Some Algebraic Geometry Notes

peter lundgaard

July 2024

Contents

1 Set & Category Theoretic Definitions

1.1 Set theory

1.1.1 ZF(C)-axioms

We introduce set theory first via the Zermelo-Frankel axioms with an added axiom of choice which will be necessary in some cases. We add to first order predicate logic a relational symbol $\in$. For a pair of objects z, X we define $z \in X := \in(z, X)$.

Axiom. *0.* $\exists \emptyset \forall z (\neg (z \in \emptyset))$

1. $\forall X \forall Y (\forall z (\forall z \in X \iff z \in Y) \Rightarrow X = Y)$.

2. $\forall x \forall y \exists P \forall z (z \in P \iff (z = x \vee z = y))$.

3. $\forall X \exists U \forall z (z \in U \iff \exists w \in X (z \in w))$

Definition 1.1.1. Let X be a set. P a predicate. We say that $P(x)$ *is true for all but finitely many* $x \in X$, if there exists a $Y \subset X$, such that $P(x)$ is true for all $x \in X \setminus y$

1.1.2 Natural Numbers & the Peano Axioms

1.1.3 Von Neumann Ordinals: A Construction of $\mathbb{N}$ (the Natural Numbers)

1.1.4 NGB Set Theory - Classes

1.2 Relations

1.2.1 Functions

1.2.2 Ordering

1.2.3 Equivalence Relations

Definition 1.2.1. Let A be a non-empty set, we define an *equivalence relation* on A to be a subset $\sim$ of $A \times A$ satisfying

1. reflexivity,
$$x \sim x \text{ for every } x \in X$$
2. symmetry,
$$x \sim y \Rightarrow y \sim x \text{ for every } x, y \in X$$

3. transitivity
$$x \sim y \ \wedge y \sim z \Rightarrow x \sim z \text{ for every } x, y, z \in X.$$

Here we define $x \sim y$ to mean $(x, y) \in \sim$. For an $x \in X$ we define the *equivalence class under $\sim$ represented by x* to be the set

$$[x]_\sim := \{y \in X : y \sim x\}.$$

We denote the set of equivalence classes under $\sim$ by $X/\sim$.

Lemma 1.2.2. *Let X be a non-empty set and $\sim$ an equivalence relation on X. Let $x, y \in X$. Then*
$$[x]_\sim = [y]_\sim \iff x \sim y$$

Proof. "$\Leftarrow$": Let $z \in [x]_\sim$, then $z \sim x$ and $z \sim y$, since also $z \in [y]_\sim$. Then $x \sim z$ (by symmetry) and $z \sim y$, implying $x \sim y$ by transitivity.
"$\Leftarrow$": If $x \sim y$, then $x \in [y]_\sim$. By symmetry $y \sim x$, hence $y \in [x]_\sim$. □

Lemma 1.2.3. *Let X be a non-empty set and $\sim$ an equivalence relation on X. The function*

$$\pi : X \to X/\sim$$
$$x \mapsto [x]_\sim$$

is a well-defined surjective function.

Proof. Suppose $x = y$, then $x \sim y$, hence by Lemma 1.2.2 $p(x) = [x]_\sim = [y]_\sim = p(y)$. Let $[x]_\sim \in X/\sim$. Then $\pi(x) = [x]_\sim$, hence π is surjective. □

1.3 Category Theory

1.3.1 Initial Definitions

Definition 1.3.1. A category $\mathcal{C}$ is a pair $(\mathbf{Ob}(\mathcal{C}), \mathbf{Hom}(\mathcal{C}))$ where

1. $\mathbf{Ob}(\mathcal{C})$ denotes a class of *objects*.

2. $\mathbf{Hom}(\mathcal{C})$ denotes a class of *morphisms*.

3. A morphism f in $\mathbf{Hom}(\mathcal{C})$ is a relation between between elements A, B in $\mathbf{Ob}(\mathcal{C})$. We denote it by $f : A \to B$.

4. For objects A,B in $\mathbf{Ob}(\mathcal{C})$ we denote the class of morphisms from A to B by $\mathbf{Hom}(A,B)$.

5. There is binary operation $\circ$ on the class of morphisms called *composition* such that for morphisms $f:B\to C$ and $g:A\to B$ we have that

$$fg := f\circ g : A\to C$$

and

$$(f\circ g)\circ h = f\circ(g\circ h)$$

where $f:C\to D$, $g:B\to C$ and $h:A\to B$ for objects A,B,C,D in $\mathbf{Ob}(\mathcal{C})$. Furthermore for each object X in $\mathbf{Ob}(\mathcal{C})$ there is a morphism $\mathbb{1}_X : X\to X$ called the *identity morphism* such that

$$\mathbb{1}_B f = f = f\mathbb{1}_A,$$

for a morphism $f:A\to B$.

Definition 1.3.2. Let $\mathcal{C}$ be a category. An *isomorphism* $f:A\to B$ is a morphism in $\mathbf{Hom}(\mathcal{C})$ such that there is another morphism $f^{-1}:B\to A$ satisfying,

$$ff^{-1} = \mathbb{1}_B \text{ and } f^{-1}f = \mathbb{1}_A.$$

Definition 1.3.3. A category $\mathcal{C}$ is called a *groupoid* if every f in $\mathbf{Hom}(\mathcal{C})$ is an isomorphism

Definition 1.3.4. A *subcategory* $\mathcal{D}$ of a category $\mathcal{C}$ is a subclass of $\mathbf{Ob}(\mathcal{C})$ together with a subclass of $\mathbf{Hom}(\mathcal{C})$ that constitutes a category

Remark 1.3.5. equivalently a subcategory of $\mathcal{C}$ is a subclass $\mathbf{Ob}(\mathcal{D})$ of $\mathbf{Ob}(\mathcal{C})$ and a subclass $\mathbf{Hom}(\mathcal{D})$ of $\mathbf{Hom}(\mathcal{C})$ such that each domain A and codomain B for a morphism in $\mathbf{Hom}(\mathcal{D})$, A,B are elements of $\mathbf{Ob}(\mathcal{D})$. In addition $\mathbf{Hom}(\mathcal{D})$ is closed under composition.

Definition 1.3.6. The *maximal groupoid* of a category $\mathcal{C}$ is the subcategory of $\mathcal{C}$ whose objects are $\mathbf{Ob}(\mathcal{C})$ and whose morphisms are the isomorphisms of $\mathbf{Hom}(\mathcal{C})$

Remark 1.3.7. The maximal groupoid is a subcategory. Indeed, The domain and codomain of a morphism is trivially contained in $\mathbf{Ob}(\mathcal{C})$. Suppose $f:A\to B$ and $g:B\to C$ are isomorphisms. Then

$$f^{-1}g^{-1}gf = f^{-1}\mathbb{1}_B f = f^{-1}f = \mathbb{1}_A$$

and

$$gff^{-1}g^{-1} = g\mathbb{1}_B g^{-1} = gg^{-1} = \mathbb{1}_C.$$

Hence $gf : A \to C$ is an isomorphism with inverse $f^{-1}g^{-1}$.

Definition 1.3.8. Let $\mathcal{C}$ be a category and $f \in \mathrm{Hom}(A,B)$, $g \in \mathrm{Hom}(B,A$. f is called a *retraction* of g and g a section of f if $fg = \mathbb{1}_A$

Lemma 1.3.9. *Let $\mathcal{C}$ be a category, $f \in \mathrm{Hom}(A,B)$. Suppose $g,h \in \mathrm{Hom}(B,A)$ are respectively a retraction and a section of f. Then $g = h$ and f is an isomorphism. It follows that a morphism can have at most one inverse*

Proof. Indeed

$$g = g\mathbb{1}_A = gfh = \mathbb{1}_B h = h.$$

Hence f is an isomorphism with $f^{-1} = g = h$. Let f_1 and f_2 be inverses of an isomorphism f. Note that both f_1 and f_2 is both a section and a retraction. Therefor, by the first statement, $f_1 = f_2$. $\square$

Example 1.3.10.

1. Let **Set** be defined by objects being sets and morphisms being functions. Indeed, letting $\circ$ be composition in the conventional way and letting $\mathbb{1}_X = \mathrm{id}_X : X \to X, x \mapsto x$, we see that this indeed defines a category.

2. Consider a pair (X,R) of a set X and a transitive, reflexive relation R on X. Let $(a,b),(b,c) \in R$. We define

$$(a,b)(b,c) := (a,c).$$

   This is indeed well defined since aRb and bRc implies aRc, in other words $(a,c) \in R$. Let another pair $(c,d) \in R$. Then $((a,b)(b,c))(c,d) = (a,c)(c,d) = (a,d)$ and $(a,b)((b,c),(c,d)) = (a,b)(b,d) = (a,d)$. We define $\mathbb{1}_a = (a,a)$, which is indeed in R. Then $(a,a)(a,b) = (a,b)$ and $(a,b)(b,b) = (a,b)$. So (X,R) indeed defines a morphism.

Definition 1.3.11. Let $\mathcal{C}$ be a category and A an object in $\mathcal{C}$. *The slice category of $\mathcal{C}$ under A* denoted $A/\mathcal{C}$ is the category whose objects are morphisms in $\mathrm{Hom}(\mathcal{C})$ with domain A and where a morphism from $f : A \to X$ and $g : A \to Y$ is a map $h : X \to Y$ such that

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

commutes. *The slice category of $\mathcal{C}$ over A* denoted $\mathcal{C}/A$ is $A/\mathcal{C}^{\mathrm{op}}$, i.e. objects are morphisms with codomain A and a morphism from $f : X \to A$ to $g : Y \to A$ is a morphism $h : X \to Y$ satisfying

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

Remark 1.3.12. Both these constructions are indeed categories: Consider morphisms h_{12} between $f_1 : A \to X$ & $f_2 : A \to Y$ and h_{23} between $f_2 : A \to Y$ & $f_3 : A \to Z$. Then we have commutative diagrams

$$\begin{tikzcd} & A \arrow[dl, "f_1"'] \arrow[dr, "f_2"] & \\ X \arrow[rr, "h_{12}"] & & Y \end{tikzcd} \qquad \begin{tikzcd} & A \arrow[dl, "f_2"'] \arrow[dr, "f_3"] & \\ Y \arrow[rr, "h_{23}"] & & Z \end{tikzcd}$$

to obtain the commutative diagram

$$\begin{tikzcd} & A \arrow[dl, "f_1"'] \arrow[d, "f_2"] \arrow[dr, "f_3"] & \\ X \arrow[r, "h_{12}"] & Y \arrow[r, "h_{23}"] & Z \end{tikzcd}$$

hence $h_{23}h_{12}$ is a morphism between f_1 and f_3. For an object $f : A \to X$ in $A/\mathcal{C}$ define the identity morphism to be $\mathbb{1}_X$. We thus get that associativity of composition and the identity morphisms being neutral with respect to composition is inherited from this being true in $\mathcal{C}$. Reversing arrows we get that $\mathcal{C}/A$ is also a category.

Definition 1.3.13. Let $\mathcal{C}_1, \mathcal{C}_2$ be categories. A *Covariant functor* from $\mathcal{C}_1$ to $\mathcal{C}_2$ is a mapping $\mathcal{F}$, denoted $\mathcal{F} : \mathcal{C}_1 \to \mathcal{C}_2$, which assigns to each object A in $\mathbf{Ob}(\mathcal{C}_1)$ to an object $\mathcal{F}(A)$ in $\mathbf{Ob}(\mathcal{C}_2)$ and to each morphism in $\mathbf{Hom}(\mathcal{C}_1)$, $f : A \to B$ a morphism in $\mathbf{Hom}(\mathcal{C}_2)$, $\mathcal{F}(f) : \mathcal{F}(A) \to \mathcal{F}(B)$ such that

1. for every object X in $\mathbf{Ob}(\mathcal{C}_1)$, $\mathcal{F}(1_X) = 1_{\mathcal{F}(X)}$.

2. for every pair of morphisms $f : B \to C$ and $g : A \to B$ in $\mathbf{Hom}(\mathcal{C}_1)$, $\mathcal{F}(fg) = \mathcal{F}(f)\mathcal{F}(g)$.

Lemma 1.3.14. *Consider two categories $\mathcal{C}_1$ and $\mathcal{C}_2$ with a functor $\mathcal{F} : \mathcal{C}_1 \to \mathcal{C}_2$. If $f : A \to B$ is an isomorphism in $\mathbf{Hom}(\mathcal{C}_1)$, then $\mathcal{F}(f) : \mathcal{F}(A) \to \mathcal{F}(A)$ is an isomorphism in $\mathbf{Hom}(\mathcal{C}_2)$.*

Proof. Indeed,

$$\mathcal{F}(f)\mathcal{F}\left(f^{-1}\right) = \mathcal{F}\left(ff^{-1}\right) = \mathcal{F}(\mathbb{1}_B) = \mathbb{1}_{\mathcal{F}(B)} \text{ and } \mathcal{F}\left(f^{-1}\right)\mathcal{F}(f) = \mathcal{F}\left(f^{-1}f\right) = \mathcal{F}(\mathbb{1}_A) = \mathbb{1}_{\mathcal{F}(A)}.$$

□

Definition 1.3.15. Let $\mathcal{C}$ be a category. We define *the opposite category of* $\mathcal{C}$ denoted $\mathcal{C}^{\text{op}}$ to be the category with $\mathbf{Ob}(\mathcal{C}^{\text{op}}) := \mathbf{Ob}(\mathcal{C})$ and where a morphism $f : A \to B$ in $\mathbf{Hom}(\mathcal{C}^{\text{op}})$ is a morphism $f : B \to A$ in $\mathbf{Hom}(\mathcal{C})$

Remark 1.3.16. The above indeed does define a category. We define $\circ^{\text{op}}$ by

$$f \circ^{\text{op}} g = g \circ f : C \to A$$

where $f : B \to C$ and $g : A \to B$ are morphisms in $\mathbf{Hom}(\mathcal{C}^{\text{op}})$. Then for morphisms $f : C \to D$, $g : B \to C$ and $h : A \to B$

$$(f \circ^{\text{op}} g) \circ^{\text{op}} h = h(gf) = (hg)f = f \circ^{\text{op}} (g \circ^{\text{op}} h).$$

Furthermore, we define the identity morphism in $\mathbf{Hom}(\mathcal{C}^{\text{op}})$ to be the identity morphism in $\mathbf{Hom}(\mathcal{C})$, hence

$$f \circ^{\text{op}} \mathbb{1}_A = \mathbb{1}_A f = f \text{ and } \mathbb{1}_B \circ^{\text{op}} f = f \mathbb{1}_B = f.$$

Definition 1.3.17. Consider categories $\mathcal{C}_1$ and $\mathcal{C}_2$. A *covariant functor* $\mathcal{F}$ between $\mathcal{C}_1$ and $\mathcal{C}_2$ is a covariant functor between $\mathcal{C}_1$ and $\mathcal{C}_2^{\text{op}}$.

Corollary 1.3.18. *Consider categories* $\mathcal{C}_1$*,* $\mathcal{C}_2$ *and a covariant functor* $\mathcal{F} : \mathcal{C}_1 \to \mathcal{C}_2^{\text{op}}$*. If* $f : A \to B$ *is an isomorphism in* $\mathbf{Hom}(\mathcal{C}_1)$*, then* $\mathcal{F}(f) : \mathcal{F}(B) \to \mathcal{F}(A)$ *is an isomorphism in* $\mathbf{Hom}(\mathcal{C}_2^{\text{op}})$

Proof. This follows immediately from Lemma 1.3.14. □

Example 1.3.19. Suppose that there, for a category $\mathcal{C}$, is a well-defined assignment $\mathcal{F}$ of objects in $\mathcal{C}$ to integers and of a morphism $A \to B$ to $\mathcal{F}(A) \leq \mathcal{F}(B)$. This will define a functor from $\mathcal{C}$ to $(\mathbb{Z}, \leq)$ called *an integer invariant on* $\mathcal{C}$. Indeed, $\mathcal{F}(A) \leq \mathcal{F}(A)$, hence $\mathcal{F}(\mathbb{1}_A) = \mathbb{1}_{\mathcal{F}(A)}$. Given morphisms $g : A \to B$, $f : B \to C$ in $\mathbf{Hom}(\mathcal{C})$,

$$\mathcal{F}(A) \leq \mathcal{F}(B) \text{ and } \mathcal{F}(B) \leq \mathcal{F}(C),$$

implying $\mathcal{F}(A) \leq \mathcal{F}(C)$, hence $\mathcal{F}(A \xrightarrow{fg} C) = \mathcal{F}(A \xrightarrow{f} B)\mathcal{F}(B \xrightarrow{g} C)$.

Definition 1.3.20. A category $\mathcal{C}$ is *locally small* if $\mathbf{Hom}(A,B)$ is a set for every object A,B in $\mathbf{Ob}(\mathcal{C})$. It is *small* if $\mathbf{Ob}(\mathcal{C})$ is a set.

Proposition 1.3.21. *Consider the class of small categories with morphisms being functors. This defines a category denoted* **Cat***.*

Proof. DO! □

Definition 1.3.22. In a category $\mathcal{C}$ a morphism $f \in \mathrm{Hom}(A,B)$ is a *monomorphism* if for every pair of morphisms $g_1, g_2 \in \mathrm{Hom}(C,A)$,

$$fg_1 = fg_2 \Rightarrow g_1 = g_2.$$

It is called an *epimorphism* if for every pair of morphisms $h_1, h_2 \in \mathrm{Hom}(B,D)$,

$$h_1 f = h_2 f \Rightarrow h_1 = h_2$$

Lemma 1.3.23. *For a category $\mathcal{C}$ a*

1.3.2 Products & Co-products

Definition 1.3.24. Let A be a set and $\{X_\alpha\}_{\alpha\in A}$ be a family of sets. We then define *the direct product of X_α over A* to be the set

$$\prod_{\alpha\in A} X_\alpha := \left\{ f : A \to \bigcup_{\alpha\in A} X_\alpha : f(\alpha) \in X_\alpha \text{ for every } \alpha \in A \right\}.$$

Remark 1.3.25. We can identify every function $f : A \to \bigcup_{\alpha\in A}$ can be identified with a set $\{r_\alpha : \alpha \in A\}$. In particular, every $f \in \prod_{\alpha\in A} X_\alpha$ can be identified with a symbol (r_α) where $r_\alpha \in X_\alpha$ for each $\alpha \in A$. Thus

$$\prod_{\alpha\in A} X_\alpha = \left\{ (r_\alpha) : r_\alpha \in X_\alpha \text{ for every } \alpha \in A \right\}.$$

Assuming the axiom of choice every such product is non-empty whenever $\{X_\alpha\}_{\alpha\in A}$ is a family of non-empty sets. For a finite family of sets $\{X_1, \ldots, X_n\}$ we can identify $\prod_1^n X_i := \prod_{i\in\{1,\ldots,n\}} X_i$ with $X_1 \times \cdots \times X_n$.

Axiom. *Let A be a non-empty set. When $\{X_\alpha\}$ is a family of non-empty sets $\prod_{\alpha\in A} X_\alpha \neq \emptyset$.*

Proposition 1.3.26. *Let A be a set and $\{X_\alpha\}_{\alpha\in A}$ be a family of non-empty sets. For each $\alpha \in A$, define $\pi_\alpha : \prod_{\alpha\in A} X_\alpha \to X_\alpha, (x_\alpha) \mapsto x_\alpha$. For $\alpha \in A$, π_α is a surjective map such that for every set Y with maps $\{f_\alpha : Y \to M_\alpha\}_{\alpha\in A}$ there is a unique map $f : Y \to \prod_{\alpha\in A} X_\alpha$ such that for every $\alpha \in A$, $\pi_\alpha \circ f = f_\alpha$*

Proof. π_α **is surjective:** Let $\alpha \in A$ and $x_\alpha \in X_\alpha$. Using the axiom of choice there is a function mapping $\beta \mapsto x_\beta$ for some $x_\beta \in X_\beta$ for each $\beta \in A \setminus \{\alpha\}$. Then $(x_\beta) \in \prod_{\beta \in A} X_\beta$. Then $\pi_\alpha((x_\beta)) = x_\alpha$.

Existence of f: We define $f(y) = (f_\alpha(y)) \in \prod_{\alpha \in A} X_\alpha$, which is easily seen to be well defined. Then for each $y \in Y$, $\alpha \in A$,

$$\pi_\alpha \circ f(y) = \pi_\alpha(f(y)) = \pi_\alpha((f_\beta(y))) = f_\alpha(y) \Rightarrow \pi_\alpha \circ f = f_\alpha$$

Uniqueness of f: Let $g : Y \to \prod_{\alpha \in A} X_\alpha$ be another map satisfying $\pi_\alpha \circ g = f_\alpha$ for each $\alpha \in A$. Let $y \in Y$. Then there is a $(x_\alpha) \in \prod_{\alpha \in A} X_\alpha$ such that $g(y) = (x_\alpha)$. Then for $\beta \in A$

$$x_\beta = \pi_\beta((x_\alpha)) = \pi_\beta(g(y)) = \pi_\beta \circ g(y) = f_\beta(y),$$

which implies that

$$f(y) = (f_\alpha(y)) = (x_\alpha) = g(y).$$

□

1.3.3 Currying

2 Algebra

The field of abstract algebra can classically be defined as the study of sets with operations and the study of maps between such objects. I.e. an object of study in algebra would be a set with some non-empty collection of functions, which somehow act upon elements in the set. An example of such an operation would be a *binary operation* on a set, M say, i.e. a function taking a pair in $M \times M$ to an element in M. Such a pair of a set and a binary operation is called a *magma*. It could also be that another set S acts on M, i.e. there is a function taking a pair of elements in $S \times M / M \times S$ to an element in M. We call such a structure an S-*act*DO WE?.
When we ask that such operation or the overlying sets adhere to certain axioms we obtain a rich family of sub-classes of objects having a certain kind of algebraic structure which will be preserved by certain maps. We will in the following be focusing broadly on the classes of rings and modules. However, it will be useful to also introduce groups and monoid in this context.
Before we begin, consider the following comment on the nature of much of this theory: There are often very few restrictions on the sets being considered in algebra while the operations will have many more restrictions. This means that to get from A to B in a proof it feels like one has to move through a very rigid structure that

doesn't allow many choices or much creativity. Sometimes the right path from A to B will require small clever tricks, but ultimately the algebraic structure will dictate a fixed path for how a proof will go. This the opposite of what of what analysis feels like. In analysis one has a lot of freedom in constructing the functions, sequences, choice of ϵ's etc. that will do the trick, which can both result really clever and elegant uses solution or solutions that are less well thought out or elegant.
The creative aspect of algebra, is that of giving of giving good definitions and having a natural eye for the most natural constructions. With the right definition it can become clearer how a difficult question should be answered. Somehow what is important in algebra, lives among the objects of study and not among the elements in these object, hence the objective is more so to find the right (class of) object(s). This may not ever become very clear from these notes, but the hope is that some shadows of this fact(?) will be present.

2.1 Monoids

2.1.1 Definitions and Basic Properties

Definition 2.1.1. A *monoid* is a set M with an operation $\circ : M \times M \to M$ where $m_1 m_2 := m_1 \circ m_2 := \circ(m_1, m_2)$ for $m_1, m_2 \in M$ that satisfy the following two axioms

1. The operation $\circ$ satisfies the *associative law*, i.e. for every $m_1, m_2, m_3 \in M$,
$$m_1(m_2 m_3) = (m_1 m_2) m_3.$$
2. There is an element $e \in M$ such that for every $m \in M$,
$$me = em = m.$$
The element e is referred to as the *neutral element with respect to* $\circ$.

The data specifying a monoid is often written as the tuple $(M, \circ)$.

Remark 2.1.2. The neutral element with respect to $\circ$ is unique. Indeed, suppose $e, e' \in M$ are neutral with respect to $\circ$. Then
$$e = ee' = e'.$$

For an element $m \in M$ and a non-negative integer n we define
$$m^n = \underbrace{m \cdots m}_{n}$$
with the convention that $m^0 = e$.

Definition 2.1.3. A *commutative monoid* is a monoid M such that for every $m_1, m_2 \in M$,

$$m_1 m_2 = m_2 m_1.$$

Definition 2.1.4. Let $(M, \circ)$ be a monoid. A subset $N \subset M$ is called a *submonoid (of M)* if

1. $e \in N$

2. For every $n_1, n_2 \in N$,

$$n_1 n_2 \in N.$$

Remark 2.1.5. $(N, \circ|_N)$ is a monoid. Indeed, Since $n_1 n_2 \in N$ for every $n_1, n_2 \in N$, the operation $\circ|_{N \times N}: N \times N \to N$ is well-defined. The operation $\circ$ is associative on N since it is associative on M. By the definition of a submonoid $e \in N$ and again clearly the property of being the neutral element with respect to $\circ$ on N is inherited by e being so with respect to $\circ$ on M.

Example 2.1.6.
1. The non-negative integers $\mathbb{N}$ is a monoid with respect to addition and multiplication.

2. $(\mathbb{Z},+), (\mathbb{Z},\cdot), (\mathbb{Q},+), (\mathbb{Q},\cdot), (\mathbb{R},+), (\mathbb{R},\cdot), (\mathbb{C},+), (\mathbb{C},\cdot)$ are monoids.

