hence we can view it as the set of polynomial functions on V:

Definition 7.4. *Let $V \subset k^n$ be an algebraic subset, then a function $f : V \to k$ is called a **polynomial function** on V if there exists a polynomial function $\tilde{f} \in \mathcal{O}(k^n)$ such that $f = \tilde{f}|_V$. We denote the set of all polynomial functions on V by $\mathcal{O}(V)$.*

The restriction induces a natural surjective morphism of k-algebras

$$r : \mathcal{O}(k^n) \to \mathcal{O}(V), \quad f \mapsto f|_V$$

with the kernel $\ker r = I(V)$, hence by the isomorphism of quotient we have

$$K[X_1, \ldots, K_n]/I(V) \cong \mathcal{O}(k^n)/I(V) \cong \mathcal{O}(V)$$

and we call $\mathcal{O}(V)$ the **coordinate ring** of the variety V, the following corollary reflects the name of "coordinate", it reflects all the information of the variety:

Proposition 7.5 (UPQ-Coordinate ring)**.**
Let $V \subset k^n$ be an algebraic subset with k algebraically closed, then for any k-algebra A, there exists a natural bijection

$$V \longleftrightarrow \mathrm{Hom}_{k-alg}(\mathcal{O}(V), k)$$

by mapping a point $a \in V$ to the evaluation map ev_a:

$$\mathrm{ev}_a : \mathcal{O}(V) \to k, \quad f \mapsto f(a)$$

Proof. In particular, if $V = k^n$, then $k^n \cong \mathrm{Hom}(\mathcal{O}(k^n), k)$ is just the dual property of polynomial ring. □

Remark 7.4. In particular, if $V = \{a\}$ is a single point, then the coordinate ring $\mathcal{O}(V) \cong k$ itself, and then **the endomorphism of a k-algerba is unique(verify!)**, i.e.

$$\mathrm{Hom}_{k-alg}(A, A) = \{id\}$$

which satisfy the fact here, furthermore we notice that

$$\mathcal{O}(k^n)/I(V) \cong \mathcal{O}(V) \cong k$$

hence $I(V)$ must be a maximal ideal, without the assumption of algebraically closed field, we can not get bijection between points and maximal ideals.

We should notice that the coordinate ring actually reflects the geometric information of variety in the way of algebra, for example we condiser two algebraic subsets

$$V_1 = \{(x, 0) \in k^2 \mid x \in k\} \quad V_2 = \{(x, x, x) \in k^3 \mid x \in k\}$$

clearly that they are all strignt lines in affine spaces with different dimensions, and we can calculate their coordiate rings:

$$\mathcal{O}(V_1) \cong k[X,Y]/(Y) \cong k[X], \quad \mathcal{O}(V_2) \cong k[X,Y,Z]/(X-Y, Y-Z) \cong k[X]$$

hence the coordiante rings reflects the the functoriality of $\mathcal{O}$, from a geometric object to an algebraic object. To complete the functoriality of morphism, we naturally define the morphism between varieties:

Definition 7.5. *Let $V \subset k^n$ and $W \subset k^m$ be two algebraic subsets, then a map $\varphi : V \to W$ is called a **polynomials** if it is a restriction of a polynomials map $k^n \to k^m$*

$$(x_1, \ldots, x_n) \mapsto (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n))$$

where $f_i \in \mathcal{O}(k^n)$.

Indeed such a map is continous under Zariski topology **(verify!)**, which allows us to construct a sub-category of **Top**, called t**he category of affine algebraic varieties over** k:

$$\mathbf{Aff}_k \left| \begin{array}{l} \text{objects: algebraic variety } V \subset k^n \\ \text{morphisms: polynomial functions } f : V \subset k^n \to W \subset k^m \end{array} \right.$$

with this definition, then as a functor $\mathcal{O}$ maps each object to a coordinate ring, exactly a k-algebra of type fini, here we denote it by $\mathbf{Alg}_k^{\text{fini}}$.

On the level of morphisms, each polynoimial function $\varphi : k^n \to k^m$ induces a morphism of k-algebra like dual map:

$$\varphi^* : \mathcal{O}(k^m) \to \mathcal{O}(k^n), \quad f \mapsto f \circ \varphi$$

hence furthermore let $\varphi|V : V \to W$ be the restriction, then it induces a morphism of coordiante rings by quoitent:

$$\begin{array}{ccc} \mathcal{O}(k^m) & \xrightarrow{\varphi^*} & \mathcal{O}(k^n) \\ \Big\downarrow{\scriptstyle \pi_W} & & \Big\downarrow{\scriptstyle \pi_V} \\ \mathcal{O}(W) & \dashrightarrow & \mathcal{O}(V) \end{array}$$

It is easy to check:

$$\begin{aligned} \varphi^*(f) \in I(V) &\iff f \circ \varphi \in I(V) \\ &\iff f(\varphi(a)) = 0, \forall a \in V \\ &\iff f(b) = 0, \forall b \in W \\ &\iff f \in I(W) \end{aligned}$$

hence it induces a unique morphism of k-algebras

$$\varphi^* : \mathcal{O}(W) \to \mathcal{O}(V), \quad f + I(W) \mapsto f \circ \varphi + I(V)$$

Hence we finish the consturction of functor completely

$$\mathcal{O} : \mathbf{Aff}_k \to \mathbf{Alg}_k^{\text{fini}}$$

without the assumption that k is algebraically closed, the functor will not be an equivalence, actually we need another condition to ensure that, which makes us can give a one-side correpsondence completely:

Theorem 7.6. Let k be an algebraically closed field, then the functor $\mathcal{O}$ gives a contravariant equivalence to the reduced k-algebras of type fini:

$$\mathbf{Aff}_k \cong (\mathbf{Alg}_k^{\text{fini, red}})^{\text{op}}$$