3. Let A be a set. Consider $\mathrm{Fun}(A,A) := \{f : A \to A\}$. This a monoid under function composition.

4. Given a non-empty set X and a monoid M the set

$$\mathbf{Fun}(X,M) := \{f : X \to M\}$$

with $fg \in \mathrm{Fun}(X,M)$ defined by $fg(x) := f(x)g(x)$ for $f,g \in \mathrm{Fun}(X,M)$ and $x \in X$ with $fg \in \mathbf{Fun}(X,M)$ defined by $fg(x) = f(x)g(x)$. Indeed, given $f,g,h \in \mathbf{Fun}(X,R)$ and $x \in M$

$$(fg)h(x) = (fg)(x)h(x) = (f(x)g(x))h(x) = f(x)(g(x)h(x)) = f(x)(gh)(x) = f(gh)(x).$$

And for the function $e : X \to M$, mapping every element in X to e_M we have that

$$ef(x) = e(x)f(x) = e_M f(x) = f(x) \text{ and } fe(x) = f(x)e(x) = f(x)e_M = f(x).$$

5. Let M be a monoid. Then $M \subset M$ is a submonoid.

6. Let M be a monoid. Then $\{e\} \subset M$ is a submonoid.

7. Let M be a monoid and $L \subset N \subset M$ be submonoids of M. Then L is a submonoid of N. Similarly if $N \subset M$ is a submonoid and $L \subset N$ is a submonoid, then $L \subset M$ is a submonoid.

2.1.2 Morphisms of Monoids

Definition 2.1.7. Let M,N be monoids. A *monoid homomorphism/map of monoids/morphism of monoids* is a map $\rho : M \to N$ such that

1. For every $m_1, m_2 \in M$
$$\rho(m_1 m_2) = \rho(m_1)\rho(m_2).$$

2. $$\rho(e_M) = e_N.$$

Denote the set of homomorphisms from M to N by $\mathbf{Hom}^{\mathrm{Mon}}(M,N)$.

Remark 2.1.8. Let **Monoid** be the class of monoids and $\mathbf{Hom}^{\mathbf{Mon}}$ the class of monoid homomorphisms. One readily verifies that $(\mathbf{Monoid}, \mathbf{Hom}^{\mathbf{Mon}})$ is a category. Potential to write more.

Remark 2.1.9. By a prior example (cf. Example 2.1.6) we have seen that $\mathbf{Fun}(M,N)$ is a monoid. $\mathbf{Hom}^{\mathrm{Mon}}(M,N) \subset \mathbf{Fun}(M,N)$ is a submonoid if N is commutative. Indeed, for $f,g \in \mathrm{Hom}^{\mathrm{Mon}}(M,N)$ and $x,y \in M$. Then

$$fg(xy) = f(xy)g(xy) = f(x)f(y)g(x)g(y) = f(x)g(x)f(y)g(y) = fg(x)fg(y) \Rightarrow fg \in \mathrm{Hom}^{\mathrm{Mon}}(M,N).$$

Furthermore we have

$$e(xy) = e_M(xy) = xy = (e_N x)(e_N y) = e(x)e(y) \Rightarrow e \in \mathrm{Hom}^{\mathrm{Mon}}(M,N).$$

Lemma 2.1.10. *Let $\rho : M \to N$ be a monoid homomorphism and $L \subset M$ a submonoid. Then $\rho(L) \subset N$ is a submonoid.*

Proof. Let $\rho(l_1), \rho(l_2) \in \rho(L)$. Then since $l_1 l_2 \in L$,

$$\rho(l_1)\rho(l_2) = \rho(l_1 l_2) \in \rho(L).$$

Clearly $e_N = \rho(e_M) \in \rho(L)$. □

Corollary 2.1.11. *The image of a monoid homomorphism $\rho : M \to N$ is a submonoid of N.*

Proof. This follows from the above lemma (cf. Example 2.1.6). □

Definition 2.1.12. Let $\rho : M \to N$ be a monoid homomorphism. We define *the kernel of* ρ to be the set

$$\mathbf{ker}\,\rho := \rho^{-1}(e_N) = \{m \in M : \rho(m) = e_N\} \subset M$$

Lemma 2.1.13. *Let* $\rho : M \to N$ *be a monoid homomorphism, and* $L \subset N$ *a submonoid. Then* $\rho^{-1}(L) \subset M$ *is a submonoid.*

Proof. Let $m_1, m_2 \in \rho^{-1}(L)$. Then since $\rho(m_1), \rho(m_2) \in L$,

$$\rho(m_1 m_2) = \rho(m_1)\rho(m_2) \in L,$$

hence $m_1 m_2 \in \rho^{-1}(L)$. Since $\rho(e_M) = e_N \in L$, it follows that $\rho^{-1}(L)$ is a submonoid of N. □

Corollary 2.1.14. *The kernel of a monoid homomorphism* $\rho : M \to N$ *is a submonoid of* M.

Proof. Since $\{e_N\}$ is a submonoid of N it follows by the above lemma that $\mathbf{ker}\ \rho = \rho^{-1}(\{e_N\})$ (cf. Example 2.1.6) is a submonoid. □

Lemma 2.1.15. *Let* M *be a commutative monoid and* $N \subset M$ *a submonoid. Then* N *is a commutative monoid.*

Proof. By Lemma 2.1.5 N is a monoid. Let $n_1, n_2 \in N$, then since $n_1, n_2 \in M$, $n_1 n_2 = n_2 n_1$. □

Lemma 2.1.16. *Let* $\rho : M \to N$ *be a monoid homomorphism. Let* $L \subset M$ *be a submonoid. If* M *is commutative, then* $\rho(L) \subset N$ *is a commutative monoid.*

Proof. Since $\rho(L) \subset N$ is a submonoid, it is a monoid. Let $\rho(l_1), \rho(l_2) \in \rho(L)$. Then since L is a commutative by Lemma 2.1.15 it follows that

$$\rho(l_1)\rho(l_2) = \rho(l_1 l_2) = \rho(l_2 l_1) = \rho(l_2)\rho(l_1).$$

□

2.1.3 Product Monoids & Restricted Product of Monoids

Theorem 2.1.17. *Let A be a set and $\{M_\alpha\}_{\alpha\in A}$ a family of monoids. We define a binary operation on the product $\prod_{\alpha\in A} M_\alpha$ by $(m_\alpha)(m'_\alpha) = (m_\alpha m'_\alpha)$ for $(m_\alpha),(m'_\alpha) \in \prod_{\alpha\in A} M_\alpha$. With this operation $\prod_{\alpha\in A} M_\alpha$ becomes a monoid. If M_α is commutative for every $\alpha \in A$ so is $\prod_{\alpha\in A} M_\alpha$.*

Proof. We define $e := (e) := (e_\alpha)$ where e_α is the neutral element in M_α for each α. Let $(m_\alpha),(m'_\alpha),(m''_\alpha) \in \prod_{\alpha\in A} M_\alpha$. We then have that

$$(m_\alpha)\big((m'_\alpha)(m''_\alpha)\big) = (m_\alpha)(m'_\alpha m''_\alpha) = (m_\alpha(m'_\alpha m''_\alpha)) = ((m_\alpha m'_\alpha)m''_\alpha) = (m_\alpha m'_\alpha)(m''_\alpha) = \big((m_\alpha)(m'_\alpha)\big)(m''_\alpha)$$

and that

$$e(m_\alpha) = (e)(m_\alpha) = (e_\alpha m_\alpha) = (m_\alpha) \text{ and } (m_\alpha)e = (m_\alpha)(e) = (m_\alpha e_\alpha) = (m_\alpha),$$

hence $(\prod_{\alpha\in A} M_\alpha, \cdot)$ is a monoid. Suppose M_α is commutative for each $\alpha \in A$. Then

$$(m_\alpha)(m'_\alpha) = (m_\alpha m'_\alpha) = (m'_\alpha m_\alpha) = (m_\alpha)(m'_\alpha).$$

□

Lemma 2.1.18. *Let A be a set and $\{M_\alpha\}_{\alpha\in A}$ a family of monoids. Then*

$$\pi_\beta : \prod_{\alpha\in A} M_\alpha \to M_\beta$$

is monoid homomorphism. Given a monoid N and a family monoid homomorphisms $\{f_\alpha : N \to M_\alpha\}_{\alpha\in A}$ then the unique map $f : N \to \prod_{\alpha\in A}$ (cf. Proposition 1.3.26) such that $\pi_\alpha \circ f = f_\alpha$ for every $\alpha \in A$ is monoid homomorphism.

Proof. Let $(m_\alpha),(m'_\alpha) \in \prod_{\alpha\in A} M_\alpha$ and fix $\beta \in A$. Then

$$\pi_\beta((m_\alpha)(m'_\alpha)) = \pi_\beta((m_\alpha m'_\alpha)) = m_\beta m'_\beta = \pi_\beta((m_\alpha))\pi_\beta((m'_\alpha)).$$

Lastly

$$\pi_\beta(e) = \pi_\beta((e_\alpha)) = e_\beta.$$

Let $n, n' \in N$. Then

$$f(nn') = (f_\alpha(nn')) = (f_\alpha(n) f_\alpha(n')) = (f_\alpha(n))(f_\alpha(n')) = f(n)f(n'),$$

and

$$f(e_N) = (f_\alpha(e_N)) = (e_\alpha) = e$$

□

Proposition 2.1.19. *Let A be a set and $\{M_\alpha\}_{\alpha\in A}$ a family of monoids. Consider a family of sets $\{N_\alpha\}_{\alpha\in A}$ such that $N_\alpha \subset M_\alpha$ is a submonoid for each $\alpha \in A$. Then*

$$\prod_{\alpha\in A} N_\alpha \subset \prod_{\alpha\in A} M_\alpha$$

is a submonoid.

Proof. Since $e_\alpha \in N_\alpha$ for each $\alpha \in A$. Then $e = (e_\alpha) \in \prod_{\alpha\in A} N_\alpha$. Let $(n_\alpha),(n'_\alpha) \in \prod_{\alpha\in A} N_\alpha$. Then since $n_\alpha n'_\alpha \in N_\alpha$ for each $\alpha \in A$. This implies that $(n_\alpha)(n'_\alpha) = (n_\alpha n'_\alpha) \in \prod_{\alpha\in A} N_\alpha$. □

Example 2.1.20. Not every submonoid of a monoid arises in this fashion. For instance consider $N = \{(n,n) \in \mathbb{N}\times\mathbb{N}\}$ which is a proper submonoid of $(\mathbb{N}\times\mathbb{N},+)$. Indeed, $0_{\mathbb{N}\times\mathbb{N}} = (0,0) \in N$ and if $(n_1,n_1),(n_2,n_2) \in N$, then $(n_1,n_1)+(n_2,n_2) = (n_1+n_2,n_1+n_2) \in N$. Show it is not product of submonoids

Remark 2.1.21. With products introduced, at this point we will introduce some notation. Consider a monoid M. Let A be a non-empty set, $\{M_\alpha\}$ a family of submonoids of M and suppose we are given $(m_\alpha) \in \prod_{\alpha\in A} M_\alpha$ such that $m_\alpha = 0$ for all but finitely many $\alpha \in A$. Then there are $\alpha_1,\dots,\alpha_n \in A$ such that $m_\alpha = 0$ for every $\alpha \in A\setminus\{\alpha_1,\dots,\alpha_n\}$. We then define

$$\prod_{\alpha\in A} m_\alpha := \prod_1^n m_{\alpha_i}.$$

We first note that $\prod_{\alpha\in A} m_\alpha$ is an element of M. Suppose $\{\beta_1,\dots,\beta_m\} \subset A$ is another subset such that $m_\alpha = 0$ for all $\alpha \in A\setminus\{\beta_1,\dots,\beta_m\}$. If $m_\alpha = e$ for all $\alpha \in A$, then clearly

$$\prod_1^m m_{\beta_j} = e = \prod_1^n m_{\alpha_i}.$$

If there is an $i \in \{1,\dots,n\}$ such that $m_{\alpha_i} \neq e$, then $\alpha_i = \beta_{j(i)}$ for some $j(i) \in \{1,\dots,m\}$, for if not, $\alpha_i \in X\setminus\{\beta_1,\dots,\beta_m\}$, which would imply $m_{\alpha_i} = e$. We can show that $i \mapsto j(i)$ is a bijection using the same argument for the non-zero m_{β_j} to show that there is a $i(j)$ such that $\beta_j = \alpha_{i(j)}$. It then follows that

$$\prod_1^n m_{\alpha_i} = \prod_{i\in\{1,\dots,n\}:m_{\alpha_i}\neq e} m_{\alpha_i} = \prod_{i\in\{1,\dots,m\}:m_{\beta_i}\neq e} m_{\beta_i} = \prod_1^m m_{\beta_i},$$

hence the notion is independent of the choice of the elements of A corresponding to possibly non-zero entries of (m_α).

A postemptive note after the above construction: The author of these notes, realizes that it is intuitively rather obvious, what is to be understood by $\prod_{\alpha\in A} m_\alpha$ and that it makes sense (is well-defined). It might it even be obvious - PERIOD! Somehow this construction just feels like a notational trick. If anyone should, by some weird coincidence, read these notes, note that the author being fixated on being (overly) precise in some instances, is a result of wanting to make sure that their understanding of what is going on, is precise AND EVEN FORMALISABLE - in some instances at least. The other instances where this seems not to be the case, it is either because the author doesn't care or that they have postponed it. Care is often given when the answer to the question seems easy enough to be done in LEAN. For example, if we knew what a monoid M and $\prod_{\alpha\in A} \bullet$ is in LEAN, it seem rather easy(?) to prove that the function

$$\prod_{\alpha\in A} M_\alpha \to M, (m_\alpha) \mapsto \prod_{\alpha\in A} m_\alpha,$$

is well-defined in a set-theoretic sense in LEAN or it should at least be easy to see how $\prod_{\alpha\in A} m_\alpha$ should be defined in LEAN from what has been written in this remark.

Definition 2.1.22. Let A be a set and $\{M_\alpha\}_{\alpha\in A}$ a family of monoids. We define the *restricted direct product* of M_α over A as the set

$$\prod_{\alpha\in A}' M_\alpha := \left\{(m_\alpha) \in \prod_{\alpha\in A} M_\alpha : m_\alpha = e_\alpha \text{ for all but finitely many } \alpha \in A\right\}$$

Lemma 2.1.23. *Let A be a set and $\{M_\alpha\}_{\alpha\in A}$ a family of monoids. $\prod'_{\alpha\in A} M_\alpha$ is a submonoid $\prod_{\alpha\in A} M_\alpha$.*

Proof. Let $(m_\alpha), (m'_\alpha) \in \prod'_{\alpha\in A} M_\alpha$. For some distinct $\alpha_1, \dots, \alpha_r \in A$ and $\beta_1, \dots, \beta_p \in A$, $m_\alpha = e_\alpha$ for every $\alpha \in A \setminus \{\alpha_1, \dots, \alpha_r\}$ and $m'_\alpha = e_\alpha$ for every $\alpha \in A \setminus \{\beta_1, \dots, \beta_p\}$. Then $m_\alpha m'_\alpha = e_\alpha$ for every $\alpha \in A \setminus \{\alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_p\}$ hence $(m_\alpha)(m'_\alpha) = (m_\alpha m'_\alpha) \in \prod'_{\alpha\in A} M_\alpha$. Clearly $e = (e_\alpha) \in \prod'_{\alpha\in A} M_\alpha$. □

2.2 Groups

2.2.1 Definition & Basic Properties

Definition 2.2.1. A *group* is a monoid $(G, \circ)$ where for every $g \in G$ there is an element $g^{-1} \in G$ such that

$$gg^{-1} = g^{-1}g = e.$$

For $g \in G$ we refer to g^{-1} as the *inverse of g with respect to i*. The data specifying a group is also often written as the tuple $(G, \circ)$.

Remark 2.2.2. For an element $g \in G$ and a non-negative integer n, we define $g^{-n} = \left(g^{-1}\right)^n$. It is easy to check that $(g^n)^{-1} = g^{-n}$.

Definition 2.2.3. A group $(G,+)$ is called *abelian* or *additive*, if it is also a commutative monoid. We denote the inverse of $g \in G$ with respect to addition by $-g$, and for $g_1, g_2 \in G$ we define

$$g_1 - g_2 := g_1 + (-g_2).$$

and $ng_1 = \underbrace{g_1 + \ldots + g_1}_{n \text{ times}}$

Lemma 2.2.4. *Let $(G,\circ)$ be a group. Let $g, g', a \in G$. If $ag = ag'$, then $g = g'$. Similary, if $ga = g'a$, then $g = g'$.*

Proof. We have that $(a^{-1}, ag) = (a^{-1}, ag')$, hence

$$g = eg = (a^{-1}a)g = a^{-1}(ag) = a^{-1}(ag') = (a^{-1}a)g' = eg' = g.$$

The proof of the other statement is dual. □

Lemma 2.2.5. *Let $(G,\circ)$ be a group. The following is true*

1. *Inverse elements are unique*

2. *For every $g, g' \in G$,*
$$(gg')^{-1} = g'^{-1}g^{-1}$$

3. *For every $g \in G$,*
$$(g^{-1})^{-1} = g$$

Proof. 1. Let $g \in G$ and consider g', g'' such that $g'g = gg' = e$ and $g''g = gg'' = e$. Then

$$gg' = e = gg'',$$

hence $g' = g''$ by the prior lemma.
2. One easily check that both $(gg')^{-1}$ and $g'^{-1}g^{-1}$ are inverse elements of gg'. It then follows from 1. that $(gg')^{-1} = g'^{-1}g^{-1}$.
3. One easily sees that $\left(g^{-1}\right)^{-1}$ and g are inverse elements of g^{-1}. It follows from 1. that $\left(g^{-1}\right)^{-1} = g$. □

Remark 2.2.6. One should note that if we in 1. for $g \in G$ only proved that elements $g' \in G$ satisfying $gg' = e$ were unique, this would still be sufficient to prove 2. and 3. This in addition means that if $gg^{-1} = e$ then

$$g^{-1}g = g^{-1}\left(g^{-1}\right)^{-1} = (g^{-1}g)^{-1} = e.$$

Since the first statement in Lemma 2.2.4 only uses $eg = g$ for every $g \in G$ and we only ever make use first statement of this lemma in 1. then we can prove that that if $eg = g$ for every $g \in G$, then

$$ge = \left(g^{-1}\right)^{-1}\left(e^{-1}\right)^{-1} = \left(e^{-1}g^{-1}\right)^{-1} = \left(eg^{-1}\right)^{-1} = \left(g^{-1}\right)^{-1} = g,$$

for every $g \in G$. In other words it is sufficient to check that $eg = g$ and $gg^{-1} = e$ for every $g \in G$, when checking the group axioms under the assumption that axiom 1. is already fulfilled.

Definition 2.2.7. Let G be a group. A subset $H \subset G$ is called a *subgroup* if it is a submonoid of G and for every $h \in H$ we have that $h^{-1} \in H$.

Remark 2.2.8. $(H, \circ)$ is a group. Indeed, $(H, \circ)$ is a monoid since $H \subset G$ is a submonoid. Since $h^{-1} \in H$ for every $h \in H$, it follows that every element in H has in inverse in H, hence H is a group.

Example 2.2.9.

1. For $n \in \mathbb{Z}$, the set $n\mathbb{Z} := \{nm : m \in \mathbb{Z}\}$ is a subgroup of $(\mathbb{Z}, +)$.
2. $\mathbb{R} \subset \mathbb{C}$ is a subgroup of $(\mathbb{C}, +)$. $\mathbb{R} \setminus \{0\} \subset \mathbb{C} \setminus \{0\}$ is a subgroup of $(\mathbb{C} \setminus \{0\}, \cdot)$
3. Let G be a group. Then G and $\{e\}$ are subgroups of G.
4. Given a set A. The set of invertible maps $A \to A$ forms a submonoid of $\mathbf{Fun}(A, A)$ under composition, furthermore it is a group, when picking the inverse elements to be inverse maps and the neutral to be the identity on A.
5. Given a non-empty set X and a group G the monoid $\mathbf{Fun}(X, G)$ forms a group. Indeed for $f \in \mathbf{Fun}(X, G)$, define $f^{-1} : X \to G$ by $f^{-1}(x) = (f(x))^{-1}$. Then
$$ff^{-1}(x) = f(x)(f(x))^{-1} = e_N = e(x).$$
6. Let G be a group and consider $I \subset H \subset G$. Then $I, H \subset G$ are subgroups if and only if $I \subset H$ and $H \subset G$ are subgroups.
7. Let X be any non-empty set and G a group. The set

Definition 2.2.10. Let G be a group and $S \subset G$. Then we define *the subgroup generated by S* to be the set

$$\langle S \rangle = \left\{s_1^{v_1} \cdots s_n^{v_n} \in G : n \geq 1, s_1, \ldots, s_n \in S, v_1, \ldots, v_n \in \{\pm 1\}\right\}.$$

Our convention will be that $\langle \emptyset \rangle = \{e\}$

Remark 2.2.11. Disallowing negative exponents in $\langle S\rangle$ gives the definition of *the submonoid generated by S*. From the following, we can derive that this is a submonoid even if we allow G to just be a monoid. If S is empty, it is clearly a subgroup. So suppose S is non-empty. Let $s_1^{v_1}\cdots s_n^{v_n}, t_1^{w_1}\cdots t_m^{w_m} \in \langle S\rangle$. Then if we define $s_i = t_{i-n}$ and $v_i = w_{i-n}$ for $i \in \{n+1,\dots,n+m\}$. Then

$$s_1^{v_1}\cdots s_n^{v_n} t_1 \cdots^{w_1} \cdots t_m^{w_m} = s_1^{v_1}\cdots s_{n+m}^{v_{n+m}} \in \langle S\rangle$$

Clearly $e \in \langle S\rangle$. We also have

$$\left(s_1^{v_1}\cdots s_n^{v_n}\right)^{-1} = s_n^{-v_n}\cdots s_1^{-v_1} \in \langle S\rangle.$$

It follows that $\langle S\rangle$ is a subgroup. Let $H \subset G$ be a subgroup containing S. Then clearly $s_1^{v_1}\cdots s_n^{v_n} \in H$. Thus $\langle S\rangle$ is the smallest subgroup containing S.
$S,T \subset G$ such that $S \subset T \subset G$. Then $\langle S\rangle \subset \langle T\rangle$. If $H \subset G$ is a subgroup, then $\langle H\rangle = H$, since H is the smallest subgroup containing H.

Definition 2.2.12. A group G is *finitely generated* if $G = \langle g_1,\dots,g_n\rangle$ for some $g_1,\dots,g_n \in G$. If G is generated by one element it is called *cyclic*.

Lemma 2.2.13. *Let G be a cyclic group. Then any subgroup of G is cyclic.*

Proof. Let $H \subset G$ be a subgroup. If $H = \{e\}$ we are done, so suppose it is not. We have that $G = \langle g\rangle$ for some g. The set $\{n > 0 : g^n \in H\}$ is non-empty and thus have a minimum by the well-ordering of the natural numbers. Call this number m. We claim that $\langle g^m\rangle = H$. The first inclusion is trivial. Let $h \in H$. Then $h = g^l$ for some $l \in \mathbb{Z}$. It is sufficient to check that $h \in \langle g^m\rangle$ the case where $l > 0$, so we assume this. By minimality $l \geq m$. Then $h = g^l = g^{qm+r}$ for some $q,r \geq 0$ and $r < m$. Then $g^r = g^{qm+r-qm} = q^{qm+r}q^{-qm} \in H$, hence by minimality $r = 0$, hence $h = g^{qm} \in \langle g^m\rangle$ □

2.2.2 Morphisms of groups

Definition 2.2.14. Let G,H be groups. A map $\rho : G \to H$ is called a *group homomorphism/map of groups/morphism of groups*, if for every $g_1, g_2 \in G$,

$$\rho(g_1g_2) = \rho(g_1)\rho(g_2).$$

Denote the set of group homomorphism between G and H by $\mathbf{Hom}^{\mathrm{Grp}}(G,H)$

Remark 2.2.15. 1. Denote the neutral elements of G and H by e_G and e_H respectively. Then $\rho(e_G) = e_H$. Indeed,

$$\rho(e_G)e_H = \rho(e_G e_G) = \rho(e_G)\rho(e_G),$$

hence by Lemma 2.2.4, $e_H = \rho(e_G)$.
We also have that $\rho\left(g^{-1}\right) = \rho(g)^{-1}$. Indeed,

$$\rho(g)\rho(g)^{-1} = e_H = \rho(e_G) = \rho\left(gg^{-1}\right) = \rho(g)\rho\left(g^{-1}\right),$$

hence by uniqueness of inverse elements $\rho\left(g^{-1}\right) = \rho(g)^{-1}$. Thus a group homomorphism is a monoid homomorphism.

2. Suppose H is commutative. Let $\rho \in \mathbf{Hom}^{\mathrm{Grp}}(G,H)$ and $x \in M$. Then

$$\rho(x)\left(\rho(x)\right)^{-1} = e_N = e(x),$$

implying that $\mathbf{Hom}^{\mathrm{Grp}}(G,H) \subset \mathbf{Fun}(G,H)$ is a subgroup. Let $f \in \mathbf{Hom}^{\mathrm{Grp}}(G,H)$ and $x,y \in G$, then

$$f^{-1}(xy) = (f(xy))^{-1} = (f(x)f(y))^{-1} = (f(x))^{-1}(f(y))^{-1} = f^{-1}(x)f^{-1}(y) \Rightarrow f^{-1} \in \mathbf{Hom}^{\mathrm{Grp}}(G,H).$$

The following lemma follows directly from Lemmas 2.1.10 and 2.1.13

Lemma 2.2.16. *Let $\rho : G \to H$ be a group homomorphism and $I \subset G$, $J \subset H$ be subgroups. Then $\rho(I) \subset H$ and $\rho^{-1}(J) \subset G$ are subgroups. In particular, the kernel and image of a ρ are subgroups of G and H respectively.*

Proof. By Lemma 2.1.10 and Lemma 2.1.13 both sets in question are submonoids of H and G respectively. Let $g \in \rho^{-1}(J)$. Then by Remark 2.2.15,

$$\rho\left(g^{-1}\right) = \rho(g)^{-1} \in J \Rightarrow g^{-1} \in \rho^{-1}(J),$$

hence $\rho^{-1}(J)$ is a subgroup of G. Let $\rho(i) \in \rho(I)$. Then by Remark 2.2.15

$$\rho(i)^{-1} = \rho\left(i^{-1}\right) \in \rho(I),$$

hence $\rho(I)$ is a subgroup of H. □

2.2.3 Product Groups, Direct Sums & and Other Enumerated Constructions

Definition 2.2.17. Let A be a set and $\{G_\alpha\}$ a family of additive groups. We define the *direct sum of G_α over A* as

$$\bigoplus_{\alpha \in A} G_\alpha = \prod_{\alpha \in A}' G_\alpha.$$

Remark 2.2.18. Let A be a set and $\{G_\alpha\}$ a family of subgroups of G. Then

$$\sum_{\alpha\in A} g_\alpha \in G,$$

for $(g_\alpha) \in \prod'_{\alpha\in A}$ is a well-defined construction (cf. Remark 2.1.21)

Lemma 2.2.19. *Let A be a set and $\{G_\alpha\}_{\alpha\in A}$ a family of groups. The direct product of G_α over A is a group. If each G_α is additive, then so is the direct product. The restricted direct product is a subgroup of the direct product, hence the direct sum is an additive group.*

Proof. All of these constructions are monoids by Theorem 2.1.17 and Lemma 2.1.23 is follows that the direct product is a monoid, that when the groups are additive that this is also the case for the direct product and lastly the restricted direct product is a submonoid of the product monoid, hence also the direct sum when the groups are additive. For the first statement it thus suffices to check that each element of $\prod_{\alpha\in A} G_\alpha$ has an inverse. Let $(g_\alpha) \in \prod_{\alpha\in A} G_\alpha$. We define $(g_\alpha)^{-1} := (g_\alpha^{-1})$. It then follows that

$$(g_\alpha)^{-1}(g_\alpha) = (g_\alpha^{-1})(g_\alpha) = (g_\alpha^{-1} g_\alpha) = (e_\alpha) = e.$$

For the last two statements it suffices to check that $\prod'_{\alpha\in A} G_\alpha$ is closed under inversion of elements. Let $(g_\alpha) \in \prod'_{\alpha\in A} G_\alpha$. Then there $g_\alpha = e_\alpha$ for each $\alpha \in A \setminus B$ for some finite subset B of A. Hence for $\alpha \in A \setminus B$, $g_\alpha^{-1} = e_\alpha$. It follows that $(g_\alpha)^{-1} = \left(g_\alpha^{-1}\right)$. □

Proposition 2.2.20. *Let A be a set and $\{G_\alpha\}_{\alpha\in A}$ a family of groups.Then*

$$\pi_\beta : \prod_{\alpha\in A} G_\alpha \to G_\beta$$

is group homomorphism. Given a group H and a family of group homomorphisms $\{f_\alpha : H \to G_\alpha\}_{\alpha\in A}$ then the unique group homomorphism $f : H \to \prod_{\alpha\in A} G_\alpha$ (cf. Lemma 2.1.18) such that $\pi_\alpha \circ f = f_\alpha$ for every $\alpha \in A$ is group homomorphism.

Proof. π_β and f being monoid homomorphism they are automatically group homomorphisms. □

Proposition 2.2.21. *Let A be a set and $\{G_\alpha\}_{\alpha\in A}$ a family of groups. Consider a family of sets $\{H_\alpha\}_{\alpha\in A}$ such that $H_\alpha \subset G_\alpha$ is a subgroup for each $\alpha \in A$. Then*

$$\prod_{\alpha\in A} H_\alpha \subset \prod_{\alpha\in A} G_\alpha$$

is a subgroup.

Proof. By Proposition 2.1.19 it follows that $\prod_{\alpha\in A} H_\alpha$ is a submonoid. It thus suffices to check that it is closed under inversion of elements. Let $(h_\alpha) \in \prod_{\alpha\in A} H_\alpha$. Then $h_\alpha^{-1} \in H_\alpha$ for each $\alpha \in A$, hence $(h_\alpha)^{-1} = \left(h_\alpha^{-1}\right) \in \prod_{\alpha\in A} H_\alpha$. □

Proposition 2.2.22. *Let G be an additive group and $H \subset G$ a subgroup. Then H is an additive subgroup.*

Proof. This follows from Lemma 2.2.8 and Lemma 2.1.15. □

Lemma 2.2.23. *Let $\rho : G \to H$ be a group homomorphism where G is an abelian group and $J \subset G$ be a subgroup. Then $\rho(J)$ is an abelian group*

Proof. Since $\rho(J) \subset H$ is a subgroup it is a group. It remains to check that $\rho(J)$ is abelian. Let $\rho(g_1)\rho(g_2) \in \rho(J)$. Then

$$\rho(g_1)\rho(g_2) = \rho(g_1 g_2) = \rho(g_2 g_1) = \rho(g_2)\rho(g_1).$$

□

Proposition 2.2.24. *Let A be a set, G a group and $\{H_\alpha\}_{\alpha\in A}$ be a family of subgroups of G. Then*

$$\bigcap_{\alpha\in A} H_\alpha$$

is a subgroup of G

Proof. Clearly $e \in H_\alpha$ for each $\alpha \in A$, hence $e \in \bigcap_{\alpha\in A} H_\alpha$. Fix a $\beta \in A$. Let $h, h' \in \bigcap_{\alpha\in A} H_\alpha$. Then $hh', h^{-1} \in H_\alpha$ for each $\alpha \in A$, hence $hh', h^{-1} \in \bigcap_{\alpha\in A} H_\alpha$. □

Proposition 2.2.25. *Let A be a set, G an additive group and $\{H_\alpha\}_{\alpha\in A}$ be a family of subgroups of G.*

$$s : \bigoplus_{\alpha\in A} H_\alpha \to G$$
$$(h_\alpha) \mapsto \sum_{\alpha\in A} h_\alpha$$

Proof. Let $(h_\alpha), (h'_\alpha) \in \bigoplus_{\alpha\in A} H_\alpha$. For suitable $\{\alpha_1, \ldots, \alpha_n\} \subset A$, $h_\alpha = h'_\alpha = 0$ and hence $h_\alpha + h'_\alpha = 0$ for every $\alpha \in A \setminus \{\alpha_1, \ldots, \alpha_n\}$. We thus find that

$$\begin{aligned} s((h_\alpha) + (h'_\alpha)) = s((h_\alpha + h'_\alpha)) &= \sum_{\alpha\in A} \left(h_{\alpha_i} + h'_{\alpha_i}\right) = \sum_1^n \left(h_{\alpha_i} + h'_{\alpha_i}\right) = \sum_1^n h_{\alpha_i} + \sum_1^n h'_{\alpha_i} \\ &= \sum_{\alpha\in A} h_\alpha + \sum_{\alpha\in A} h'_\alpha = s((h_\alpha)) + s((h'_\alpha)). \end{aligned}$$

□

Definition 2.2.26. Let A be a set, G an additive group and $\{H_\alpha\}_{\alpha\in A}$ be a family of subgroups of G. We define *the sum of H_α over A* set

$$\sum_{\alpha\in A} H_\alpha := s\left(\bigoplus_{\alpha\in A} H_\alpha\right) = \left\{\sum_{\alpha\in A} h_\alpha : (h_\alpha) \in \bigoplus_{\alpha\in A} H_\alpha\right\},$$

which by the above proposition and Lemma 2.2.16 is a subgroup of G.

Remark 2.2.27. 1. The kernel of s is contained in

$$\left\{(h_\alpha) \in \bigoplus_{\alpha\in A} : h\alpha \in H_\alpha \cap \sum_{\beta\in A\setminus\{\alpha\}} H_\beta \text{ for each } \alpha \in A\right\}.$$

Indeed, let $(h_\alpha) \in \ker s$. Then $\sum_1^n h_{\alpha_i} = \sum_\alpha h_\alpha = 0$. Let $\alpha \in A$. If $h_\alpha = 0$, then it is trivially in $H_\alpha \cap \sum_{\beta\in A\setminus\{\alpha\}} H_\beta$. Otherwise $\alpha = \alpha_i$ for some $i \in \{1,\ldots,n\}$. Then $h_{\alpha_i} = \sum_{\beta\in A\setminus\{\alpha_i\}} h_\alpha \in \sum_{\beta\in A\setminus\{\alpha_i\}} H_\beta$, hence $h_{\alpha_i} \in H_{\alpha_i} \cap \sum_{\beta\in A\setminus\{\alpha_i\}} H_\beta$.

2. One should note that

$$\sum_{\alpha\in A} H_\alpha = \left\langle \bigcup_{\alpha\in A} H_\alpha \right\rangle.$$

Indeed, $h_\alpha \in \bigcup_{\alpha\in A} H_\alpha$ for every $\alpha \in A$, hence

$$\sum_{\alpha\in A} h_\alpha = \sum_1^n h_\alpha \in \left\langle \bigcup_{\alpha\in A} H_\alpha \right\rangle.$$

Let $\sum_1^n m_i h_{\alpha_i} \in \langle \bigcup_{\alpha\in A} H_\alpha \rangle$ where $m_i \geq 0$, $\alpha_i \in A$, $h_{\alpha_i} \in H_{\alpha_i}$. For each i we then have that

$$m_i h_{\alpha_i} = \sum_{j=1}^{m_i} h_{\alpha_i} \in H_{\alpha_i},$$

Hence upon putting $h'_\alpha = 0$ for $\alpha \in A \setminus \{\alpha_1,\ldots,\alpha_n\}$ and $h'_{\alpha_i} = m_i h_{\alpha_i}$, implying

$$\sum_1^n m_i h_{\alpha_i} = \sum_{\alpha\in A} h'_\alpha \in \sum_{\alpha\in A} H_\alpha.$$

2.2.4 Quotient Groups

Definition 2.2.28. Let G be a group and $X, Y \subset G$ we then define

$$XY := \circ(X \times Y) = \{xy \in G : (x,y) \in X \times Y\},$$

and

$$X^{-1} := i(X) = \left\{x^{-1} \in G : x \in X\right\}.$$

Remark 2.2.29. One easily sees for $X, Y, Z \subset G$ that $X(YZ) = (XY)Z$ and that $(XY)^{-1} = Y^{-1}X^{-1}$.

Definition 2.2.30. Let G be a group, $H \subset G$ be a subgroup and $g \in G$. The *left coset of H with respect to g* is defined to be the set

$$gH := \{g\}H = \{gh : h \in H\}$$

The *right coset of H with respect to g* is defined to be the set

$$Hg := H\{g\} = \{hg : h \in H\}.$$

Remark 2.2.31. Note that clearly $hH = H = Hh$ for any $h \in H$. If $gH \neq H$, then $gh \notin H$ for some $h \in H$, meaning $g \notin H$. Since $eg = g = ge$, it also follows that $g \in gH$ and $g \in Hg$ for every $g \in G$. It is also easy to check that $H^{-1} = H$.

Proposition 2.2.32. *Let G be a group and $H \subset G$ a subgroup. Then the sets*

$$\sim_l := \{(g_1, g_2) \in G \times G : g_1 H = g_2 H\} \ \ \& \ \ \sim_r := \{(g_1, g_2) \in G \times G : Hg_1 = Hg_2\}$$

define equivalence relations. We define $G/H := G/\sim_l$ and $G \setminus H := G/\sim_r$.

Proof. We only check the left case, since the right case is dual. Let $g_1, g_2, g_3 \in G$. Obviously $g_1 H = g_1 H$, hence $g_1 \sim_l g_1$. Suppose $g_1 \sim_l g_2$. Then $g_1 H = g_2 H$, hence $g_2 H = g_1 H$, meaning $g_2 \sim_l g_1$. Suppose $g_1 \sim_l g_2$ and $g_2 \sim_l g_3$. Then $g_1 H = g_2 H = g_3 H$, hence $g_1 \sim_l g_3$. ☐

Lemma 2.2.33. *Let G be a group and $H \subset G$. Then for $g, g' \in G$*

$$g \sim_l g' \iff g^{-1} g' \in H.$$

Proof. Indeed,

$$g \sim_l g' \iff gH = g'H \iff g^{-1}(gH) = g^{-1}(g'H) \iff H = \left(g^{-1}g\right)H = \left(g^{-1}g'\right)H$$
$$\iff g^{-1}g' \in H,$$

where the last bi-implication follows from Remark 2.2.31. ☐

Proposition 2.2.34. *A group G with a subgroup $H \subset$ is the disjoint union of the elements of G/H respectively $G \setminus H$.*

Proof. We check the left case. The right case is dual. Since $g \in gH$ for every $g \in G$, it follows that

$$G = \bigcup_{gH \in G/H} gH.$$

Let $g_1, g_2 \in G$. Suppose $g_1 H \cap g_2 H \neq \emptyset$. Then there is an element $x \in g_1 H \cap g_2 H$, hence $g_1 h_1 = x = g_2 h_2$ for suitable $h_1, h_2 \in H$. This implies that

$$g_1^{-1} g_2 = h_1 h_2^{-1} \in H \Rightarrow g_1 H = g_2 H.$$

Thus if $g_1 H \neq g_1 H$, then $g_1 H \cap g_2 H = \emptyset$. ☐

Definition 2.2.35. A subgroup $H \subset G$ is *normal* if

$$gNg^{-1} = N$$

for every $g \in G$.

Remark 2.2.36. Note that

$$gNg^{-1} = N \iff gN = gN(g^{-1}g) = (gNg^{-1})g = Ng,$$

Hence any subgroup of an abelian group is normal. Furthermore $\sim_l = \sim_r$. Thus we may define $\sim = \sim_l = \sim_r$. Let $X \subset G$. Then $XN = NX$. Indeed, if $xn \in XN$, then $xn \in xN = Nx$, hence $XN \subset NX$. The other inclusion is shown in a similar way.

Lemma 2.2.37. *The kernel of a group homomorphism $\rho : G \to H$ is a normal subgroup of G.*

Proof. Let $g \in G$ and $k \in \textbf{ker } \rho$. Then

$$\rho\left(gkg^{-1}\right) = \rho(g)\rho(k)\rho(g)^{-1} = \rho(g)e\rho(g)^{-1} = e,$$

hence $g(\textbf{ker } \rho)g^{-1} \subset \textbf{ker } \rho$. Conversely $k \in g(\textbf{ker } \rho)g^{-1}$ since $k = g\left(g^{-1}kg\right)g^{-1}$ and $g^{-1}kg \in \textbf{ker } \rho$ by the above computation. □

Proposition 2.2.38. *Let G be a group and $H, N \subset G$ be subgroup, where N is normal. Then $HN, NH \subset G$ are subgroups of G.*

Proof. Let $h_1n_1, h_2n_2 \in HN$. Then

$$h_1n_1h_2n_2 \in (HN)(HN) = (NH)(HN) = (N(HH))N = (NH)N = (HN)N = H(NN) = HN.$$

Since $e \in H$ and $e \in N$, $ee \in HN$. Thus HN is a subgroup of G. Since $HN = NH$, it follows that NH is also a subgroup of G. □

Proposition 2.2.39. *Let G be a group and $H \subset G$ a normal subgroup. Then the operation*

$$\cdot : G/H \times G/H \to G/H$$

given by $(g_1H)(g_2H) := g_1g_2H := (g_1g_2)H$ for $g_1H, g_2H \in G/H$ is well-defined and $(G/H, \cdot)$ is a group.

Proof. We first need to check that the group operation is well-defined. Let $g_1, g_2 \in G$. We need to check that $(g_1H)(g_2H) = (g_1g_2)H$. Let $g_1h_1g_2h_2 \in g_1Hg_2H$. Then $h_1g_2 \in Hg_2 = g_2H$, hence $h_1g_2h_2 \in g_2Hh = g_2H$, hence $g_1(h_1g_2h_2) = g_1(g_2H) = (g_1g_2)H$. If $g_1g_2h \in g_1g_2H$, then $g_1g_2h = g_1eg_2h \in (g_1H)(g_2H)$. Thus the operation is well-defined, since if $(g_1H, g_2H) = (g'_1H, g'_2H)$, then trivially

$$(g_1g_2)H = (g_1H)(g_2H) = (g'_1H)(g'_2H) = (g'_1g'_2)H.$$

For $g_1H, g_2H, g_3H \in G/H$, it follows by Remark 2.2.29 that

$$(g_1H)((g_2H)(g_3H)) = (g_1Hg_2H)(g_3H).$$

Define the neutral element with respect to $\cdot$ to be eH. Indeed

$$(eH)(gH) = (eg)H = gH,$$

for every $gH \in G/H$. We inverse element of $gH \in G/H$ to be $(g^{-1}H) = H^{-1}g^{-1} = Hg^{-1} = g^{-1}H$. Indeed

$$(g^{-1}H)(gH) = (g^{-1}g)H = eH.$$

□

Corollary 2.2.40. *If $(G, +)$ is an additive group and $H \subset G$ is a subgroup, then $(G/H, +)$ (where + is defined as in the above proposition) is an additive group*

Proof. By Remark 2.2.36 H is normal and by the above proposition G/H is a group. Let $g_1 + H, g_2 + H \in G/H$ be given. Then

$$(g_1 + H) + (g_2 + H) = (g_1 + g_2) + H = (g_2 + g_1) + H = (g_2 + H) + (g_1 + H).$$

□

Lemma 2.2.41. *Let G be a group and N a normal subgroup. Let $N \subset H \subset G$ a subgroup. Then $N \subset H$ is normal and $H/N \subset G/N$ is a subgroup.*

Proof. For every $h \in H$, $hNh^{-1} = N$, since $h \in G$. Let $h_1N, h_2N \in H/N$. Then since $h_1h_2 \in H$,

$$(h_1N)(h_2N) = h_1h_2N \in H/N.$$

Furthermore since $e \in H$, $eN \in H/N$. □

Proposition 2.2.42. *Let G be a group and N a normal subgroup. Then $S = \{H \subset G : H$ is a subgroup of $G, N \subset H\}$ is in one-to-one correspondence with the set $S' = \{K \subset G/N : K$ a subgroup of $G/N\}$. Any subgroup $K \subset G/N$ is of the form H/N for some $H \in S$.*

Proof. We show that $u : S \to S'$, $H \mapsto H/N$ is a bijection. This map is well-defined by the above lemma. For $K \in S'$, let $H(K) = \{g \in G : gN \in K\}$. We check that $H(K)$ is a subgroup G containing N. Let $h_1, h_2 \in H(K)$. Then $h_1N, h_2N \in K$, hence $h_1h_2N \in K$, implying $h_1h_2 \in H(K)$. Clearly $eN \in K$, hence $e \in H(K)$. Let $n \in N$. Then $nN = eN \in K$, hence $n \in H(K)$. Then the map $u' : S' \to S$, $K \to H(K)$ is well-defined. We check that u and u' are mutual inverses. Let $K \in S'$. We need to check that $uu'(K) = H(K)/N = K$. Let $k \in K$, then $k = gN$ for some $g \in G$, then $g \in H(K)$, hence $k = gN \in H(K)/N$. Let $hN \in H(K)/N$. Then by definition $hN \in K$. Let $H \in S$. Then we need to check that $u'u(H) = H(H/N) = H$. Let $h \in H(H/N)$, then $hN \in H/N$, hence $h \in H$. Let $h \in H$. Then $hN \in H/N$, hence $h \in H(H/N)$. ☐

Proposition 2.2.43. *Let G be a group and $N \subset G$ a normal subgroup. The surjection $\pi : G \to G/N, g \mapsto gN$ is a group map*

Proof. Let $g_1, g_2 \in G$. Then $\pi(g_1g_2) = g_1g_2N = g_1Ng_2N = \pi(g_1)\pi(g_2)$. ☐

2.3 Rings

2.3.1 Definition & Basic Properties

Definition 2.3.1. A *ring (with unity)* is a set R with operations $+ : R \times R \to R$ and $\cdot : R \times R \to R$ called *multiplication* such that $(R, +)$ is an additive group, $(R, \cdot)$ is a monoid and for $r_1, r_2, r_3 \in R$

$$r_1(r_2 + r_3) = r_1r_2 + r_1r_3 \ \& \ (r_1 + r_2)r_3 = r_1r_3 + r_2r_3.$$

We denote the neutral element with respect to multiplication by 1. The data specifying a ring is often written $(R, +, \cdot)$.

Lemma 2.3.2. *Let R be a ring and $r \in R$. The following identities are true for rings*

1. $0 \cdot r = 0$.
2. $(-1)r = -r$, $r(-1) = -r$.

Proof. 1. follows from the following computation.

$$0r = 0r + 0 = 0r + r - r = 0r + 1r - r = (0+1)r - r = 1r - r = r - r = 0$$

2. follows from the following computation.

$$(-1)r = (-1)r + 0 = (-1)r + r - r = (-1)r + 1r - r = (-1+1)r - r = 0r - r = 0 - r = -r$$

the other statement is proven similarly. ☐

Definition 2.3.3. Let R be a ring. If $(R,\cdot)$ is a commutative monoid, then R is called a *commutative ring.*

Definition 2.3.4. Let R be a ring. A *subring* is a subset $S \subset R$ such that S is a subgroup of $(R,+)$ and a submonoid of $(R,\cdot)$.

Remark 2.3.5. $(S,+,\cdot)$ is a ring. Indeed, clearly $(S,+)$ is an additive group since $S \subset R$ is a subgroup and $(S,\cdot)$ is a monoid since $S \subset R$ is a submonoid. Let $r_1,r_2,r_3 \in S$, then since $S \subset R$,

$$r_1(r_2+r_3) = r_1r_2 + r_1r_2 \;\&\; (r_1+r_2)r_3 = r_1r_3 + r_2r_3.$$

One should also note that R is commutative then S is commutative

Example 2.3.6. 1. $(\mathbb{Z},+,\cdot)$, $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$, $(\mathbb{C},+,\cdot)$ are rings.

2. Given a ring R we may form the opposite ring $(R^{(\mathrm{op})},+,*)$ where $(R^{(\mathrm{op})},+) = (R,+)$ and multiplication is defined by $r * r' = r' \cdot r$ for $r,r' \in R$. checking that this is a ring is easy. Clearly $(R^{(\mathrm{op})},+)$ is an additive group. Let $r_1,r_2,r_3 \in R^{(\mathrm{op})}$. Then
$$r_1 * (r_2 * r_3) = (r_3r2)r_1 = r_3(r_2r_1) = (r_1 * r_2) * r_3$$
and
$$r_1 * 1 = 1r_1 = r_1 = r_11 = 1 * r_1$$
and lastly
$$r_1 * (r_2+r_3) = (r_2+r_3)r_1 = r_2r_1 + r_3r_1 = r_1 * r_2 + r_1 * r_3,$$
where last identity to be checked is omitted as it is dual to the one above. One also easily verifies that $\left(R^{(\mathrm{op})}\right)^{(\mathrm{op})} = R$

3. For a non-empty set X and a ring R, the set $\mathbf{Fun}(X,R)$ is a monoid and an additive group with respect to multiplication and addition defined earlier. One easily verifies that it is also a ring.

4. For rings R,S, if $\mathbf{Hom}^{\mathrm{Ring}}(R,S) \neq \emptyset$ is never a subring of $\mathbf{Fun}(R,S)$. Indeed, note that the zero map is never a ring homomorphism since it maps 1 to 0.

2.3.2 Morphisms of Rings

Definition 2.3.7. Let R, S be rings. A map $\sigma : R \to S$ is called a *ring homomorphism/map of rings/morphism of rings* if σ is a group homomorphism between $(R,+)$ and $(S,+)$ and a monoid homomorphism between $(R,\cdot)$ and $(S,\cdot)$. The set of ring homomorphisms between R to S is denoted $\mathbf{Hom}^{\mathrm{Ring}}(R,S)$.

Here are some examples of rings

Lemma 2.3.8. *Let $\sigma : R \to S$ be a ring homomorphism and $T \subset R$, $U \subset S$ be subrings. Then $\sigma(T) \subset S$ and $\sigma^{-1}(U)$ are subrings. If T is commutative then so is $\sigma(T)$.*

Proof. Prior lemmas ensure that these sets are appropriate additive subgroups and submonoids. □

2.3.3 Product Rings

Proposition 2.3.9. *Let A be a set, $\{R_\alpha\}_{\alpha\in A}$ a family of rings. Then $(\prod_{\alpha\in A} R_\alpha,\cdot)$ is a monoid and $(\prod_{\alpha\in A} R_\alpha,+)$ an additive group by Theorem 2.1.17 resp. Proposition 2.2.19. In addition $(\prod_{\alpha\in A} R_\alpha,+,\cdot)$ is a ring.*

Proof. It remains to check that multiplication distributes over addition. Let $(r_\alpha),(r'_\alpha),(r''_\alpha) \in \prod_{\alpha\in A} R_\alpha$. Then

$$\begin{aligned}(r_\alpha)\big((r'_\alpha)+(r''_\alpha)\big) &= (r_\alpha)(r'_\alpha+r''_\alpha) = (r_\alpha(r'_\alpha+r''_\alpha)) = (r_\alpha r'_\alpha + r_\alpha r''_\alpha) = (r_\alpha r'_\alpha)+(r_\alpha r''_\alpha)\\ &= (r_\alpha)(r'_\alpha)+(r_\alpha)(r''_\alpha).\end{aligned}$$

□

Remark 2.3.10. The above ring is called *the product ring (of $\{R_\alpha\}_{\alpha\in A}$ over A).*

Corollary 2.3.11. *The direct sum of rings is a subring of the direct product.*

Proof. This follows from Lemma 2.1.23 and Lemma 2.2.19. □

Proposition 2.3.12. *Let A be a set and $\{R_\alpha\}_{\alpha\in A}$ a family of rings. Then*

$$\pi_\beta : \prod_{\alpha\in A} R_\alpha \to R_\beta$$

is ring homomorphism. Given a ring S and a family of ring homomorphisms $\{f_\alpha : S \to R_\alpha\}_{\alpha\in A}$ then the unique group and monoid homomorphism $f : S \to \prod_{\alpha\in A} R_\alpha$ (cf. Lemma 2.1.18 and Proposition 2.2.20) such that $\pi_\alpha \circ f = f_\alpha$ for every $\alpha \in A$ is a ring homomorphism.

Proof. This follows immediately from the fact both π_β and f are both group and monoid homomorphisms. □

2.3.4 The Set of Integers: $\mathbb{Z}$

Definition 2.3.13. For $(a,b),(c,d) \in \mathbb{N}$ we define $(a,b) \sim (c,d)$ if $a+d = b+c$. On easily checks that this is an equivalence relation. We define

$$\mathbb{Z} := \mathbb{N}^2/\sim .$$

Proposition 2.3.14. *On $\mathbb{Z}$ we define*

$$[(a,b)]+[(c,d)] := [(a+c,b+d)]$$

and

$$[(a,b)][(c,d)] := [(ac+bd,ad+bc)].$$

Moreover we define $0 := [(0,0)]$, $-[(a,b)] := [(b,a)]$ and $1 := [(1,0)]$. With these definitions, $\mathbb{Z}$ becomes a commutative ring. The $\{[(a,0)] \in \mathbb{Z} : a \in \mathbb{N}\}$ is a sub semi-ring isomorphic to $\mathbb{N}$.

Proof. Suppose first that $([(a,b)],[(c,d)]) = ([(x,y)],[(v,w)])$. Then

$$a+y = b+x, \quad c+w = d+v$$

hence

$$(a+c)+(y+w) = (a+y)+(c+w) = (b+x)+(d+v) = (b+d)+(x+v) \Rightarrow [(a+c,b+d)] = [(x+v,y+w)].$$

So addition is well-defined. We also have that We check that $\mathbb{Z}$ is a group. Associativity of addition on $\mathbb{Z}$ readily follows from associativity of addition on $\mathbb{N}$. Let $[(a,b)],[(c,d)] \in \mathbb{Z}$ be arbitrary. Then

$$[(a,b)]+[(0,0)] = [(a+0,b+0)] = [(a,b)]$$

and

$$[(a,b)]+(-[(a,b)]) = [(a,b)]+[(b,a)] = [(a+b,a+b)] = [(0,0)] = 0$$

and

$$[(a,b)]+[(c,d)] = [(a+c,b+d)] = [(c+a,d+b)] = [(c,d)]+[(a,b)]$$

hence $\mathbb{Z}$ is a commutative group. It is easy to check that $\{[(a,0)] \in \mathbb{Z} : a \in \mathbb{N}\}$ is a submonoid isomorphic to $\mathbb{N}$. Suppose again $([(a,b)],[(c,d)]) = ([(x,y)],[(v,w)])$. We find that

$$\begin{aligned} ac+bd+xw+yv+yc+xd+xc+yd = \\ &= c(a+y)+d(b+x)+x(c+w)+y(d+v) \\ &= c(b+x)+d(a+y)+x(d+v)+y(c+w) \\ &= ad+bc+xv+yw+yc+xd+xc+yd. \end{aligned}$$

implying that, using a fact from group theory,

$$ac+bd+xw+yv=ad+bc+xv+yw \Rightarrow [(ac+bd,ad+bc)]=[(xv+yw,xw+yv)].$$

We now find that

$$\begin{aligned}([(a,b)][(c,d)])[(e,f)] &= [(ac+bd,ad+bc)][(e,f)]\\ &= [(ace+bde+adf+bcf,acf+bdf+ade+bce)]\\ &= [(a(ce+df)+b(cf+de),a(cf+de)+b(ce+df)]\\ &= [(a,b)][(ce+df,cf+de)]=[(a,b)]([(c,d)][(e,f)]).\end{aligned}$$

and

$$[(a,b)][(1,0)]=[(a+b\cdot 0,a\cdot 0+b)]=[(a,b)]$$

and easily we check that

$$[(a,b)][(c,d)]=[(c,d)][(a,b)].$$

Lastly,

$$\begin{aligned}[(a,b)]([(c,d)]+[(e,f)]) = [(a,b)][(c+e,d+f)] &= [(ac+ae+bd+bf,ad+af+bc+be)]\\ &= [(ac+bd,ad+bc)]+[(ae+bf,af+be)]\\ &= [(a,b)][(c,d)]+[(a,b)][(e,f)],\end{aligned}$$

making $\mathbb{Z}$ a commutative ring. One readily verifies that $\{[(a,0)]\in\mathbb{Z}:a\in\mathbb{N}\}$ is a submonoid of $\mathbb{Z}$ with respect to multiplication. The isomorphism from $\mathbb{N}$ is given by $a\mapsto[(a,0)]$. $\square$

2.4 Modules

2.4.1 Initial Definitions, Basic Properties & Constructions

Definition 2.4.1. Let R be a ring. A *left R-module* is an additive group $(M,+)$ with a *left scalar multiplication* $\cdot : R\times M\to M$, where $rm := r\cdot m := \cdot(r,m)$ for $(r,m)\in R\times M$ satisfying the following axioms

1. For every $r\in R$, $m,m'\in M$,

$$r(m+m')=rm+rm'.$$

2. For every $r,r'\in R$, $m\in M$,

$$(r+r')m=rm+r'm.$$

3. For every $r, r' \in R$, $x \in M$,
$$(rr')m = r(r'm).$$

4. For every $m \in M$,
$$1m = m.$$

To emphasise that a module M is a left R-module, we may write ${}_RM := M$.
A *right R-module* is an additive group $(M,+)$ with a *right scalar multiplication* $\cdot : M \times R \to M$, where $mr := m \cdot r := \cdot(m,r)$, satisfying axioms dual to ones for left scalar multiplication. To emphasise that a module M is a right R-module, we may write $M_R := M$
Let S be a ring. An (R,S)-bimodule is an additive group $(M,+)$, that is a left R-module and a right S-module satisfying

$$(rm)s = r(ms),$$

for every $r \in R, s \in S$, $m \in M$. To emphasise that a module M is an (R,S)-bimodule, we may write ${}_RM_S := M$.

Lemma 2.4.2. *Let M be an additive group and R a ring. Then M is a left R-module if and only if M is a right $R^{(op)}$-module.*

Proof. "$\Rightarrow$": 'pose M is a left R-module. We define a right scalar multiplication of $R^{(\mathrm{op})}$ on M by $mr = rm$. Checking the first 3 axioms is straight forward. For the 4th axiom, let $r_1, r_2 \in R^{(\mathrm{op})}$ and $m \in M$ be given. Then

$$m(r_1 * r_2) = (r_2r_1)m = r_2(r_1m) = r_2(mr_1) = (mr_1)r_2.$$

"$\Leftarrow$": This is very similar. ☐

The consequence of the above lemma is that any theorem about right R-modules that is true for left R-modules, can be proven by applying said left case theorem to $R^{(\mathrm{op})}$. Using the fact that $R^{(\mathrm{op})} = R$ when R is commutative, implies that left/right R-modules coincide. In this case left/right R-modules will be referred to simply as *R-modules*.
For a field K, we call a K-module a *vector space over K* We give simple initial examples of modules.

Definition 2.4.3. Let M be a left/right R-module. A left/right R-submodule is a subset $N \subset M$ such that N is a subgroup of $(M,+)$ and for every $r \in R$, $n \in N$ we have that $rn \in N$ resp. $nr \in N$.

Remark 2.4.4. A left/right R-submodule $N \subset M$ is a left/right R-module. Indeed $(N,+)$ is group since it is a subgroup of $(M,+)$. N being closed under left/right scalar multiplication ensures that $\cdot := \cdot|_{R\times N}: R \times N \to N$ respectively $\cdot := \cdot|_{N\times R}: N \times R \to N$ are well-defined and M being a left/right R-module, these left/right actions respect the axioms for left/right scalar multiplication. If M is a (R,S)-bimodule and N is a left R-submodule and a right S-submodule of M then it is also an (R,S)-bimodule. Indeed, if $r \in R$, $n \in N$ and $s \in S$, then since $n \in M$, $r(ns) = (rn)s$.

Example 2.4.5.

1. Let $(G,+)$ be an additive group. Then G is a $\mathbb{Z}$-module under the left/right scalar multiplication

$$ng = \sum_{1}^{n} g \;\&\; gn = \sum_{1}^{n} g.$$

Under this definition $gn = ng$ hence if G is a left $\mathbb{Z}$-module, it is automatically a $\mathbb{Z}$-module. Let $n,m \in \mathbb{Z}$ and $g,g' \in G$. Then

1. $n(g+g') = \sum_{1}^{n}(g+g') = \sum_{1}^{n} g + \sum_{1}^{n} g' = ng + ng'$
2. $(n+m)g = \sum_{1}^{n+m} g = \sum_{1}^{n} g + \sum_{n+1}^{n+m} g = ng + \sum_{1}^{m} g = ng + mg$
3. $n(mg) = n\sum_{1}^{m} g = \sum_{1}^{n}\sum_{1}^{m} g = \sum_{1}^{nm} g = (nm)g$
4. $1g = \sum_{1}^{1} g = g$

To prove the second to last equality 3. one should really use induction in m. Note that the induction start uses 4.

2. Let $(R,+,\cdot)$ be a ring. Then $(R,+)$ is an additive group, which becomes an (R,R)-bimodule under the action $rx := r\cdot x$ and $xs := x \cdot s$ for $r,s,x \in R$.

3. A left/right R-submodule of $I \subset R$ is called a *left/right ideal in* R. If I is an (R,R)-bimodule, it is called a *both-sided ideal in* R. If R is commutative a left/right ideal is simply referred to as an *ideal in* R.

4. Let A be a set and R a ring, then $\prod_{\alpha\in A} R$ is an R-module, under the left/right scalar multiplication given by $a(r_\alpha) := (ar_\alpha)/(r_\alpha)a = (r_\alpha a)$ for $a \in R$ and $(r_\alpha) \in \prod_{\alpha\in A} R$. This is easily checked using 2. together with the fact that $(a)(r_\alpha) = a(r_\alpha)$ and $(r_\alpha)(a) = (r_\alpha)a$ for every $a \in R$, $(r_\alpha) \in \prod_{\alpha\in A} R$. It in particular follows that the matrix ring is an R-module.

5. Let R be a ring and $I \subset R$ an ideal. Then R/I is an R-module, when equipped with the left/right scalar multiplication $r(a+I) := (ra+I)$. One sees this from the fact that $(r+I)(a+I) = r(a+I)$.

6. Let S be a ring and R a subring of S. Then $R \times S \ni (r,s) \mapsto rs \in S$ defines a left scalar multiplication of R on S, hence S is a left R-module. One turns S into a right R-module via $S \times R \ni (s,r) \mapsto sr \in S$

Lemma 2.4.6. *Let R be a ring and M a left/right R-module.*

1. $0m = 0$ *for every* $m \in M$.

2. $(-1)m = -m$ *for every* $m \in M$.

3. $r(-m) = -rm$ *for every* $r \in R$, $m \in M$.

4. $r0 = 0$ *for every* $r \in R$.

Proof. 1. Really this is just a generalization of Lemma 2.3.2 1. Indeed,

$$0m = 0m + m - m = 0m + 1m - m = (0+1)m - m = 1m - m = m - m = 0.$$

2. Really this is just a generalization of Lemma 2.3.2 2. Indeed,

$$(-1)m = (-1)m + m - m = (-1)m + 1m - m = (-1+1)m - m = 0m - m = 0 - m = -m$$

3. Indeed,

$$r(-m) = r((-1)m) = (r(-1))m = ((-1)r)m = -rm.$$

4. Indeed,

$$r0 = r(0-0) = r0 + r(-0) = r0 - r0 = (r-r)0 = 0 \cdot 0 = 0.$$

□

Definition 2.4.7. Let R be a ring. An element m of a left/right R-module M is called a torsion *torsion element* if there is an $r \in R$ that is not a left/right zero-divisor satisfying $rm = 0$ (resp. $mr = 0$). A module is a torsion module if every element is a torsion element and *torsion free* if the only torsion element is 0.

Remark 2.4.8. An example of a torsion free module is a domain R.

Definition 2.4.9. Let ${}_RM, {}_RN$ be left R-modules. A map $\rho : M \to N$ is a *left R-module homomorphism/map of left R-modules/morphism of left R-modules* if for every $r \in R$, $m, m' \in M$,

$$\rho(rm + m') = r\rho(m) + \rho(m').$$

right R-module homomorphisms are defined in a dual manner. *(R,S)-bimodule homomorphisms* is a map that is both a left R-module homomorphism and a right S-module homomorphism.

Remark 2.4.10. A left/right/bimodule module homomorphisms $(M,+) \to (N,+)$ are automatically group homomorphisms. Similarly for every $r \in R$ and $m \in M$,

$$\rho(rm) = r\rho(m),$$

which is seen by setting $m' = 0$ when $\rho : M \to N$ is a left R-module. This is also true from the right if ρ was a right R-module.

Lemma 2.4.11. *Let M,N be additive groups, R a ring and $\rho : M \to N$ a group homomorphism. Then ρ is a left R-module homomorphism if and only if ρ is a right $R^{(op)}$-module.*

Proof. "$\Rightarrow$": We make M and N right $R^{(\mathrm{op})}$-modules as in Lemma 2.4.2. Let $r \in R^{(\mathrm{op})}$ and $m \in M$. It is sufficient to check that $\rho(mr) = \rho(m)r$. Indeed,

$$\rho(mr) = \rho(rm) = r\rho(m) = \rho(m)r.$$

"$\Leftarrow$": This is proven by using $\left(R^{(\mathrm{op})}\right)^{(\mathrm{op})} = R$ and applying "$\Rightarrow$". ☐

The consequence of the above lemma is that we get a way of automatically check a theorem for R-module homomorphisms, whenever we have a proof of the left case. This akin to what we gained in Lemma 2.4.2.

Lemma 2.4.12. *If $\rho : M \to N$ is a left R-module/right S-module homomorphism, then for a left R-submodule/right S-module $L \subset M$, $\sigma(L) \subset N$ is a left R-submodule/ right R-submodule. If M and N are (R,S)-bimodules and ρ is an (R,S)-bimodule homomorphism, then $\sigma(L)$ is an (R,S)-bimodule. The two statements present are thus in particular true for the image of ρ.*

Proof. We only check left case, as the right case is dual. We already know that $\sigma(L)$ is a subgroup. Let $r \in R$ and $\sigma(l) \in \sigma(L)$. Then

$$r\sigma(l) = \sigma(rl) \in \sigma(L).$$

Let $r \in R$, $\sigma(l) \in \sigma(L)$, $s \in S$. Then

$$r(\sigma(l)s) = r\sigma(ls) = \sigma(r(ls)) = \sigma((rl)s) = \sigma(rl)s = (r\sigma(l))s.$$

☐

Lemma 2.4.13. *Let R,S be rings, A a set and $\{M_\alpha\}_{\alpha\in A}$ be a family of left R-modules/right S-modules/(R,S)-bimodules. Then*

$$\prod_{\alpha\in A} M_\alpha$$

is a left R-modules/right S-module/(R,S)-bimodule.

Proof. Note that by Proposition 2.2.21 the direct product of a family of left/right modules is an additive group. We check the left case. Let $r_1, r_2 \in R$, $(m_\alpha),(m'_\alpha) \in \prod_{\alpha\in A} M_\alpha$. Then

1. $(r_1+r_2)(m_\alpha) = ((r_1+r_2)m_\alpha) = (r_1 m_\alpha + r_2 m_\alpha) = (r_1 m_\alpha) + (r_2 m_\alpha) = r_1(m_\alpha) + r_2(m_\alpha)$.
2. $r_1((m_\alpha)+(m'_\alpha)) = r_1(m_\alpha + m'_\alpha) = (r_1 m_\alpha + r_1 m'_\alpha) = (r_1 m_\alpha) + (r_1 m'_\alpha) = r_1(m_\alpha) + r_1(m'_\alpha)$.
3. $1(m_\alpha) = (1 m_\alpha) = (m_\alpha)$.
4. $(r_1 r_2)(m_\alpha) = ((r_1 r_2) m_\alpha) = (r_1(r_2 m_\alpha)) = r_1(r_2 m_\alpha) = r_1(r_2(m_\alpha))$.

Suppose $\{M_\alpha\}_{\alpha\in A}$ is a family of (R,S)-modules and let $r\in R$, $s \in S$. Then

$$r((m_\alpha)s) = r(m_\alpha s) = (r(m_\alpha s)) = ((r m_\alpha)s) = (r m_\alpha)s = (r(m_\alpha))s.$$

□

Proposition 2.4.14. *Let A be a set, R a ring and $\{M_\alpha\}_{\alpha\in A}$ a family of left/right modules. Then*

$$\pi_\beta : \prod_{\alpha\in A} M_\alpha \to M_\beta$$

is a left/right R-module homomorphism. Given a left/right R-module N and a family of left/right R-module homomorphisms $\{f_\alpha : N \to M_\alpha\}_{\alpha\in A}$ then the unique group homomorphism $f : N \to \prod_{\alpha\in A} M_\alpha$ (cf. Proposition 2.2.20) such that $\pi_\alpha \circ f = f_\alpha$ for every $\alpha \in A$ is a left/right R-module homomorphism.

Proof. Let $r \in R$ and $(m_\alpha) \in \prod_{\alpha\in A} M_\alpha$, $n \in N$. Then for $\beta \in A$,

$$\pi_\beta(r(m_\alpha)) = \pi_\beta((r m_\alpha)) = r m_\beta = r\pi_\beta((m_\alpha)).$$

Furthermore, we have that

$$f(rn) = (f_\alpha(rn)) = (r f_\alpha(n)) = r(f_\alpha(n)) = r f(n).$$

□

Lemma 2.4.15. *Let R be a ring, M a left/right R-module, A a set and $\{M_\alpha\}_{\alpha\in A}$ a family of left/right R-submodules. Then*

$$\bigcap_{\alpha\in A} M_\alpha$$

is a left/right R-submodule.

Proof. From Proposition 2.2.24 we already know that $\bigcap_{\alpha\in A} M_\alpha$ is an additive subgroup. Let $r \in R$ and $m \in \bigcap_{\alpha\in A} M_\alpha$ then $rm \in M_\alpha$ for every $\alpha \in A$, meaning $rm \in \bigcap_{\alpha\in A} M_\alpha$. □

Proposition 2.4.16. *Let A be a set, R a ring and $\{M_\alpha\}_{\alpha\in A}$ a family of left/right R-modules. Then*

$$\bigoplus_{\alpha\in A} M_\alpha$$

is a submodule of $\prod_{\alpha\in A} M_\alpha$.

Proof. We already know it to be an additive subgroup. Let $r \in R$ and $(m_\alpha) \in \bigoplus_{\alpha\in A} M_\alpha$. Then for some finite subset $B \subset A$, $m_\alpha = 0$ for every $\alpha \in A \setminus B$. Hence $rm_\alpha = 0$ for every $\alpha \in A \setminus B$. It thus follows that $r(m_\alpha) = (rm_\alpha) \in \bigoplus_{\alpha\in A} M_\alpha$. □

Lemma 2.4.17. *Let R be a ring M be a left/right R-module. Let $N \subset M$ be a left/right R-submodule. Then M/N is a left/right submodule under the left/right scalar multiplication $r(m+N) := rm+N$ resp. $(m+N)r := mr+N$. If M is an (R,S)-bimodule for some ring S and N is a left R-submodule and a right S-submodule. Then M/N is an (R,S)-bimodule.*

Proof. Let $r_1, r_2 \in R$, $m+N, m'+N \in M/N$. Then

1. $(r_1+r_2)(m+N) = (r_1+r_2)m+N = (r_1m+r_2m)+N = (r_1m+N)+(r_2m+N)$ $= r_1(m+N)+r_2(m+N)$.
2. $r_1((m+m')+N) = r_1(m+m')+N = (r_1m+r_1m')+N = (r_1m+N)+(r_1m'+N)$ $= r_1(m+N)+r_1(m'+N)$.
3. $1(m+N) = 1m+N = m+N$
4. $(rr')(m+N) = (rr')m+N = r(r'm)+N = r(r'm+N) = r(r'(m+N))$

Suppose M is an (R,S)-bimodule. Let $r \in R$, $s \in S$. Then

$$r((m+N)s) = r(ms+N) = r(ms)+N = (rm)s+N = (rm+N)s = (r(m+N))s.$$

□

Corollary 2.4.18. *The canonical surjective group map $\pi : M \to M/N$ is a left/right module map*

Proof. Let $r \in R$, $m \in M$. Then $\pi(rm) = rm + N = r(m+N) = r\pi(m)$. □

Lemma 2.4.19. *Let R be a ring and M a left/right R-module and $N \subset M$ a submodule. Then there is one-to-one correspondence between the sets*

$$U = \{L \subset M : L \text{ is a submodule of } M \text{ containing } N\}$$

and

$$U' = \{K \subset M/N : K \text{ is a submodule of } M/N\}.$$

Proof. This is a corollary of Proposition 2.2.42. Note that by Lemma 2.4.17, $u : U \to U'.L \mapsto L/N$ is well-defined- Note that $U \subset S$ and $U' \subset S'$. Let $K \in U'$, we check that $L(K) := \{m \in M : m + N \in K\}$ is a submodule. Let $r \in R$, $m \in L(K)$. Then $(rm+N) = r(m+N) \in M/N$, hence $rm \in L(K)$. Then $u' : U' \to U, K \mapsto L(K)$ is well-defined. One easily verifies that u and u' are mutual inverses. □

Lemma 2.4.20. *Let R be a ring, M a left/right R-module, A a set and $\{M_\alpha\}_{\alpha \in A}$ a family of left/right R-submodules of M. Then $s : \bigoplus_{\alpha \in A} M_\alpha \to M$ (cf. Proposition 2.2.25), hence*

$$\sum_{\alpha \in A} M_\alpha$$

is a left/right R-submodule.

Proof. Indeed for $r \in R$, $(m_\alpha) \in \bigoplus_{\alpha \in A} M_\alpha$,

$$s(r(m_\alpha)) = s((rm_\alpha)) = \sum_{\alpha \in A} rm_\alpha = \sum_1^n rm_{\alpha_i} = r\sum_1^n m_{\alpha_i} = r\sum_{\alpha \in A} m_\alpha = rs((m_\alpha)),$$

where $\alpha_1, \ldots, \alpha_n \in A$ are chosen suitably. □

Remark 2.4.21. We define $\sum_1^n M_1 = \sum_{i \in \{1,\ldots,n\}} M_i$ for left/right R-submodules $M_1, \ldots, M_n$ of M and $M_1 + M_2 := \sum_1^2 M_i$.

Lemma 2.4.22. *Let R be a ring, $I, J \subset R$ be a left resp. right ideal, M a left/right R-module and $m \in M$. Then*

$$Im := \{rm : r \in I\} \;\&\; mJ := \{mr : r \in J\}$$

is a left resp. right R-submodule of M. Let $X \subset M$. Then

$$IX := \sum_{x \in X} Rx \;\&\; XJ := \sum_{x \in X} xR$$

is a left resp. right R-submodule of M.

Proof. Indeed for the first statement let $a,b \in I$ and $r \in R$. Then $ra \in I$, hence

$$r(am) = (ra)m \in Im.$$

Furthermore, since $a + b \in I$, hence

$$am + bm = (a + b)m \in Im.$$

The right case follows from J being a left $R^{(\mathrm{op})}$-module hence mJ is a left $R^{(\mathrm{op})}$-module, hence mJ is a right R-module. IX, XJ being left/right modules follows from the first statement and Lemma 2.4.20. □

Definition 2.4.23. Let R be a ring and M a left/right R-module. Then M is said to be *finitely generated over* R if there is a finite sequence $m_1, \ldots, m_n \in M$ such that $M = \sum_1^n Rm_i$.

Definition 2.4.24. Let A be a set, R a ring and $\{M_\alpha\}_{\alpha \in A}$ a family of left/right R-modules. We say that $\sum_{\alpha \in A} M_\alpha$ is *direct*, if for every $\beta \in A$,

$$M_\beta \cap \sum_{\alpha \in A \setminus \{\beta\}} M_\alpha = 0.$$

Lemma 2.4.25. *Let A be a set, R a ring and $\{M_\alpha\}_{\alpha \in A}$ a family of left/right R-modules such that $\sum_{\alpha \in A} M_\alpha$ is direct. Then*

$$\sum_{\alpha \in A} M_\alpha \simeq \bigoplus_{\alpha \in A} M_\alpha.$$

Proof. We define the map

$$\rho : \bigoplus_{\alpha \in A} M_\alpha \to \sum_{\alpha \in A} M_\alpha$$
$$(m_\alpha) \mapsto \sum_{\alpha \in A} m_\alpha,$$

where $\sum_{\alpha \in A} m_\alpha$ is defined to be the sum of non-zero entries of (m_α). Let $\sum_1^n m_{\alpha_i}$, where $n \geq 1$ and $\alpha_1, \ldots, \alpha_n \in A$. One easily finds that this is a module homomorphism. For $\alpha \in A$, we then define $m_\alpha = m_{\alpha_i}$ if $\alpha = \alpha_i$ for some i and $m_\alpha = 0$ if not. Then Clearly

$$\sum_1^n m_{\alpha_i} = \sum_{\alpha \in A} m_\alpha = \rho((m_\alpha)),$$

which means ρ is surjective. Suppose $(m_\alpha) \in \ker\ \rho$. Then

$$0 = \rho((m_\alpha)) = \sum_{\alpha \in A} m_\alpha = \sum_1^n m_{\alpha_1},$$

for some distinct $\alpha_1,\ldots,\alpha_n \in A$. Let $j \in \{1,\ldots,n\}$. Then

$$-m_{\alpha_j} = \sum_{i\in\{1,\ldots,n\}\setminus\{j\}} m_{\alpha_i} \in \sum_{\alpha\in A\setminus\{\alpha_j\}} M_\alpha.$$

This implies $m_{\alpha_j} \in M_{\alpha_j} \cap \sum_{\alpha\in A\setminus\{\alpha_j\}} M_\alpha = 0$, hence $m_{\alpha_j} = 0$, which means $m_\alpha = 0$ for each $\alpha \in A$ and so $(m_\alpha) = 0$. By the 1st Isomorphism Theorem for modules it follows that $\sum_{\alpha\in A} M_\alpha \simeq \bigoplus_{\alpha\in A} M_\alpha$. □

Definition 2.4.26. Let R be a ring and M a left/right R-module. A subset $X \subset M$ is said to be *left/right linearly independent over* R (or if R is commutative just *linearly independent*), if for every finite sequence $m_1,\ldots,m_n \in M$ and every finite sequence $r_1,\ldots,r_n \in R$,

$$\sum_1^n r_i m_i = 0 \iff r_i = 0 \;\forall i \in \{1,\ldots,m\} \text{ resp. } \sum_1^n m_i r_i = 0 \iff r_i = 0 \;\forall i \in \{1,\ldots,m\}$$

Remark 2.4.27. One should note that $0 \notin X$, since $1\cdot 0 = 0$ and $1 \neq 0$.

Proposition 2.4.28. *Let R be a ring, M a left/right R-module and $X \subset M$ a subset. Then if X is left/right linearly independent over R, $\sum_{x\in X} Rx$ is direct.*

Proof. When X is empty the statement is trivial, hence suppose $X \neq \emptyset$.
Let $y \in X$ and let $m \in Ry \cap \sum_{x\in X\setminus\{y\}} Rx$. Then

$$r_{n+1}y = m = \sum_1^n r_i x_i,$$

for suitable $x_1,\ldots,x_n \in X\setminus\{y\}$ and $r_1,\ldots,r_{n+1} \in R$. Thus, we have that

$$r_{n+1}y + \sum_1^n r_i x_i = 0 \Rightarrow 0 = r_1 = r_2 = \cdots = r_{n+1} \Rightarrow m = 0.$$

We then conclude that $Ry \cap \sum_{x\in X\setminus\{y\}} Rx = 0$. □

Definition 2.4.29. Let R be a ring, M a left/right R-module. A subset $X \subset M$ is called a *basis of M over R* if X is linearly independent over R and $M = RX$ respectively $M = XR$. If X is finite and a basis of M over R it is called a *finite basis*.

Proposition 2.4.30. *Let S be a ring and $R \subset S$ a subring. Consider $(M,\cdot,+)$, a left/right S-module, then $rm := r\cdot m$, for $r \in R$, defines a structure of left/right R-modules. If in addition Q is a subring of a ring T and M is an (S,T)-bimodule, then M is an (R,Q)-bimodule.*

Proof. Let $m_1, m_2 \in M$, $r_1, r_2 \in R$. Then

1. $r_1(m_1+m_2) = r_1 \cdot (m_1+m_2) = r_1 \cdot m_1 + r_1 \cdot m_2 = r_1 m_1 + r_1 m_2$,
2. $(r_1+r_2)m_1 = (r_1+r_2) \cdot m_1 = r_1 \cdot m_1 + r_2 \cdot m_1 = r_1 m_1 + r_2 m_1$,
3. $(r_1 r_2)m_1 = (r_1 r_2) \cdot m_1 = r_1(r_2 \cdot m_1) = r_1(r_2 m_1)$,
4. $1m_1 = 1 \cdot m_1 = m_1$.

Let $r \in R$, $q \in Q$, $m \in M$. Then

$$r(mq) = r \cdot (m \cdot q) = (r \cdot m) \cdot q = (rm)q.$$

□

2.4.2 Ideals

Definition 2.4.31. Recall that a *left/right ideal* in a ring R, is a left/right R-submodule of R. If it is an (R,R)-module it is called a *both-sided ideal.* If R is a commutative a left/right ideal is simply refered to as an ideal.

Definition 2.4.32. Let $\sigma : R \to S$ be a ring homomorphism. When we refer to the kernel of σ, we refer to the kernel of σ when seen as a group homomorphism between $(R,+)$ and $(S,+)$, i.e. $\ker \sigma := \sigma^{-1}(0)$.

Lemma 2.4.33. *Let $\sigma : R \to S$ be a ring homomorphism and $I \subset S$ be a left/right/both-sided ideal. Then $\sigma^{-1}(I) \subset R$ is a left/right/both-sided ideal.*

Proof. By Lemma 2.2.16 it follows that $\sigma^{-1}(I) \subset R$ is an additive subgroup. Let $r \in R$ and $a \in \sigma^{-1}(I)$. Then

$$\sigma(ra) = \sigma(r)\sigma(a) \in I,$$

hence $ra \in \sigma^{-1}(I)$. □

Corollary 2.4.34. *The kernel of a ring homomorphism $\sigma : R \to S$ is an ideal in R*

Proof. This follows immediately from the above lemma. □

Lemma 2.4.35. *Let $\sigma : R \to S$ be a surjective ring homomorphism and $I \subset R$ be a left/right/both-sided ideal. Then $\sigma(I) \subset S$ is a left/right/both-sided ideal.*

Proof. By Lemma 2.2.16 $\sigma(I)$ is an additive subgroup of S. Let $s \in S$ and $\sigma(a) \in \sigma(I)$. Then for some $r \in R$, $s = \sigma(r)$. It follows that

$$s\sigma(a) = \sigma(r)\sigma(a) = \sigma(ra) \in \sigma(I).$$

□

Remark 2.4.36. We call RX and XR *the left/right ideal generated by* X. *The ideal generated by* X is the ideal $\langle X\rangle := R(XR) = (RX)R$.
Suppose R is commutative. For $M \subset R$, one can easily check that $RM = MR$, and hence the left/right ideal generated by M over R is a two-sided ideal. Thus $\langle M\rangle = RM = MR$.

Example 2.4.37. One may note that quite clearly $R = R1 = 1R$ and hence that $R = \langle 1\rangle$.

Lemma 2.4.38. *Let* R *be a ring and* $I \subset R$ *a left/right/both-sided ideal. Then*

$$1 \in I \iff I = R.$$

Proof. We need only work with the assumption that I is a left ideal.
"$\Rightarrow$": Let $r \in R$. Then $r = r1 \in I$, hence $R \subset I \Rightarrow R = I$.
"$\Leftarrow$": This is trivial, since $1 \in R = I$. ☐

Definition 2.4.39. An ideal in a ring generated by only one element is called a *principal ideal.* A ring in which every ideal is principal is called a *principal ideal domain* or a *PID*

An example of a PID is $\mathbb{Z}$. Indeed $(\mathbb{Z},+)$ is a cyclic group generated by 1, thus every subgroup of the form $\langle n\rangle = n\mathbb{Z}$ for some $n \geq 0$, in particular every ideal is of this form.

Proposition 2.4.40. *Let* S *be a ring and* $R \subset S$ *a subring. Consider a left/right ideal* $I \subset S$. *Then* $I \cap R \subset R$ *is an ideal.*

Proof. R and I are both left/right R-submodules of S, hence so is $R \cap I$. ☐

Lemma 2.4.41. *Let* R *be a commutative ring and* $I, J \subset R$ *ideals. Then*

$$IJ \subset I \cap J.$$

Proof. Let $ij \in IJ$. Then since $j \in J$, $ij \in J$ and since $i \in I$, $ij \in I$, hence $ij \in I \cap J$. ☐

2.4.3 Quotient Rings

Proposition 2.4.42. *Let* R *be a ring and* $I \subset R$ *an ideal. View* I *as a subgroup of the additive group* $(R,+)$. *Then* $(R/I,+)$ *is an additive group by Corollary 2.2.40. Define*

$$\cdot : R/I \times R/I \to R/I$$

by $(r+I)(r'+I) := (rr'+I)$. *This is a well-defined operation and* $(R/I,+,\cdot)$ *is a ring. It is also commutative if* R *is commutative.*

Proof. Let $(r_1+I, r_2+I) = (r'_1+I, r'_2+I) \in R/I \times R/I$. Then

$$r_1r_2 - r'_1r'_2 = r_1r_2 - r_1r'_2 + r_1r'_2 - r'_1r'_2 = r_1(r_2 - r'_2) + (r_1 - r'_1)r'_2 \in I,$$

hence $r_1r_2 + I = r'_1r'_2 + I$ and it follows that $\cdot$ is well-defined. Let $r_1+I, r_2+I, r_3+I \in R/I$. Then

$$\begin{aligned}(r_1+I)((r_2+I)(r_3+I)) &= (r_1+I)(r_2r_3+I) = (r_1(r_2r_3)) + I = (r_1r_2)r_3 + I \\ &= (r_1r_2+I)(r_3+I) = ((r_1+I)(r_2+I))(r_3+I).\end{aligned}$$

We also have that

$$(1+I)(r+I) = (1r+I) = (r1+I) = r+I.$$

Furthermore

$$\begin{aligned}(r_1+I)((r_2+I)+(r_3+I)) &= (r_1+I)((r_2+r_3)+I) = (r_1(r_2+r_3)) + I = (r_1r_2+r_1r_3)+I \\ &= (r_1r_2+I)+(r_1r_3+I) = (r_1+I)(r_2+I)+(r_1+I)(r_3+I).\end{aligned}$$

Suppose R is commutative. Then

$$(r_1+I)(r_2+I) = (r_1r_2)+I = (r_2r_1)+I = (r_2+I)(r_1+I).$$

□

Corollary 2.4.43. *The canonical surjective group map $\pi : R \mapsto R/I$ is a ring map.*

Proof. Let $r_1, r_2 \in R$. Then $\pi(r_1r_2) = r_1r_2 + I = (r_1+I)(r_2+I) = \pi(r_1)\pi(r_2)$. Lastly $\pi(1) = 1 + I$. □

2.4.4 Noetherian Modules and Noetherian Rings

Definition 2.4.44. Let X be a set. Let $\leq \subset X \times X$, where we write $x \leq y$ if $(x,y) \in \leq$. $\leq$ is a *partial order*, if it is 1. reflexive, 2. antisymmetric and 3. transitive, i.e. for $x, y, z \in X$

1. $x \leq x$,
2. $x \leq y \;\&\; y \leq x \Rightarrow x = y$,
3. $x \leq y \;\&\; y \leq z \Rightarrow x \leq z$.

We write $x \geq y$ if $y \leq x$.

Remark 2.4.45. Given a partial order $\leq$, we have that $\geq$ is a partial order as well.

Example 2.4.46. Let X be a set and $\mathscr{X} \subset 2^X$. Then $\subset := \{(A,B) \in \mathscr{X} \times \mathscr{X} : \forall x (x \in A \Rightarrow x \in B)\}$ defines a partial order on $\mathscr{X}$.
Another example is that of $\leq$ on $\mathbb{N}$,$\mathbb{Z}$ or $\mathbb{Q}$.

Definition 2.4.47. Let X be a set with a partial order and $\{x_i\}_{i\in\mathbb{N}} \subset X$ be a sequence. We say that $\{x_i\}_{i\in\mathbb{N}}$ is *descending with respect to* $\leq$ if $x_i \geq x_{i+1}$ for every $i \geq 0$ and *ascending with respect to* $\leq$ if $x_i \leq x_{i+1}$ for every $i \geq 0$. A sequence $\{x_i\}_{i\in\mathbb{N}}$ is said to *stabilize* if there is a non-negative integer n such that $x_n = x_{n+d}$ for every $d \geq 0$.

Definition 2.4.48. Let X be a set with a partial order $\leq$. A subset Y of X is called a *chain* if for every $c,d \in Y$, $c \leq d$ or $d \leq c$.

Remark 2.4.49. Any ascending/descending sequence $\{x_i\}_{i\in\mathbb{N}}$ is a chain and is denoted

$$x_1 \leq x_2 \leq \ldots \text{ respectively } x_1 \geq x_2 \geq \ldots,$$

these are called *ascending/descending chains*

Definition 2.4.50. Let M be a left/right R-module and

$$\mathscr{M} := \left\{N \in 2^M : N \text{ is a left/right submodule of } M\right\}$$

be a chain. We say that M is *left/right noetherian* if every ascending chain stabilizes and *left/right artinian* if every descending chain stabilizes. A ring S is left/right noetherian/artinian, if it is noetherian/artinian as a left/right S-module or simply artinian/noetherian if is both left and right artinian/noetherian.

Definition 2.4.51. A simple left/right R-module is one whose only submodules are 0 and M.

Example 2.4.52. Any simple left/right R-module M is noetherian/artinian. A family of submodules of M is of the form $\{0,M\},\{M\}$ or $\{0\}$. These are all finite sets, hence any chain C is finite and thus has an upper/lower bound in C.
Simple rings are simple modules. This means division rings and fields are both noetherian/artinian.

Lemma 2.4.53. *Let M be a left/right R-module. Consider a chain C of submodules of M. Then*

$$\bigcup_{N\in C} N$$

is a submodule.

Proof. Let $m_1, m_2 \in \bigcup_{N \in C} N$. $m_1 \in N_1$ and $m_2 \in N_2$ for some $N_1, N_2 \in C$. WLOG $N_1 \subset N_2$, hence $m_1 \in N_2$, which means $m_1 + m_2 \in N_2 \subset \bigcup_{N \in C} N$. Clearly $0 \in \bigcup_{N \in C} N$. Let $r \in R$, clearly $rm_1 \in \bigcup_{N \in C} N$. □

We give the following axiom which one check is equivalent to the axiom of choice

Axiom. *(Zorn's Lemma) Let $X \neq \emptyset$ be a set with a partial order $\leq$ such that for every chain $C \subset X$ there exists an $x \in C$ such that $c \leq x$ for every $c \in C$, (i.e. there is an upper bound x for C in C). Then there is a maximal element in $m \in X$, i.e. for every $y \in X$ if $m \leq y$, then $m = y$.*

Example 2.4.54. In certain situations we do not need to assume Zorn's Lemma.

1. Suppose X is a non-empty finite set with n elements and a partial order $\leq$. Then X has a maximal element. Indeed, this is easily proven by induction in n. If X has one element this is trivially maximal. Consider for $n \geq 1$ $X = \{x_1, \ldots, x_{n+1}\}$. Then by induction $\{x_1, \ldots, x_n\}$ has a maximal element x_i. Then $\max_{\leq}(x_i, x_{n+1})$ is a maximal element of X.

2. A topology τ on some set X has X as a maximal element

We give a reformulation of every chain having a maximal/minimal element

Lemma 2.4.55. *Let $X \neq \emptyset$ be a set with a partial ordering $\leq$. Every ascending/descending sequence in X stabilizes if and only if every chain C in X has a upper/lower bound in C.*

Proof. We only check the ascending case since a descending sequence is just an ascending sequence with respect to $>$ and a minimal element is just a maximal element with respect to $>$.

"$\Rightarrow$":We prove the contrapositive. Suppose $C \subset X$ is a chain that does not have an upper bound in C. Let $c_1 \in C$. Then there exists $c_2 \in C$ such that $c_1 < c_2$. Continuing this process recursively we get a sequence $\{c_i\}_{i \in \mathbb{N}}$ such that $c_i < c_{i+1}$ for every $i \geq 0$, hence this is a sequence in X that does not stabilize.

"$\Leftarrow$": Let $\{x_i\}_{i \in \mathbb{N}}$ be an ascending sequence in X. Then $\{x_i\}_{i \in \mathbb{N}}$ is a chain. Then there exists a $x_n \in \{x_i\}_{i \in \mathbb{N}}$ such that $x_j \leq x_n$ for every $j \geq 1$. Now since $x_n \leq x_{n+d}$ for every $d \geq 0$ it follows that $x_n = x_{n+d}$ for every $d \geq 0$, hence $\{x_i\}_{i \in \mathbb{N}}$ stabilizes. □

The consequence of the above lemma is the following

Proposition 2.4.56. *If M is a left/right R-module and X is a non-empty set of submodules of M, and M is noetherian/artinian then X has maximal/minimal element*

Proof. If M is noetherian/artinian, then every chain C in X has an upper/lower bound in C. Using Zorn's Lemma, this implies that X has a maximal/minimal element. ☐

Corollary 2.4.57. *Every left/right noetherian ring R has a maximal left/right ideal. In particular every noetherian ring has a maximal ideal.*

Proof. Let

$$X = \{I \subsetneq R : I \text{ a left ideal in } R\}.$$

Then it follows by the above proposition that this set has a maximal element, which by definition will be a maximal left/right ideal of R. Defining

$$Y = \{I \subsetneq R : I \text{ an ideal in } R\},$$

the above proposition implies the existence of a maximal ideal. ☐

Theorem 2.4.58. *Let M be a left/right R module. Then M is left/right noetherian if ad only if every submodule of M is finitely generated over R.*

Proof. "$\Rightarrow$": Suppose M is not finitely generated. Let $m_1 \in M$. Then $M_1 := Rm_1 \subsetneq M$ since M is not finitely generated. We then recursively define $M_{n+1} = Rm_{n+1} + M_n$ where $m_{n+1} \in M \setminus M_n$ which we can do since every M_n is finitely generated and thus by assumption a proper submodule of M. We thus obtain an strictly ascending chain of submodules

$$M_1 \subsetneq M_2 \subsetneq M_3 \subsetneq \dots$$

Hence this is an ascending chain that does not stabilize.
"$\Leftarrow$": Suppose every submodule of M is finitely generated over R. Let an ascending chain of submodules, say

$$M_1 \subset M_2 \subset \dots$$

be given. Then by Lemma 2.4.53 $N := \bigcup_1^\infty M_n$ is a submodule. By assumption $N = \sum_1^m Rn_k$ for some $n_1, \dots, n_m \in N$. For each $k \in \{1, \dots, m\}$ there is a $j(k) \geq 0$ such that $n_k \in M_{j(k)}$. Let $p = \max\{j(1), \dots, j(k)\}$. Thus we have that $n_1, \dots, n_m \in M_p$. Hence, $M_p = \sum_1^m Rn_k = N$. Hence for $d \geq 0$, $N = M_p \subset M_{p+d}$, implying $M_p = M_{p+d}$. In other words, every ascending chain of submodules of M stabilizes. ☐

Corollary 2.4.59. *A PID is a Noetherian ring.*

Lemma 2.4.60. *Let a left/right R-module M and a submodule $N \subset M$ be given. Then M is left/right noetherian/artinian if and only if N and M/N is left/right noetherian/artinian.*

Proof. We prove the left noetherian version.
"$\Rightarrow$": An ascending chain of submodules of N, is in particular an ascending chain of submodules of M, which by assumption stabilizes. A chain of submodules in M/N is of the form

$$L_1/N \subset L_2/N \subset L_3/N \subset \dots$$

where $L_i \subset M$ is a submodule of M containing N by Lemma 2.4.19. For some $n \geq 0$, $L_n = L_{n+d}$ for every $d \geq 0$, hence $L_n/N = L_{n+d}/N$ for every $d \geq 0$.
"$\Leftarrow$": Let $M_1 \subset M_2 \subset M_3 \subset \dots$. Then

$$M_1 + N \subset M_2 + N \subset \dots$$

is an ascending chain of submodules of M containing N. Hence

$$(M_1 + N)/N \subset (M_2 + N)/N \subset \dots$$

is an ascending chain of submodules of M/N. Hence for some $n \geq 1$, $M_n + N = M_{n+d} + N$ for every $d \geq 0$. Then by Lemma 2.4.19 $M_n + N = M_{n+d} + N$ for every $d \geq 0$. We also have that

$$M_1 \cap N \subset M_2 \cap N \subset \dots$$

is an ascending chain of submodules of N. Thus for some $m \geq 1$, $M_m \cap N = M_{m+d} \cap N$ for every $d \geq 0$. Put $k = n + m$ and let $d \geq 0$ be given. Let $x \in M_{k+d}$. Then $x \in M_{k+d} + N = M_k + N$. Hence $x = y + z$ for some $y \in M_k$ and $z \in N$. It thus follows that $z = x - y \in N \cap M_{k+d} = N \cap M_k$. In particular, $z \in M_k$, hence $x = z + y \in M_k$, hence $M_k = M_{k+d}$. Thus $M_1 \subset M_2 \subset \dots$ stabilizes
We proceed to prove the left artinian case. "$\Rightarrow$" is dual to the noetherian case.
"$\Leftarrow$": Consider a descending chain of submodules of M,

$$M_1 \supset M_2 \supset \dots$$

Similarly as above $M_1 + N \supset M_2 + N \supset \dots$ and $M_1 \cap N \supset M_2 \cap N \supset \dots$ give descending chain stabilizing some positive n respectively m. Put $k = n + m$ and let $d \geq 0$. Consider $x \in M_k$. Then in particular $x \in M_k + N = M_{k+d} + N$. Thus $x = y + z$ for some $y \in M_{k+d}$ and $z \in N$. Then $z = x - y \in M_k \cap N = M_{k+d} \cap N$, hence $x = z + y \in M_{k+d}$, hence $M_k = M_{k+d}$, meaning $M_1 \subset M_2 \subset \dots$ stabilizes. $\square$

2.4.5 A First Look at Algebras over Rings

Definition 2.4.61. By a *ring extension* S *over* R we mean a ring S containing a ring R as a subring. Such a pair is denoted $S \supset R$. Such a pair is a *field extension* if S is a field and R is a subfield.

Remark 2.4.62. This defines a partial order. Indeed any ring is a ring extension of itself. If $S \supset R$ and $R \supset S$ then $R = S$. Lastly, if $T \supset S$ and $S \supset R$ are ring extensions, then $T \supset R$ and R is a subring of T.

Definition 2.4.63. Let R be a ring. We define *the center of R* to be the set

$$Z(R) := \{x \in R : xy = yx \textit{ for every } y \in R\}$$

Remark 2.4.64. The center R is a subring of R. Indeed, let $x, x' \in Z(R)$. Then given $y \in R$,

$$y(x + x') = yx + yx' = xy + x'y = (x + x')y \Rightarrow\ x + x' \in Z(R).$$

We also have that $y0 = 0 = 0y$ for every $y \in R$, hence $0 \in Z(R)$. In addition,

$$y(-x) = -(yx) = -xy \Rightarrow\ -x \in Z(R).$$

This means $Z(R)$ is a subgroup of R. Furthermore,

$$y(xx') = (yx)x' = (xy)x' = x(yx') = x(x'y) = (xx')y \Rightarrow\ xx' \in Z(R)$$

and $y1 = y = 1y$, hence $1 \in Z(R)$. We thus see that $Z(R)$ is the largest commutative subring of R

Definition 2.4.65. Let R be a commutative ring. A ring A is an R-algebra, if $(A, +)$ is an R-module such that

$$r(a_1a_2) = (ra_1)a_2$$

for every $r \in R$, $a_1, a_2 \in A$.

Remark 2.4.66. An equivalent definition is that an algebra is a ring A together with ring homomorphism $R \to A$ whose image is contained in $Z(A)$.

Definition 2.4.67. Let A be an R-algebra. A subset B of A is *an R-subalgebra of A* if it is a subring of A and an R-submodule of A.

Remark 2.4.68. One easily checks that indeed an R-subalgebra is itself an R-algebra.

Definition 2.4.69. Let S be an algebra over commutative ring R and $s_1, \ldots, s_n \in Z(S)$. We define the *algebra over R generated by $s_1, \ldots, s_n$* to be the set

$$R[s_1, \ldots, s_n] := \left\{ \sum_{v=(v_1,\ldots,v_n)\in\mathbb{N}^n} a_v s_1^{v_1} \cdots s_n^{v_n} : a_v \in R \forall v \in \mathbb{N}^n, a_v = 0 \text{ for all but finitely many } v \in \mathbb{N}^n \right\}.$$

If $S \supset R$, then the above set is the *ring extension over R generated by $s_1 \ldots, s_n$*

Lemma 2.4.70. *Let S be an algebra over a commutative ring R and $s_1,\dots,s_n \in Z(S)$. $R[s_1,\dots,s_n]$ is an R-subalgebra of S. Furthermore, $R[s_1,\dots,s_n]$ is the smallest subalgebra over R containing $s_1,\dots,s_n$.*

Proof. Clearly $1,0 \in R[s_1,\dots,s_n]$. Let $\sum_{v\in\mathbb{N}^n} a_v s_1^{v_1}\cdots s_n^{v_n}, \sum_{v\in\mathbb{N}^n} b_v s_1^{v_1}\cdots s_n^{v_n} \in R[s_1,\dots,s_n]$. Then

$$\begin{aligned}\sum_{v\in\mathbb{N}^n} a_v s_1^{v_1}\cdots s_n^{v_n} + \sum_{v\in\mathbb{N}^n} b_v s_1^{v_1}\cdots s_n^{v_n} &= \sum_{v\in\mathbb{N}^n}\left(a_v s_1^{v_1}\cdots s_n^{v_n} + b_v s_1^{v_1}\cdots s_n^{v_n}\right)\\ &= \sum_{v\in\mathbb{N}^n}(a_v+b_v)s_1^{v_1}\cdots s_n^{v_n} \in R[s_1,\dots,s_n].\end{aligned}$$

and

$$\begin{aligned}\left(\sum_{v\in\mathbb{N}^n} a_v s_1^{v_1}\cdots s_n^{v_n}\right)\left(\sum_{w\in\mathbb{N}^n} b_w s_1^{w_1}\cdots s_n^{w_n}\right) &= \sum_{v\in\mathbb{N}^n}\sum_{w\in\mathbb{N}^n} a_v b_w s_1^{v_1+w_1}\cdots s_n^{v_n+w_n}\\ &= \sum_{u\in\mathbb{N}^n}\left(\sum_{v,w\in\mathbb{N}^n : v+w=u} a_v b_w\right) s_1^{u_1}\cdots s_n^{u_n} \in R[s_1,\dots,s_n]\end{aligned}$$

Let $r \in R$. Then

$$r\sum_{v\in\mathbb{N}^n} a_v s_1^{v_1}\cdots s_n^{v_n} = \sum_{v\in\mathbb{N}^n}(ra_v)s_1^{v_1}\cdots s_n^{v_n} \in R[s_1,\dots,s_n].$$

□

Lemma 2.4.71. *Let $S \supset R$ be a ring extension of commutative rings and $s_1,\dots,s_n,t_1,\dots,t_m \in S$. Then $R[s_1,\dots,s_n,t_1,\dots,t_m] = R[s_1,\dots,s_n][t_1,\dots,t_m]$.*

Proof. We already have that $R[s_1,\dots,s_n,t_1,\dots,t_m] \supset R[s_1,\dots,s_n]$, hence $R[s_1,\dots,s_n,t_1,\dots,t_m] \supset R[s_1,\dots,s_n][t_1,\dots,t_m]$. Let $\sum_{(v,w)\in\mathbb{N}^{n+m}} a_v s_1^{v_1}\cdots s_n^{v_n} t_1^{w_1}\cdots t_m^{w_m} \in R[s_1,\dots,s_n,t_1,\dots,t_m]$. We then see that

$$\sum_{(v,w)\in\mathbb{N}^{n+m}} a_{vw} s_1^{v_1}\cdots s_n^{v_n} t_1^{w_1}\cdots t_m^{w_m} = \sum_{w\in\mathbb{N}^m}\left[\sum_{v\in\mathbb{N}^n} a_{vw} s_1^{v_1}\cdots s_n^{v_n}\right] t_1^{w_1}\cdots t_m^{w_m} \in R[s_1,\dots,s_n][t_1,\dots,t_m].$$

□

Definition 2.4.72. Let R be a ring and consider R-algebras A,B. A map $\sigma : A \to B$ is an *R-algebra homomorphism* if it is both a ring homomorphism and an R-module homomorphism.

Lemma 2.4.73. *Let R be a ring, consider R-algebras A,B and a map $\sigma : A \to B$. Then σ is an R-algebra homomorphism if and only if σ is a multiplicative map fixing R, i.e. for every $a_1,a_2 \in A$, $\sigma(a_1a_2) = \sigma(a_1)\sigma(a_2)$ and for every $r \in R$, $\sigma(r) = r$.*

Proof. DO AT SOME POINT. □

2.5 Abelian Categories

2.5.1 Preadditive Categories

Definition 2.5.1. A category $\mathcal{C}$ is *preadditive* if every $\mathbf{Hom}(A,B)$ is an additive group and

$$f(g+h) = fg + fh, \quad (f+g)h = fh + gh.$$

Example 2.5.2. 1.

2.5.2 Initial Objects, Terminal Objects & Zero Objects

Definition 2.5.3. An *initial object* I in a category $\mathcal{C}$ is an object in C such that for every object X, $\mathbf{Hom}(I,X)$ is singleton. A *terminal object* in $\mathcal{C}$ is a an initial object in $\mathcal{C}^{\mathrm{op}}$. An element in $\mathcal{C}$ is a *zero object* if it is both initial and terminal.

2.5.3 Additive, Pre-abelian & Abelian Categories

Definition 2.5.4. A preadditive category $\mathcal{C}$ with zero objects is called *additive* if it admits binary coproducts.by some theorem

Definition 2.5.5. In an additive category $\mathcal{C}$ of a pair of objects A,B in $\mathcal{C}$ a *biproduct* of A and B is an object $A \oplus B$ such that

$$\begin{tikzcd} A \arrow[r, shift left, "p_1"] & A \oplus B \arrow[l, shift left, "i_1"] \arrow[r, shift left, "p_2"] & B \arrow[l, shift left, "i_2"] \end{tikzcd}$$

$i_1 p_1 = \mathbb{1}_A$, $p_2 i_2 = \mathbb{1}_B$ and $p_1 i_1 + i_2 p_2 = \mathbb{1}_{A \oplus B}$

Lemma 2.5.6. *Any additive category admits a biproduct. In particular it admits a product and finite coproducts and products are isomorphic.*

Proof. □

Definition 2.5.7. In an additive category $\mathcal{C}$ a *kernel* of a morphism $f : A \to B$ is an object K and a morphism $k : K \to A$ such that

$$\begin{tikzcd} K \arrow[r, "0"] \arrow[d, "k"] & B \\ A \arrow[ur, "f"] & \end{tikzcd}$$

commutes. A *cokernel* in $\mathcal{C}$ is a kernel in $\mathcal{C}^{\mathrm{op}}$

2.6 Homological Algebra

2.6.1 Exact Sequences

Definition 2.6.1. *A sequence of left/right R-modules,* is a collection of pairs

$$\{(M_i, \rho_i) : i \in \mathbb{Z}, M_i \text{ is a left/right } R\text{-module }, \rho_i \in \mathrm{Hom}(M_i, M_{i+1})\}.$$

A sequence is finite if $M_i = 0$ for every but finitely many i. A sequence is in general denoted

$$\cdots \xrightarrow{\rho_{i-1}} M_i \xrightarrow{\rho_i} M_{i+1} \xrightarrow{\rho_{i+1}} \cdots$$

When the maps are obvious we opt to not explicitly denote them. When a sequence is finite we of often opt to denote it

$$0 \longrightarrow M_s \longrightarrow M_{s+1} \longrightarrow \cdots \longrightarrow M_{b-1} \longrightarrow M_b \longrightarrow 0$$

Where s, b are respectively the largest and smallest index for which $M_s \neq 0$ and $M_b \neq 0$.

Definition 2.6.2. A finite sequence

$$M \xrightarrow{\rho} M' \xrightarrow{\rho'} M''$$

is said to be *exact (at M')* if $\mathbf{im}\ \rho = \ker \rho'$. In general a sequence

$$\cdots \xrightarrow{\rho_{i-1}} M_i \xrightarrow{\rho_i} M_{i+1} \xrightarrow{\rho_{i+1}} \cdots$$

is said to be *exact* if

$$M_{i-1} \xrightarrow{\rho_i} M_i \xrightarrow{\rho_i} M_{i+1}$$

is exact for each $i \in \mathbb{Z}$

Remark 2.6.3. Equivalently a sequence $M \xrightarrow{\rho} M' \xrightarrow{\rho'} M''$ is exact if $\rho' \circ \rho = 0$.

Lemma 2.6.4. *Consider a sequence*

$$0 \longrightarrow M \xrightarrow{\rho} M' \xrightarrow{\rho'} M'' \longrightarrow 0$$

The following are equivalent:

1. *The sequence is exact*

2. *ρ is injective and ρ' is surjective*

Proof. "1.$\Rightarrow$ 2.": by exactness $\mathbf{ker}\,\rho = \mathbf{im}\ 0 = 0$ and $\mathbf{im}\ \rho' = \mathbf{ker}\ 0 = M''$. "2.$\Rightarrow$ 1.": $\mathbf{im}\ 0 = 0 = \mathbf{ker}\ \rho$ and $\mathbf{ker}\ 0 = M'' = \mathbf{im}\ \rho'$ ☐

Corollary 2.6.5. *Given a left/right R-module map $\rho : M \to N$, the sequence*

$$0 \longrightarrow \mathbf{ker}\ \rho \hookrightarrow M \twoheadrightarrow \mathbf{im}\ \rho \longrightarrow 0$$

is exact.

2.6.2 Isomorphism Theorems

We are going to construct ways of identifying certain algebraic structures given certain maps. We will develop these theorems for groups, rings, modules and algebra homomorphisms in a sense separately. However, in a certain categorical setting which I don't know we would be able to develop them all at once.

Definition 2.6.6. An *isomorphism of groups* is a bijective group homomorphism. A *isomorphism of rings* is a bijective ring homomorphism. If there exists an isomorphism between groups/rings $G,H/R,S$ then $G,H/R,S$ are said to be *isomorphic as groups/rings* and we write $G \simeq H/R \simeq S$, when this does not lead to confusion.

Remark 2.6.7. One easily check that the inverse of of a bijective group/ring homomorphism is automatically a group/ring homomorphism itself. Indeed, if $\rho : G \to H$ is a bijective monoid map, let $h,h' \in H$, then for some $g,g' \in G$, $h = \rho(g)$ and $h' = \rho(g')$. Then

$$\rho^{-1}(hh') = \rho^{-1}(\rho(g)\rho(g')) = \rho^{-1}(\rho(gg')) = gg' = \rho^{-1}(h)\rho^{-1}(h').$$

Lastly

$$\rho^{-1}(e_H) = \rho^{-1}(\rho(e_G)) = e_G.$$

Example 2.6.8. Let R be a commutative ring. Consider the identity map

$$\mathrm{id} : (R,\cdot) \to (R,*) = R^{(\mathrm{op})}, r \mapsto r$$

Then this clearly a bijective map of groups. Let $r,r' \in R$ then

$$\mathrm{id}(rr') = rr' = r'r = \mathrm{id}(r')\mathrm{id}(r) = \mathrm{id}(r) * \mathrm{id}(r'),$$

hence $R \simeq R^{(\mathrm{op})}$.

Lemma 2.6.9. *Let $\rho : G \to H$ be a group homomorphism. Then* $\mathbf{ker}\ \rho = \{e\}$ *if and only if ρ is injective.*

Proof. "$\Rightarrow$": Let $g_1, g_2 \in G$ be given such that $\rho(g_1) = \rho(g_2)$. Then

$$\rho\left(g_1 g_2^{-1}\right) = \rho(g_1)\rho(g_2)^{-1} = e \Rightarrow g_1 g_2^{-1} = e \Rightarrow g_1 = g_2.$$

"$\Leftarrow$": Let $k \in \textbf{ker } \rho$. Then $\rho(k) = e = \rho(e)$, implying $k = e$, hence $k \in \{e\}$. $\square$

Corollary 2.6.10. *Let $\sigma : R \to S$ be a ring homomorphism. Then* $\textbf{ker } \sigma = 0$ *if and only if σ is injective.*

Lemma 2.6.11. *Let G, H be groups, $\rho : G \to H$ a group homomorphism and $I \subset G$ $J \subset H$ be normal subgroups. Then $\varrho : G/I \to H/J$, $gI \mapsto \rho(g)J$ is a well-defined group homomorphism if and only if $\rho(I) \subset J$.*

Proof. "$\Rightarrow$": Suppose ϱ is a well-defined group homomorphism. Let $\rho(i) \in \rho(I)$. Then since $iI = eI$,

$$\rho(i)J = \varrho(iI) = \varrho(eI) = \rho(e)J = eJ,$$

hence $\rho(i) \in J$.
"$\Leftarrow$": Suppose $\rho(I) \subset J$. Let $g_1 I = g_2 I \in G/I$. Then $g_1 g_2^{-1} \in I$, hence

$$\varrho(g_1 I)\varrho(g_2 I)^{-1} = \rho(g_1)\rho(g_2)^{-1}J = \rho\left(g_1 g_2^{-1}\right)J = eJ \Rightarrow \varrho(g_1 I) = \varrho(g_2 I).$$

We now check that ϱ is a group homomorphism. Let $g_1 I, g_2 I \in G/I$. Then

$$\varrho((g_1 I)(g_2 I)) = \varrho(g_1 g_2 I) = \rho(g_1 g_2) = \rho(g_1)\rho(g_2) = \varrho(g_1 I)\varrho(g_2 I).$$

$\square$

Corollary 2.6.12. *Let R, S be rings, $\sigma : R \to S$ a ring homomorphism and $I \subset R$, $J \subset S$ be left/right ideals. Then the $\vartheta : R/I \to S/J$, $r + I \mapsto \sigma(r) + J$ is a well-defined ring homomorphism if and only if $\sigma(I) \subset J$*

Proof. By the above proposition ϑ being a well-defined ring homomorphism implies $\sigma(I) \subset J$. Conversely if $\sigma(I) \subset J$ it remains to check that ϑ is a ring homomorphism. Let $r_1 + I, r_2 + I \in R/I$. Then

$$\vartheta((r_1 + I)(r_2 + I)) = \vartheta(r_1 r_2 + I) = \sigma(r_1 r_2) + J = \sigma(r_1)\sigma(r_2) + J = \vartheta(r_1 + I)\vartheta(r_2 + I).$$

Furthermore,

$$\vartheta(1 + I) = \sigma(1) + J = 1 + J.$$

$\square$

Corollary 2.6.13. *Let R be a subring of a ring S. Let $I \subset R$ be a left/right ideal. Then $\sigma : R/I \to S/IS, r+I \mapsto r+SI$ is a well defined ring map in the left case and so is $r+I \mapsto r+IS$ in the right case.*

Proof. Consider $\iota : R \hookrightarrow S, r \mapsto r$. Then since $\iota(I) \subset SI$, it follows that $\sigma : R/I \mapsto S/SI, r+I \mapsto \iota(r)+SI = r+SI$ is a well-defined ring map. □

Corollary 2.6.14. *Let R be a ring, M,N left/right R-modules, $\rho : M \to N$ a left/right module homomorphism and $L \subset M$, $K \subset N$ be submodules. Then the $\vartheta : M/L \to N/K$, $m+L \mapsto \rho(m)+K$ is a well-defined module homomorphism if and only if $\rho(L) \subset K$*

Proof. The above proposition again tells us that if ϑ is a well defined module homomorphism, then $\rho(L) \subset K$. Conversely if $\rho(L) \subset K$, we just need to check the map is homogeneous of degree 1. Indeed let $r \in R$, $m \in M$. Then

$$\vartheta(rm+L) = \sigma(rm)+K = r(\sigma(m)+K) = r\vartheta(m+L).$$

□

Proposition 2.6.15. *Let G,H be groups, $\rho : G \to H$ a group homomorphism and $N \subset G$ a normal subgroup such that $N \subset \mathbf{ker}\ \rho$. Consider the canonical surjection $\pi : G \twoheadrightarrow G/N$, i.e. $g \mapsto gN$. Then $\varrho : G/N \to H$, $gN \mapsto \rho(g)$ is the unique group homomorphism with the property that $\rho = \varrho\pi$. In other words the diagram*

$$\begin{tikzcd} G \arrow[r, two heads, "\pi"] \arrow[dr, "\rho"'] & G/N \arrow[d, dotted, "\exists!\varrho"] \\ & H \end{tikzcd}$$

commutes.

Proof. The assumption that $I \subset \mathbf{ker}\ \rho$ implies that $\rho(I) = \{e\}$ hence the above lemma shows that ϱ is a well-defined group homomorphism. Since $H/\{e\} = H$. Indeed, for $g \in G$,

$$\varrho\pi(g) = \varrho(\pi(g)) = \varrho(gI) = \rho(g) \Rightarrow \varrho\pi = \rho.$$

Uniqueness: Consider another group homomorphism $\varrho' : G/I \to H$ such that $\varrho\pi = \rho$. Let $gI \in G/I$. Then

$$\varrho'(gI) = \varrho'(\pi(g)) = \rho(g) = \varrho(\pi(g)) = \varrho(gI) \Rightarrow \varrho' = \varrho.$$

□

Corollary 2.6.16. *Let R,S be rings, $\sigma : R \to S$ and $I \subset R$ an ideal such that $I \subset \ker\ \sigma$. Define $\pi : R \to R/I$ to be the canonical surjection, i.e. $r \mapsto r+I$. Then there is a unique ring homomorphism $\varrho : R/I \to S$ such that*

$$\sigma = \varrho\pi.$$

Proof. This follows from the above proposition together with corollary 2.6.12. ☐

Corollary 2.6.17. *Let R be a ring and M,N be left/right R-modules. Consider $\rho : M \to N$ a left/right module map and $L \subset M$ a submodule such that $L \subset \ker\ \sigma$. Define $\pi : M \to M/L$ to be the canonical surjection, i.e. $r \mapsto r+L$. Then there is a unique ring homomorphism $\varrho : M/L \to N$ such that*

$$\rho = \varrho\pi.$$

Proof. This follows from the above proposition together with corollary 2.6.14. ☐

Theorem 2.6.18. *(1st Isomorphism Theorem for Groups)*
Let G,H be a group and $\rho : G \to H$ a group homomorphism. Then $G/\ker\ \rho \simeq \rho(G)$ via the group homomorphism $\varrho : G/\ker\ \rho \to H$, $g(\ker\ \rho) \mapsto \rho(g)$

Proof. By Proposition 2.6.15, $\varrho : G/\ker\ \rho \to H$, $g(\ker\ \rho) \mapsto \rho(g)$ is a well-defined group homomorphism. Then $\overline{\varrho} : G/\ker\ \rho \to \overline{\varrho}(G/\ker\ \rho)$, $g(\ker\ \rho) \mapsto \varrho(g)$ is a surjective group homomorphism. We check that $\varrho(G/\ker\ \rho) = \rho(G)$. Indeed, let $\varrho(g(\ker\ \rho)) \in \varrho(G/\ker\ \rho)$. Then $\varrho(g(\ker\ \rho)) = \rho(g) \in \rho(G)$. Similarly, if $\rho(g) \in \rho(G)$, then $\rho(g) = \varrho(g(\ker\ \rho)) \in \varrho(G/\ker\ \rho)$. It remains to check that $\overline{\varrho}$ is injective. Suppose $0 = \overline{\varrho}(g(\ker\ \rho))$. Then $\rho(g) = 0$, which implies $g \in \ker\ \rho$, hence $g(\ker\ \rho) = e(\ker\ \rho)$, meaning $\ker \overline{\varrho} = 0$. By Lemma 2.6.9 $\overline{\varrho}$ is injective. We thus conclude that $\overline{\varrho}$ is a bijective group homomorphism, which means

$$G/\ker\ \rho \simeq \overline{\varrho}(G/\ker\ \rho) = \varrho(G/\ker\ \rho) = \rho(G).$$

☐

Corollary 2.6.19. *(1st Isomorphism Theorem for Rings)*
Let R,S be rings and $\sigma : R \to S$ a ring homomorphism. Set $I = \ker\ \sigma$. Then $R/I \simeq \sigma(R)$ via the ring homomorphism $\vartheta : R/I \to \sigma(R)$, $r+I \mapsto \sigma(r)$.

Proof. By Corollary 2.6.16 and the above theorem ϑ is a bijective ring homomorphism, hence $R/I \simeq \sigma(R)$. ☐

Corollary 2.6.20. *(1st Isomorphism Theorem for Modules) Let R be a ring, M,N be a left/right R-modules and consider a left/right R-module homomorphism $\rho : M \to N$. Then $M/\text{ker } \rho \simeq \rho(M)$ via the left/right R-module homomorphism $\varrho : M/\text{ker } \rho \to \rho(M)$, $m + \text{ker } \rho \mapsto \rho(m)$*

Proof. Theorem 2.6.18 and Corollary 2.6.17 ensures that $\varrho : M/\text{ker } \rho \to \rho(M)$ is a bijective module homomorphism. ☐

Corollary 2.6.21. *(1st Isomorphism Theorem for Algebras) Let R be a commutative ring and A,B be R-algebras. Consider an R-algebra homomorphism $\sigma : A \to B$. Then $A/\text{ker } \sigma \simeq \sigma(A)$ via the R-algebra homomorphism $\vartheta : A/\text{ker } \sigma \to \sigma(A), a + \text{ker } \sigma \mapsto \sigma(a)$.*

Proposition 2.6.22. *Let G be a group, $N', N \subset G$ normal subgroups with $N' \subset N$. Then*

$$\frac{G/N'}{N/N'} \simeq G/N.$$

Proof. Consider the surjective group map

$$\pi : G \to G/N, g \mapsto gN'.$$

Then by Lemma 2.6.11

$$\varpi : G/N' \to G/N, gN' \mapsto gN,$$

is a surjective group map, since $N' \subset \text{ker } \pi$. Clearly $N'/N \subset \text{ker } \varpi$. Let $gN' \in \text{ker } \varpi$. Then $gN = 0$, hence $g \in N$. Thus $gN' \in N/N'$. Then $\text{ker } \varpi = N/N'$, hence by the 1st isomorphism theorem

$$\frac{G/N'}{N/N'} = \frac{G/N'}{\text{ker } \varpi} \overset{\varpi}{\simeq} G/N.$$

☐

Corollary 2.6.23. *Let R be a ring $I,J \subset R$ left/right ideals with $J \subset I$. Then*

$$\frac{R/J}{I/J} \simeq R/I.$$

Proof. Follows from the 1st isomorphism for rings and the fact that π is also a ring homomorphism. ☐

Corollary 2.6.24. *Let R be a ring M a left/right R-module. Consider submodules $N, N' \subset M$ with $N' \subset N$. Then*

$$\frac{M/N'}{N/N'} \simeq M/N.$$

Proof. Follows from the 1st isomorphism theorem for modules together with the fact that π is also a module homomorphism. $\square$

Lemma 2.6.25. *Consider* ${}_RL \leq {}_RN \leq {}_RM$. *Then*

$$0 \longrightarrow N/L \hookrightarrow M/L \twoheadrightarrow M/N \longrightarrow 0$$

is exact

Proof. This is an immediate consequence of Lemma 2.6.4 $\square$

Theorem 2.6.26. *Let* $N, L \leq M$. *Then* $N/(N \cap L) \simeq (N+L)/N$.

Proof. We get the chain of modules $N \cap L \leq N \leq N+L$, hence we have an exact sequence

$$0 \longrightarrow N/(N \cap L) \xrightarrow{\rho} (N+L)/(N \cap L) \xrightarrow{\rho'} (N+L)/N \longrightarrow 0$$

Substituting $(N+L)/(N \cap L)$ for $\mathbf{im}\ \rho$ we preserve exactness, i.e.

$$0 \longrightarrow N/(N \cap L) \xrightarrow{\overline{\rho}} \mathbf{im}\ \rho \xrightarrow{\rho'|_{\mathrm{im}\ \rho}} (N+L)/N \longrightarrow 0$$

where $\overline{\rho}(n + N \cap L) = \rho(n + N \cap L)$. Note that for $n + N \cap L \in \mathbf{im}\ \rho$,

$$\rho'|_{\mathrm{im}\ \rho}(n + N \cap L) = \rho'(n + N \cap L) = n + N = 0 \Rightarrow \rho'|_{\mathrm{im}\ \rho} = 0.$$

It follows that $\mathbf{im}\ \overline{\rho} = \mathbf{ker}\ \ \rho'|_{\mathrm{im}\ \rho} = \mathbf{ker}\ \ 0 = (N+L)/N$, hence ρ is surjective. By exactness ρ is also injective hence $N/(N \cap L) \overset{\rho}{\simeq} (N+L)/N$. $\square$

2.6.3 Free Modules

2.7 Vector Spaces

2.7.1 Finite Dimensional Vector Spaces

Definition 2.7.1. Let V be an n-dimensional vector space over a field K with basis $\mathcal{V} = \{v_1, \ldots, v_n\}$. Let $\mathcal{W} = \{w_1, \ldots, w_n\} \subset V$. Write $w_i = \sum_1^n a_{ij} v_j$ for suitable (unique!) $a_{ij} \in K$ and consider the matrix

$${}_{\mathcal{V}}T_{\mathcal{W}} := (a_{ij})^T \in M_n(K).$$

When ${}_{\mathcal{V}}T_{\mathcal{W}}$ is invertible we call it *the basis transformation of* $\mathcal{V}$ *to* $\mathcal{W}$ or *a change-of-basis*

Remark 2.7.2. We canonically identify ${}_{\mathcal{V}}T_{\mathcal{W}}$ with an endomorphism on V:

$${}_{\mathcal{V}}T_{\mathcal{W}} : V \to V, v = \sum_1^n \alpha_i v_i \mapsto \sum_1^n \left(\sum_1^n a_{ji} \alpha_j \right) v_i$$

Note that

$${}_{\mathcal{V}}T_{\mathcal{W}} v_i = \sum_1^n \left(\sum_1^n a_{kj} \delta_{ki} \right) v_j = \sum_1^n a_{ij} v_j = w_i$$

Theorem 2.7.3. *Let V be an n-dimensional vector space over a field K with basis $\mathcal{V} = \{v_1, \ldots, v_n\}$. Consider $\mathcal{W} = \{w_1, \ldots, w_n\} \subset V$. Then $\mathcal{W}$ is a basis of V over K if and only if ${}_{\mathcal{V}}T_{\mathcal{W}}$ is invertible.*

Proof. "$\Rightarrow$": write $v_i = \sum_1^n b_{ij} w_j$. Note that $w_i =_{\mathcal{V}} T_{\mathcal{W}} v_i =_{\mathcal{V}} T_{\mathcal{W}\mathcal{W}} T_{\mathcal{V}} w_i$, and that $v_i =_{\mathcal{W}} T_{\mathcal{V}} w_i =_{\mathcal{W}} T_{\mathcal{V}\mathcal{V}} T_{\mathcal{W}} v_i$. Since a module homomorphism is uniquely characterized by its behavior on the basis elements to be written it follows that ${}_{\mathcal{V}}T_{\mathcal{W}\mathcal{W}}T_{\mathcal{V}} =_{\mathcal{W}} T_{\mathcal{V}\mathcal{V}} T_{\mathcal{W}} = I_n$, hence ${}_{\mathcal{V}}T_{\mathcal{W}}$ is invertible with inverse ${}_{\mathcal{W}}T_{\mathcal{V}}$.
"$\Leftarrow$": by An invertible linear map, maps a basis to a basis, hence $\mathcal{W} =_{\mathcal{V}} T_{\mathcal{W}}(\mathcal{V})$ is a basis of V. □

Lemma 2.7.4. *Given an exact sequence of vector spaces*

$$0 \longrightarrow V \xrightarrow{\rho} V' \xrightarrow{\rho'} V'' \longrightarrow 0$$

we have that

$$\dim V = \dim V' + \dim V''$$

Proof. This follows from the above Lemma 2.6.4 and theorem not yet written about finite dimensional vector spaces. □

Proposition 2.7.5. *For sequence of vector spaces*

$$0 \longrightarrow V_1 \xrightarrow{\rho_1} V_2 \xrightarrow{\rho_2} V_3 \xrightarrow{\rho_3} V_4 \longrightarrow 0$$

we have that $\dim V_4 = \dim V_3 - \dim V_2 + \dim V_1$.

Proof. Set $W := \operatorname{im} \rho_2 = \ker \rho_3$. Then

$$0 \longrightarrow V_1 \xrightarrow{\rho_1} V_2 \xrightarrow{\rho_2} W \longrightarrow 0$$

and

$$0 \longrightarrow W \hookrightarrow V_3 \xrightarrow{\rho_3} V_4 \longrightarrow 0$$

are exact, hence by the above lemma $\dim V_2 = \dim V_1 + \dim W$ implying $\dim W = \dim V_2 - \dim V_1$. Moreover,

$$\dim V_3 = \dim V_4 + \dim W = \dim V_4 + \dim V_2 - \dim V_1,$$

hence $\dim V_4 = \dim V_3 - \dim V_2 + \dim V_1$. □

Lemma 2.7.6. *Let $U \subset W \subset V$ be vector spaces where V/U finite dimensional. Then*

$$\dim\ V/U = \dim\ V/W + \dim\ W/U,$$

In particular we get that V/W and W/U are finite dimensional.

Proof. This follows directly from the above proposition and Lemma 2.6.25. $\square$

Lemma 2.7.7. *Let*

$$0 \xrightarrow{\rho_0} V_1 \xrightarrow{\rho_1} \cdots \xrightarrow{\rho_{n-1}} V_n \xrightarrow{\rho_n} 0$$

be an exact sequence of finite dimensional vector spaces. Then $\sum_1^n (-1)^i \dim\ V_i = 0$

Proof. In the case $n = 1$, it's easy to see that $V_1 = 0$. Denote the first 0-map In general for a finite sequence of elements in an additive group, $a_0, \ldots, a_n$, say $\sum_1^1 (-1)^i (a_{i-1} + a_i) = -a_0 + (-1)^{n-1} a_n$. By the rank nullity theorem and exactness we have for each $i \in \{1, \ldots, n\}$ that $\dim\ V_i = \dim\ \ker\ \rho_i + \dim\ \mathrm{im}\ \rho_i = \dim\ \mathrm{im}\ \rho_{i-1} + \dim\ \mathrm{im}\ \rho_i$. Hence picking $a_i = \dim\ \mathrm{im}\ \rho_i$, it follows that

$$\begin{aligned}\sum_1^n (-1)^i \dim\ V_i = \sum_1^n (-1)^i (a_{i-1} + a_i) = -a_0 + (-1)^n a_n &= -\dim\ \mathrm{im}\ \rho_0 + (-1)^n \dim\ \mathrm{im}\ \rho_n \\ &= -\dim\ \mathrm{im}\ 0 + (-1)^n \dim\ \mathrm{im}\ 0 = 0.\end{aligned}$$

$\square$

Lemma 2.7.8. *Let V be a finite dimensional vector space and $V_1, \ldots, V_n$ be subspaces. Then*

$$\mathrm{codim}\ \bigcap_1^n V_i \le \sum_1^n \mathrm{codim}\ V_i.$$

Proof. Pick a basis of V, B and bases of $V_1, \ldots, V_n, \bigcap_1^n V_i$, denoted respectively $\mathcal{V}_1, \ldots, \mathcal{V}_n, \mathcal{V} \subset B$. Then

$$\mathrm{codim}\ \bigcap_1^n V_i = \#(B \setminus \mathcal{V}) = \#\left(\bigcup_1^n B \setminus \mathcal{V}_i\right) \le \sum_1^n \#(B \setminus \mathcal{V}_i) = \sum_1^n \mathrm{codim}\ V_i.$$

$\square$

2.7.2 Projective Space

Definition 2.7.9. Let V be a vector space over some field K. For $v, w \in V \setminus 0$ we write $v \sim w$ if there exists a $\lambda \in K \setminus 0$ such that $w = \lambda v$

Remark 2.7.10. This an equivalence relation. Indeed $v = 1 \cdot v$, hence $v \sim v$. If $v \sim w$, then $w = \lambda v$, hence $v = \lambda^{-1} w$, meaning $w \sim v$. Suppose $v \sim w$ and $w \sim u$. Then $w = \lambda v$ and $u = \mu w$, hence $u = \mu\lambda v$, implying $v \sim u$.

Definition 2.7.11. We define *the projective space of V over K* to be the set

$$\mathbb{P}(V) := (V \setminus 0)/\sim .$$

We furthermore define $\mathbb{P}^n := \mathbb{P}^n(K) := \mathbb{P}(K^{n+1})$ which is called *the projective n-space over K*. We denote an element $[v] = [(v_1, \ldots, v_{n+1})] \in \mathbb{P}^n$ by $[v_1, \ldots, v_n]$. We call $\mathbb{P}^1$ *the projective line over K* and $\mathbb{P}^2$ *the projective plane over K*.

Remark 2.7.12. Note that $[\lambda v] = [v]$ for every $\lambda \in K \setminus 0$ and $[v] \in \mathbb{P}(V)$, hence $[\lambda v_1, \ldots, \lambda v_{n+1}] = [v_1, \ldots, v_{n+1}]$ for every $[v_1, \ldots, v_{n+1}] \in \mathbb{P}^n$.
Consider the category of Vector Spaces with morphism being maps that are homogeneous of degree 1. Consider also category of sets P with a $(K^*, \cdot)$-action, satisfying $p = \lambda p$ for every $p \in P$, with morphisms being maps that are homogeneous of degree 1 with respect to this K^*-action. Then $V \mapsto \mathbb{P}(V)$ and $f : V \to W \mapsto \widehat{f} : \mathbb{P}(V) \to \mathbb{P}(W), [v] \mapsto [f(v)]$, defines a functor. Restricting all sets to be topological spaces (with vector spaces being topological vector spaces) and all maps to be continuous, we get two subcategories of category of topological spaces such that the functor $(V, f) \mapsto (\mathbb{P}(f), \widehat{f})$ restricts to a functor between these categories. Indeed if $f : X \to Y$ is continuous and $\pi : X \to X/\sim_X$, $\tau : Y \to Y/\sim_Y$, denotes quotient maps to some quotient spaces and $\widehat{f} : X/\sim_X \to Y/\sim_Y, [x] \mapsto [f(x)]$ is well-defined, then $\widehat{f}\pi = \tau f$, hence $\widehat{f}$ is continuous, by the universal property of quotient space perhaps write some point set topology?.

Definition 2.7.13. For each $i \in \{1, \ldots, n+1\}$ we define *the i'th copy of K^n in $\mathbb{P}^n$* to be the set

$$U_i := \left\{[v_1, \ldots, v_{n+1}] \in \mathbb{P}^n : v_i \neq 0\right\}.$$

We furthermore define *the i'th hyperplane at infinity in $\mathbb{P}^n$* to be the set

$$H_{\infty,i} := \left\{[v_1, \ldots, v_{n+1}] \in \mathbb{P}^n : v_i = 0\right\}.$$

We define the hyperplane at infinity in $\mathbb{P}^n$ to be $H_\infty := H_{\infty,n+1}$

Remark 2.7.14. 1. Suppose V is a topological vector space. Consider the map $m_\lambda : \mathbb{P}(V) \to \mathbb{P}(V), [v] \mapsto [\lambda v]$ for $\lambda \in K \setminus 0$. One clearly has that $m_\lambda = \mathbf{id}$, hence it is continuous.

2. One notes that $\mathbb{P}^n = \bigcup_1^{n+1} U_i$. Note that $\varphi : K^n \to U_i, v \mapsto [v_1, \dots, v_{i-1}, 1, v_{i+1}, \dots, v_n]$ is a bijection. Indeed the map

$$\varphi^{-1} : U_i \to K^n, [v_1, \dots, v_i, \dots, v_{n+1}] \mapsto (v_1/v_i, \dots, v_{i-1}/v_i, v_{i+1}/v_i, \dots, v_{n+1}/v_i)$$

is well-defined, since

$$((\lambda v_1)/(\lambda v_i), \dots, (\lambda v_{i-1})/(\lambda v_i), (\lambda v_{i+1})/(\lambda v_i), \dots, (\lambda v_{n+1})/(\lambda v_i)) = \\ (v_1/v_i, \dots, v_{i-1}/v_i, v_{i+1}/v_i, \dots, v_{n+1}/v_i).$$

Clearly φ and φ^{-1} are mutual inverses. Suppose K is a topological field. Then K^m becomes a topological vector space and we can endow $\mathbb{P}^m$ with the quotient topology. Note that φ is continuous, since it is given by pre-composition of $\iota : K^n \to \pi^{-1}(U_i) \setminus 0, v \mapsto (v_1, \dots, v_{i-1}, 1, v_{i+1}, \dots, v_{n+1})$ with $\pi\mid_{\pi^{-1}(U_i)}: \pi^{-1}(U_i) \to U_i$, which are continuous maps. Let $U \subset K^n$ be open. Let $Q := \{v \in K^{n+1} : v_i \in K, (v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_{n+1}) \in KU\} \simeq_{\mathcal{S}_{n+1}} K \times KU$. Define $O := Q \setminus 0$. One easily verifies that $O = Q \cap K^{n+1} \setminus 0$ and that $\pi^{-1}(\pi(O)) = O$, implying that $\pi(O)$ is open. One checks that $\varphi(U) = U_i \cap \pi(O)$, hence $\varphi(U)$ is open in U_i, hence φ^{-1} is continuous. We thus conclude that K^n is homeomorphic to U_i for each i. Hence $\mathbb{P}^n$ is locally homeomorphic to K^n.

3. Consider the map $\pi\mid_{S^n(\mathbb{R})}: S^n(\mathbb{R}) \to \mathbb{P}^n(\mathbb{R})$, $S^n(K) = \{v \in \mathbb{R}^{n+1} : \|v\| = 1\}$. Then for $[v] \in \mathbb{P}^n(\mathbb{R})$, $[v] = [1/\|v\| v] = \pi_{S^n(\mathbb{R})}(1/\|v\| v)$, hence $\mathbb{P}^n(\mathbb{R})$ is compact. Since $\mathbb{C}^n \simeq \mathbb{R}^{2n}$, it follows from functoriality that $\mathbb{P}^n(\mathbb{C}) \simeq \mathbb{P}^{2n+1}(\mathbb{R})$. Moreover for every $[v] = [w]$ for $v, w \in S^n(\mathbb{R})$, then $w = \lambda v$, hence $1 = \|\lambda v\| = |\lambda| \|v\| = |\lambda|$, hence $w = \pm v$. So $\mathbb{P}^n(\mathbb{R})$ is homeomorphic to the northern hemisphere, i.e. $S^n(\mathbb{R})/(x \sim -x)$.

4. Another thing to note is that $\mathbb{P}^n = U_i \sqcup H_{\infty,i}$

2.7.3 The Projective Span

Definition 2.7.15. For $[v_1], \dots, [v_m] \in \mathbb{P}^n$ we define

$$\mathrm{Span}([v_1], \dots, [v_m]) := \left\{ \left[\sum_1^m \lambda_i v_i \right] : (\lambda_1, \dots, \lambda_m) \in K^m \setminus 0 \right\}$$

Remark 2.7.16. Of course one should ask if this is well-defined. Suppose we are given $\alpha_1, \dots, \alpha_m \in K \setminus 0$. Then for any $(\lambda_1, \dots, \lambda_m) \in K^m \setminus 0$,

$$\sum_1^m \lambda_i (\alpha_i v_i) = \sum_1^m (\lambda_i \alpha_i) v_i \in \left\{ \left[\sum_1^m \lambda_i v_i \right] : (\lambda_1, \dots, \lambda_m) \in K^m \setminus 0 \right\}$$

and conversely

$$\sum_1^m \lambda v_i = \sum_1^m (\lambda_i \alpha_i^{-1}) \alpha_i v_i \in \left\{ \left[\sum_1^m \lambda_i (\alpha_i v_i) \right] : (\lambda_1, \dots, \lambda_m) \in K^m \setminus 0 \right\}.$$

A further thing to note with this construction, is that if $v_1, \dots, v_m \in \mathbb{A}^{n+1} \setminus 0$ span $\mathbb{A}^{n+1}$ (this can only happen for $m \geq n+1$, then $\mathrm{Span}([v_1], \dots, [v_m]) = \mathbb{P}^n$.

Lemma 2.7.17. *$v_1, \dots, v_m \in \mathbb{P}^n$ are linearly independent if and only if $[v_i] \notin \mathrm{Span}([v_1], \dots, \widehat{[v_i]}, \dots, [v_m])$ for any i.*

Proof. "$\Rightarrow$": Suppose there is a $[v_i] \in \mathrm{Span}([v_1], \dots, \widehat{[v_i]}, \dots, [v_m])$. Then for some $\lambda_1, \dots, \lambda_m \in K^m$ with $\lambda_i \neq 0$,

$$-\lambda_i v_i = \sum_{j \neq i} \lambda_j v_j \Rightarrow \sum_1^m \lambda_j v_j = 0,$$

hence $v_1, \dots, v_m$ are not linearly independent. "$\Leftarrow$": Suppose there are $(\lambda_1, \dots, \lambda_m) \in K^m \setminus 0$ such that $\sum_1^m \lambda_i v_i = 0$. Then for some $j \in \{1, \dots, m\}$, $\lambda_j \neq 0$, hence

$$v_j = \lambda_j^{-1} \sum_{i \neq j} \lambda_i v_i \in \mathrm{Span}([v_1], \dots, \widehat{[v_j]}, \dots, [v_m])$$

□

2.7.4 Normed Vector Spaces

Definition 2.7.18. An *ordered field* is a field K together with a total ordering $\leq$ on K, satisfying, for every $a, b, c \in K$

1. $a \leq b \Rightarrow a + c \leq b + c$
2. $0 \leq a, 0 \leq b \Rightarrow 0 \leq ab$

Example 2.7.19. In this example we endow $\mathbb{Q}$ with the structure of ordered field. For $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$ we define $\frac{a}{b} \leq \frac{c}{d}$ if $ad \leq cb$ and $b, d > 0$. Note that we can always find a representative of a rational number whose numerator is greater than 0. It is easy to check that this is a partial order on $\mathbb{Q}$. Given any two $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$ with $b, d > 0$ we get that $bd > 0$ and $ad \leq bc$ or $ad \geq bc$, since $(\mathbb{Z}, \leq)$ is totally ordered. It follows that $\leq$ is a total ordering. Suppose $\frac{x}{y}, \frac{z}{w}, \frac{v}{u} \in \mathbb{Q}$ with $y, w, u > 0$ are given. Suppose $\frac{x}{y} \leq \frac{z}{w}$. Then $yu, wu > 0$ and

$$wu(xu + vy) \leq u(yzu + vwy) = yu(zu + vw) \Rightarrow \frac{x}{y} + \frac{v}{u} = \frac{xu + vy}{yu} \leq \frac{zu + vw}{wu} = \frac{z}{w} + \frac{v}{u}.$$

Suppose instead now that $0 \leq \frac{x}{y}$ and $0 \leq \frac{z}{w}$. Then $0 \leq x$ and $0 \leq z$, hence $0 \leq xz$, meaning $0 \leq \frac{xz}{yw}$.

Definition 2.7.20. On an ordered field K, we define *the absolute value* to be the function $|\cdot|: K \to K_{\geq 0} := \{a \in K : a \geq 0\}$ to be given by

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases}$$

Definition 2.7.21. A *normed vector space* is a vector space V over an ordered field K with a map $\|\cdot\|: V \to K$ satisfying

1. For every $v \in V$, $\|v\| \geq 0$.
2. For every $v \in V$, $\|v\| = 0 \iff v = 0$.
3. For every $v \in V$, $a \in K$, $\|av\| = |a|\|v\|$.
4. For every $v, w \in V$, $\|v + w\| \leq \|v\| + \|w\|$.

We call such a function *a norm on V over K*.

Lemma 2.7.22. *On an ordered field K, the absolute value defines a norm on K over K, making K a normed vector space over K.*

Proof. 1. This is trivial, since if $a < 0$, then $-a > 0$.
2. Suppose $|a| = 0$. Then $a \geq 0$, hence $a = |a| = 0$.
3. Let $a, b \in K$. Then if $a, b \geq 0$ we have that $ab \geq 0$, hence $|ab| = ab = |a||b|$. If $a, b < 0$, then $-a, -b > 0$, hence $ab = (-a)(-b) > 0$. It follows that $|ab| = ab = (-a)(-b) = |a||b|$. If $a \geq 0$ and $b < 0$, then $ab \leq 0$, hence $|ab| = -ab = a(-b) = |a||b|$. The case $a < 0$ and $b \geq 0$ is symmetric.
4. Let $a, b \in K$. Observe that in any case $-c, c \leq |c|$ for any $c \in K$. If $a + b \geq 0$, then $|a + b| = a + b \leq |a| + |b|$. In the other case $|a + b| = -a - b \leq |a| + |b|$. □

2.8 Ring theory

2.8.1 Matrix Rings

Definition 2.8.1. Let R be a ring and n, m be positive integers. We define *the set of $n \times m$ (n by m) matrices over R* to be the set

$$M_{n\times m}(R) = \prod_{i\in\{1,\ldots,n\}, j\in\{1,\ldots,m\}} R.$$

For an element $(a_{i,j}) \in M_{n\times m}(R)$ we define $(a_{ij}) := (a_{i,j})$, when no disambiguity arises from this notation. An element of $M_{n\times m}(R)$ is called an $(n \times m)$ *matrix.* We define $M_n(R) := M_{n\times n}(R)$.

Remark 2.8.2. By Lemma 2.2.19 $M_{n,m}(R)$ is an additive group.

Example 2.8.3. Let $R = \mathbb{Z}$, $n = 2$ and $m = 3$ and consider $a_{11} = 1, a_{12} = 2, a_{13} = 3$ and $a_{21} = 2, a_{22} = 3, a_{23} = 2$. We opt to write the element (a_{ij}) as table with 2 rows and 3 columns, i.e.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \end{pmatrix} := (a_{ij}).$$

For arbitrary rings positive integers n, m, we in general can write an element $(a_{ij}) \in M_{n\times m}$ as a table with n rows and m columns, i.e.

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} := (a_{ij})$$

Lemma 2.8.4. *Let R be a ring and n, m, l be positive integers. We define matrix multiplication to be the operation*

$$\cdot : M_{n\times m}(R) \times M_{m\times l}(R) \to M_{n\times l}(R)$$

defined by

$$(a_{ij})(b_{ij}) := \left(\sum_{k=1}^{m} a_{ik} b_{kj} \right).$$

Suppose we have an additional positive integer p and let $(a_{ij}) \in M_{n\times m}(R), (b_{ij}) \in M_{m\times l}(R), (c_{ij}) \in M_{l\times p}(R)$. Then

$$(a_{ij})\big((b_{ij})(c_{ij})\big) = \big((a_{ij})(b_{ij})\big)(c_{ij}).$$

Suppose $(a_{ij}) \in M_{n\times m}(R), (b_{ij}), (c_{ij}) \in M_{m\times l}(R)$. Then

$$(a_{ij})\big((b_{ij}) + (c_{ij})\big) = (a_{ij})(b_{ij}) + (a_{ij})(c_{ij})$$

Proof. Indeed for $(a_{ij}) \in M_{n\times m}(R), (b_{ij}) \in M_{m\times l}(R), (c_{ij}) \in M_{l\times p}(R)$,

$$\begin{aligned} (a_{ij})\big((b_{ij})(c_{ij})\big) &= (a_{ij}) \left(\sum_{k=1}^{l} b_{yk} c_{kz} \right) = \left(\sum_{h=1}^{m} a_{xh} \sum_{k=1}^{l} b_{hk} c_{kz} \right) = \left(\sum_{h=1}^{m} \sum_{k=1}^{l} a_{xh} (b_{hk} c_{kz}) \right) \\ &= \left(\sum_{k=1}^{l} \sum_{h=1}^{m} (a_{xh} b_{hk}) c_{kz} \right) = \left(\sum_{k=1}^{l} \left(\sum_{h=1}^{m} a_{xh} b_{hk} \right) c_{kz} \right) = \\ &= \left(\sum_{h=1}^{m} a_{xh} b_{hk} \right) (c_{ij}) = \big((a_{ij})(b_{ij})\big)(c_{ij}) = \big((a_{ij})(b_{ij})\big)(c_{ij}). \end{aligned}$$

Furthermore, for $(a_{ij}) \in M_{n\times m}(R), (b_{ij}), (c_{ij}) \in M_{m\times l}(R)$,

$$\begin{aligned}(a_{ij})\big((b_{ij})+(c_{ij})\big) &= (a_{ij})(b_{ij}+c_{ij}) = \left(\sum_{k=1}^{m} a_{ik}(b_{kj}+c_{kj})\right) = \left(\sum_{k=1}^{m} a_{ik}b_{kj}+a_{ik}c_{kj}\right)\\ &= \left(\sum_{k=1}^{m} a_{ik}b_{kj}+\sum_{k=1}^{m} a_{ik}c_{kj}\right) = \left(\sum_{k=1}^{m} a_{ik}b_{kj}\right)+\left(\sum_{k=1}^{m} a_{ik}c_{kj}\right)\\ &= (a_{ij})(b_{ij})+(a_{ij})(c_{ij})\end{aligned}$$

□

Lemma 2.8.5. *Let R be a ring and n a positive integer. Then $(M_n(R),+,\cdot)$, where $\cdot$ is matrix multiplication of matrices in $M_n(R)$ and $M_n(R)$, is a ring called the $(n\times n)$ matrix ring (over R).*

Proof. $M_n(R)$ is an additive group by Remark 2.8.2. Let $I := 1 := (\delta_{ij})$ (where $\delta_{ii} = 1$ for $i \in \{1,\dots,n\}$ and $\delta_{ij} = 0$ for $i,j \in \{1,\dots,n\}$ with $i \neq j$). Then for $(r_{ij}) \in M_n(R)$

$$1(r_{ij}) = \left(\sum_{k=1}^{n} \delta_{ik}r_{kj}\right) = \left(\delta_{ii}a_{ij} + \sum_{k\in\{1,\dots,n\}\setminus\{i\}} \delta_{ik}a_{kj}\right) = (1a_{ij}+0) = (a_{ij}).$$

In a dual way one can prove that

$$(a_{ij})1 = (a_{ij}).$$

By Lemma 2.8.4 it follows that $(M_n(R),+,\cdot)$ is a ring. □

Example 2.8.6. A matrix ring is never a commutative ring: Take $(a_{ij}) \in M_n(R)$ where $a_{11} = 1$ and $a_{ij} = 0$ when $a_{ij} = 0$ for $i \neq 1$ or $j \neq 1$. Take $(b_{ij}) \in M_n(R)$ where $b_{1m} = 1$ and $b_{ij} = 0$ when $i \neq 1$ or $j \neq m$. Then it is easy to check that $(a_{ij})(b_{ij}) = (b_{ij})$ while $(b_{ij})(a_{ij}) = 0$.

2.8.2 Fields, Integral Domains & some Important Ideals

Definition 2.8.7. Let R be a ring. An element $r \in R$ is called a *unit* if there is an element $r' \in R$ such that $rr' = r'r = 1$. We denote the set of units of R by R^*.

Remark 2.8.8. I. Suppose there are two elements r_1, r_2 such that $rr_i = r_ir = 1$. Then

$$r(r_1-r_2) = rr_1 - rr_2 = 1-1 = 0 \Rightarrow r_1 - r_2 = r_1r(r_1-r_2) = 0 \Rightarrow r_1 = r_2.$$

Hence an element satisfying $rr' = r'r = 1$ is unique. We denote it by r^{-1} and refer to it as the multiplicative inverse of r.

II. $(R^*,\cdot)$ is a group. We check that R^* is a submonoid of R and hence a monoid. Clearly $1 \in R^*$. Let $r, r' \in R^*$. Then

$$\left(rr'\right)\left(r'^{-1}r^{-1}\right) = r\left(r'\left(r'^{-1}r^{-1}\right)\right) = r\left(\left(r'r'^{-1}\right)r^{-1}\right) = r\left(er^{-1}\right) = rr^{-1} = 1,$$

and similarly one can check that $\left(r'^{-1}r^{-1}\right)(rr') = 1$, hence $rr' \in R^*$ a. Let $r \in R^*$. Then $r^{-1}r = rr^{-1} = 1$, hence r^{-1} is a unit, hence R^* is a group.

III. If R is commutative it is sufficient to check that $rr' = 1$ to verify that r is a unit.

Definition 2.8.9. A commutative ring K where $K^* = K \setminus \{0\}$ is called a *field*. Removing the restriction of K being commutative, K is called a *division ring* or *skew field*.

Definition 2.8.10. An left/right ideal $M \subsetneq R$ is called a *left/right maximal* ideal, if for every left/right ideal $I \subset R$ with $M \subset I$, either $I = M$ or $I = R$. If M is an ideal with aforementioned property it is called a *maximal ideal*.

Definition 2.8.11. A ring R is called *simple* if the only ideals in R are the trivial ones, i.e. 0 and R.

Lemma 2.8.12. *Any division ring is simple.*

Proof. Let D be a division ring and consider a non-zero ideal $I \subset R$. Then there is an $x \in I \setminus 0$. Since D is a division ring, there exists x^{-1} s.t. $1 = x^{-1}x \in I$, meaning $I = D$. □

Lemma 2.8.13. *Let R be a ring. Let $I \subset R$ be an ideal. Then I is a maximal ideal if and only if R/I is a simple.*

Proof. "$\Rightarrow$": Let $J/I \subset R/I$ be a non-zero ideal, i.e. assume $I \subsetneq J \subset R$ for some ideal J in R. Then $J = R$, hence $J/I = R/I$, implying R/I is simple.
"$\Leftarrow$": Conversely, consider and ideal $J \subset R$ such that $I \subsetneq J$. Then $0 \neq J/I \subset R/I$, implying that $J/I = R/I$ and hence that $J = R$. □

Proposition 2.8.14. *Let R be a commutative ring. Let $I \subset R$ be an ideal. Then I is a maximal ideal if and only if R/I is a field.*

Proof. "$\Rightarrow$": We need to prove that every non-zero element of R/I is a unit. Consider $a + I \in R/I \setminus \{0 + I\}$, i.e. an element in R/I, where $a \notin I$. Since I is maximal we have that $I + Ra = R = R1$. This implies that we can find $b \in I$ and $r \in R$ such that $1 = b + ra$, hence

$$(r + I)(a + I) = (ra + I) + (b + I) = (ra + b) + I = 1 + I,$$

and since R/I is commutative, this implies $a+I$ is a unit.
"$\Leftarrow$": Since R/I is a field, it is in particular a division ring. Then by Lemma 2.8.12 R/I is simple. By Lemma 2.8.14 I is maximal. □

Definition 2.8.15. Let R be a ring, $r \in R$. $d \in R$ is called a *left divisor of* r if $r \in Rd$ and a *right divisor of* r if $r \in dR$. If d is both a left and a right divisor of r, we write $d \mid r$. Hence if R is commutative, $d \mid r \iff r \in \langle d \rangle$.

Definition 2.8.16. Let R be a ring. An element $a \in R$ is called a *left/right zero divisor* if there is an element $r \in R \setminus 0$ such that $ar = 0$ respectively $ra = 0$. In a commutative ring an element is a left zero divisor if and only if it is a right zero divisor, hence we just call a left/right zero divisor in a commutative ring a *zero divisor*.

Definition 2.8.17. Let R be a ring. If the only left/right zero divisor of R is 0, then R is called a *left/right domain*. If R is a commutative ring and a domain it is called an *integral domain*.

Lemma 2.8.18. *Suppose $S \supset R$ is a ring extension where S is a left/right domain. Then so is R.*

Proof. Let $a \in R \setminus 0$, hence in particular in $S \setminus 0$ then for every $R \setminus 0$, $ar \neq 0$. □

Proposition 2.8.19. *A division ring D is a domain. Hence a field is an integral domain.*

Proof. Let $a \in D \setminus 0$. Then $\langle a \rangle = D$ since D is simple, hence $1 \in \langle a \rangle$, meaning there is some $b \in a$ such that $ba = ab = 1$. □

Lemma 2.8.20. *Let $x, y, a \in R$ where a is not a zero divisor. If $ax = ay$, then $x = y$. The same result can be proven if $xa = ya$. In particular, if R is a domain, then for $x, y \in R$, $a \in R \setminus 0$, $ax = ay$ xor $xa = ya$ implies $x = y$.*

Proof. We have that $a(x-y) = ax - ay = 0$ implies $x - y = 0$. The other result is dual. □

Definition 2.8.21. Let R be a commutative ring. An ideal $I \subsetneq R$ is said to be *prime* if for any $a, b \in R$ such that $ab \in I$, then $a \in I$ or $b \in I$

Lemma 2.8.22. *Let R be a commutative ring and $I \subset R$ and ideal. Then I is prime if and only if R/I is an integral domain*

Proof. "$\Rightarrow$": Let $a+I, b+I \in R/I$ be given such that $ab+I = 0+I$. Then $ab \in I$, hence by assumption $a \in I$ or $b \in I$, meaning $a+I = 0$ or $b+I = 0$.
"$\Leftarrow$": Let $a, b \in R$ such that $ab \in I$. Then $ab+I = 0+I$, hence by assumption $a+I = 0$ or $b+I = 0$, hence $a \in I$ or $b \in I$. $\square$

Corollary 2.8.23. *A maximal ideal M in commutative ring R is prime.*

Proof. Indeed R/M is a field and a field is an integral domain, hence M is prime. $\square$

Proposition 2.8.24. *Let $S \supset R$ be a commutative ring extension. Let $I \subset S$ be prime. Then $I \cap R \subset R$ is prime.*

Proof. By Proposition 2.8.22 S/I is an integral domain. Note that $S/I \supset (R+I)/I$ (cf. res saying $R+I$ is subring). Hence from Lemma 2.8.18 $(R+I)/I$ is an integral. It follows from Isomorphism theorem not yet written that $(R+I)/I \simeq R/(I \cap R)$, hence $I \cap R$ is prime. $\square$

Definition 2.8.25. Let R be a commutative ring. A non-unit and non-zero element $p \in R$ is called a *prime element* if for every $a, b \in R$, $p \mid ab$ implies $p \mid a$ or $p \mid b$.

Lemma 2.8.26. *Let R be a commutative ring. Then for a non-zero $p \in R \setminus R^*$, p is prime if and only if $\langle p \rangle$ is prime.*

Proof. This is seen by the fact that $x \in \langle p \rangle$ is by definition equivalent to $p \mid x$. $\square$

Definition 2.8.27. Consider a commutative ring R and an ideal $I \subset R$. We define the *radical* of I to be the set

$$\mathbf{rad}(I) = \left\{ r \in R : r^n \in I \text{ for some } n > 0 \right\}.$$

An ideal with the property that $I = \mathbf{rad}(I)$ is called a *radical ideal.*

Remark 2.8.28. Trivially we have that if $a \in I$, then $a = a^1 \in I$, hence $a \in \mathbf{rad}(I)$. In other words, $I \subset \mathbf{rad}(I)$.

Lemma 2.8.29. *The radical of an ideal $I \subset R$ is an ideal in R.*

Proof. Let $a, b \in \mathbf{rad}(I)$ and $r \in R$. Clearly $0^1 = 0 \in I$, hence $0 \in \mathbf{rad}(I)$. For some $n, m > 0$, $a^n, b^m \in I$. Thus, we also have that

$$(a+b)^{n+m} = \sum_{0}^{n+m} \binom{n+m}{k} a^{n+m-k} b^k, \tag{1}$$

For $k \in \{0,\dots,m\}$, $n+m-k \geq n$, implying $a^{n+m-k} \in I$. For $k \in \{m+1,\dots,n\}$, $b^k \in I$. Then using (1), it follows that $(a+b)^{n+m} \in I$ and hence that $a+b \in \mathrm{rad}(I)$. Finally we also have that

$$(ra)^n = r^n a^n \in I \Rightarrow ra \in \mathrm{rad}(I).$$

□

Lemma 2.8.30. *Let $I \subsetneq R$ be a prime ideal. Then I is a radical ideal.*

Proof. Let $a \in I$ and $n > 0$. We prove by induction in n that if $a^n \in I$ then $a \in I$. For $n = 1$, $a = a^1 \in I$. Suppose $a^{n+1} \in I$. Then $aa^n \in I$. Using that I is prime we get that $a \in I$ or $a^n \in I$. If we land in the first case, we are done. In the second case it follows by induction that $a \in I$. From the above it follows that if $a \in \mathrm{rad}(I)$, then $a \in I$. Hence it follows from Remark 2.8.28 that $I = \mathrm{rad}(I)$. □

The following definition will be important way later on.

Definition 2.8.31. Let R be a commutative ring. We define the *spectrum of R* to be the set

$$\mathrm{Spec}\, R := \{I \subset R : I \text{ is a prime ideal}\}$$

Proposition 2.8.32. *Let R be a commutative ring and $I \subset R$ an ideal. Then there is a one-to-one correspondence between radical/prime/maximal ideals in R containing I and radical/prime/maximal ideals in R/I*

Proof. **Radical:** Let $J \subset R$ be a radical ideal containing I. Let $x+I \in \mathrm{rad}(J/I)$, then for some $n \geq 1$, $x^n + I \in J/I$, hence $x^n \in J$, implying $x \in \mathrm{rad}(J) = J$. This means $x+I \in J/I$.
Let $K \subset R/I$ be a radical ideal. Then $K = J/I$ for some ideal $J \subset R$ containing I. Let $x \in \mathrm{rad}(J)$. Then for some $n \geq 1$, $x^n + I \in J/I$, hence $x+I \in \mathrm{rad}(J/I) = J/I$, implying $x \in J$.
Prime: J/I is prime if and only if $R/J \simeq \frac{R/I}{J/I}$ is an integral domain which is equivalent to J being prime.
Maximal: J/I is maximal if an only if $R/J \simeq \frac{R/I}{J/I}$ is maximal which is equivalent to J being maximal. □

Lemma 2.8.33. *Let I, J be ideals in a commutative ring R. Suppose $I = \langle a_1 \dots, a_m \rangle$ for suitable $a_1, \dots, a_m \in I$ and $I \subset \mathrm{rad}(J)$. Then $I^n \subset J$ for some $n \geq 0$.*

Proof. Let $n_i \geq 0$ be given such that $a_i^{n_i} \in J$. Let $n = \sum_1^m n_i$. We prove the statement by induction in m. For $m = 1$ the statement is trivial.

$$\lambda_{i,j} \in R \quad (i \in \{1,\ldots,m\}, j \in \{1,\ldots,2n\})$$

Then

$$\prod_1^n \left(\sum_1^m \lambda_{i,j} a_i\right) = \sum_{v \in \mathbb{N}^m} \mu_v a_1^{v_1} \cdots a_m^{v_m}.$$

A simple induction argument shows that if $v_i < n_i$ for some i then $v_j > n_j$ for some j, hence $a_i^{v_1} \cdots a_m^{v_m} \in J$ for each $v \in \mathbb{N}^m$ with $\sum_1^m v_i = n$. It follows that $\prod_1^n \left(\sum_1^m \lambda_{i,j} a_i\right) \in J$, hence $I^n \subset J$. □

2.8.3 Comaximal ideals

In this subsection every ring will be assumed commutative.

Definition 2.8.34. Let R be a ring. A pair of ideals I, J in R are said to be *comaximal* if $I + J = R$.

Lemma 2.8.35. *Let I, J be comaximal ideals in a ring R. Then*

$$IJ = I \cap J$$

Proof. The first inclusion is implied by Lemma 2.4.41. Let $a \in I \cap J$. Since I and J are comaximal we can write $1 = i + j$ for suitable $i \in I$ and $j \in J$. Then

$$a = a(i + j) = ai + aj = ia + aj \in IJ,$$

since $ia, aj \in IJ$. □

Lemma 2.8.36. *Let I, J be comaximal ideals in a ring R. Then I^n and J^m are comaximal for every $n, m \geq 1$.*

Proof. **Claim 1:** We first show that I, J^m are comaximal for every $m \geq 1$ by way of induction in m. The base case is true by assumption. Let $m \geq 1$ and $x \in R$. By induction $R = I + J^m$, hence $x = a + b$ for suitable $a \in I$ and $b \in J^m$. Moreover, $1 = i + j$ for suitable $i \in I$ and $j \in J$. Then

$$x = a + b = a + b(i + j) = (a + bi) + bj \in I + J^{m+1}.$$

We now fix $m \geq 1$ it follows by a similar induction argument in n that I^n and J^m are comaximal. □

Lemma 2.8.37. *Let R be a ring. Consider ideals $I_1, \ldots, I_N$ in R and set $J_i := \bigcap_{j \neq i} I_j$. Suppose I_i and J_i are comaximal for each i. Then*

$$\bigcap_1^N I_i^n = \left(\prod_1^N I_i\right)^n = \left(\bigcap_1^N I_i\right)^n$$

Proof. Let $n \geq 1$. Note for each $N \geq 1$, $I_1 J_1 = I_1 \cap J_1$, by lemma 2.8.35, hence by induction we have that $\prod_1^N I_i = \bigcap_1^N I_i$, hence for each $n \geq 1$, $\left(\prod_1^N I_i\right)^n = \left(\bigcap_1^N I_i\right)^n$. By assumption and induction I_1^n and $\bigcap_2^{N+1} I_i^n = \left(\bigcap_2^{N+1} I_i\right)^n$ are comaximal, hence

$$\bigcap_1^{N+1} I_i = I_1^n \cap \bigcap_2^{N+1} I_i^n = \prod_1^{N+1} I_i^n = \left(\prod_1^{N+1} I_i\right)^n.$$

□

2.8.4 Greatest Common Divisor and Least Common Multiples

Definition 2.8.38. A *greatest common divisor* of two elements a, b in a commutative ring is an element d where for every $c \in R$ such that $c \mid a$ and $c \mid b$ we have that $c \mid d$.

Remark 2.8.39. Note that $\gcd(a, 0) = a$ since if $a \mid a$ and any element divides 0.

Definition 2.8.40. A *Least common multiple* of two elements a, b in a commutative ring is an element $m \in R$ where for every $c \in R$ such that $a \mid c$ and $b \mid c$ we have that $m \mid c$.

Remark 2.8.41. $\mathrm{lcm}(a, 0) =$ since 0 is the only element that is a multiple of 0.

2.8.5 Unique Factorization Domains and Euclidean Domains

In our exploration of unique factorization domains and Euclidean Domains we will mean fix an integral domain R.

Definition 2.8.42. Let a non-unit and non-zero element $a \in R$ be given. a is said to be an *irreducible element* if for every $b, c \in R$

$$a = bc \Rightarrow b \in R^* \text{ or } c \in R^*.$$

Lemma 2.8.43. *A prime element is irreducible.*

Proof. Let $p \in R$ be prime. Consider $a, b \in R$ such that $p = ab$, then $p \mid ab$, hence either $p \mid a$ or $p \mid b$. Suppose $p \mid a$. Then $a = pr$ for some $r \in R$. This means $p = prb$, which by Lemma 2.8.20 means $rb = 1$, hence that b is a unit. In the case $p \mid b$, we can similarly show that a is a unit. It thus follows that p is irreducible. □

Definition 2.8.44. R is called a *Unique factorization domain (UFD)* if for every $r \in R \setminus \{R^* \cup \{0\}\}$ has unique factorization into irreducible elements, i.e. there are distinct irreducible elements $p_1, \ldots, p_n \in R$ unique and $v_1, \ldots, v_n \geq 1$ such that

$$r = \prod_1^n p_i^{v_i}.$$

Remark 2.8.45. By uniqueness we more precisely mean that given $q_1, \ldots, q_m \in R$ another sequence of distinct irreducible elements and $w_1, \ldots, w_m \geq 1$ such that

$$r = \prod_1^m q_i^{w_i},$$

then $m = n$ and there is some bijection $\omega : \{1, \ldots, n\} \to \{1, \ldots, n\}$ (i.e. a permutation $\omega \in \mathscr{S}_n$) and units $a_1, \ldots, a_n \in R$ such that

$$p_i = a_i q_{\tau(i)} \text{ and } v_i = w_{\tau(i)},$$

for each $i \in \{1, \ldots, n\}$.

Proposition 2.8.46. *Let R be a ring in which every element that is not zero or a unit can be written as a product of irreducible elements. R is a UFD if and only if every irreducible element is a prime.*

Proof. "$\Rightarrow$": Let $p \in R$ be irreducible and suppose there are $a, b \in R$ such that $p \mid ab$. We aim to prove that $p \mid a$ or $p \mid b$. Since $p \mid 0$, we are done if $a = 0$ or $b = 0$. So assume $a, b \neq 0$. In general for $x, y, z \in R$ if $x \mid y$, then $x \mid yz$. Hence if a is a unit, then $p \mid b = a^{-1}(ab)$, and similary if b is a unit then $p \mid a = b^{-1}(ab)$. So assume that a and b are not units. For some $q \in R$, $pq = ab$. q is not a unit, for otherwiser $p = abq^{-1}$ contradicting the irreducibltiy of p. We can then find irreducible $q_1, \ldots, q_n, p_1, \ldots, p_m, p_{m+1}, \ldots, p_l \in R$ and $v_1, \ldots, v_m, w_1, \ldots, w_m, w_{m+1}, \ldots, v_l \geq 1$ such that

$$q = \prod_1^n q_i^{v_i} \text{ and } a = \prod_1^m p_i^{w_i} \text{ and } b = \prod_{m+1}^l p_i^{w_i}.$$

From this it follows that

$$p \prod_1^n q_i^{v_i} = \prod_1^l p_i^{w_i}.$$

Since the above is a factorization into irreducible it follows from the assumption that R is a UFD that there exists an $i \in \{1, \ldots, l\}$ and a unit $s \in R$ such that $w_i = 1$ and $p = sp_i$. If $i \in \{1, \ldots, m\}$ then

$$a = \prod_1^m p_j^{w_j} = s^{-1} p \prod_{j \in \{1, \ldots, m\} \setminus \{i\}} p_j^{w_j} \Rightarrow p \mid a.$$

By a similar argument, if $i \in \{m+1,\dots,l\}$, then $p \mid b$. "$\Leftarrow$": Suppose there are irreducible elements $p_1,\dots,p_n,q_1,\dots,q_m \in R$ and positive integers $v_1,\dots,v_n,w_1,\dots,w_m \geq 1$ such that

$$\prod_1^n p_k^{v_k} = \prod_1^m q_k^{w_k}.$$

Let $i \in \{1,\dots,n\}$. Then $p_i \mid \prod_1^m q_k^{w_k}$. By assumption p_i is prime, hence $p_i \mid q_{\tau(i)}$ for some $\tau(i) \in \{1,\dots,m\}$, hence for some $s_i \in R$, $q_{\tau(i)} = s_i p_i$, by irreducibility, we get that s_i is a unit. Similarly for $j \in \{1,\dots,m\}$, we can find $\omega(j) \in \{1,\dots,n\}$ and $t_j \in R^*$ such that $p_{\omega(j)} = t_j q_j$. Thus $p_i = s_i q_{\tau(i)} = s_i t_j p_{\omega(\tau(i))}$, hence $\omega(\tau(i)) = i$. Conversely one can show that $\tau(\omega(j)) = j$, hence $n = m$ and τ is a bijection. Now we show that $v_i = w_{\tau(i)}$ for each i. WLOG $v_i \geq w_{\tau(i)}$, Then

$$p_i^{v_i - w_{\tau(i)}} p^{w_{\tau(i)}} \prod_{k \in \{1,\dots,n\} \setminus \{i\}} p_k^{v_k} = p_i^{v_i} \prod_{k \in \{1,\dots,n\} \setminus \{i\}} p_k^{v_k} = p_i^{w_{\tau(i)}} \prod_{k \in \{1,\dots,n\} \setminus \{\tau(i)\}} p_k^{w_k},$$

which implies that

$$p_i^{v_i - w_{\tau(i)}} \prod_{k \in \{1,\dots,n\} \setminus \{i\}} p_k^{v_k} = \prod_{k \in \{1,\dots,n\} \setminus \{\tau(i)\}} p_k^{w_k} \Rightarrow p^{v_i - w_{\tau(i)}} \mid \prod_{k \in \{1,\dots,n\} \setminus \{\tau(i)\}} p_k^{w_k},$$

and if $v_i - w_{\tau(i)} \neq 0$ then since p_i is prime $p_i \mid p_k$ for some $k \in \{1,\dots,n\} \setminus \{\tau(i)\}$ which is not possible since p_i and p_k are distinct. So we conclude that $v_i = w_{\tau(i)}$. $\square$

Definition 2.8.47. Let R be a UFD $\mathscr{P}$ be the set of prime/irreducible elements in R. We say two element $p,q \in \mathscr{P}$ are *associated* if there is a unit $a \in R$ such that $p = aq$.

Remark 2.8.48. Write $p \sim q$ if p and q are associated. Being associated is an equivalence relation on $\mathscr{P}$. Indeed for $p,q,r \in \mathscr{P}$. If $p = 1p$ implying $p \sim p$. $p \sim q$, then for some $a \in R^*$, $p = aq$, hence $q = a^{-1}p$, hence $q \sim p$. Suppose $p \sim q$, $q \sim p$, then for $a,b \in R^*$, $p = aq$ and $q = br$. Then $p = (ab)r$, hence $p \sim r$. We may then write any element as a product over $\mathscr{P}$

$$\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p},$$

where v_p is equal to 0 for all but finitely many p.

Lemma 2.8.49. *Let R be a UFD. Then $\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p} \mid \prod_{[p]_\sim \in \mathscr{P}/\sim} p^{w_p}$ if and only if $v_p \leq w_p$ for every $p \in \mathscr{P}$.*

Proof. We that

$$\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{w_p} = \left(\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p} \right) \left(\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{u_p} \right) = \prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p + u_p}$$

for suitable $u_p \geq 0$. By uniqueness $v_p + u_p = w_p$ implying that $v_p \leq w_p$ for every $p \in \mathscr{P}$. Conversely if $v_p \leq w_p$ then there is some $u_p \geq 0$ such that $v_p + u_p = w_p$ and hence

$$\left(\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p}\right)\left(\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{u_p}\right) = \left(\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p+u_p}\right) = \left(\prod_{[p]_\sim \in \mathscr{P}/\sim} p^{w_p}\right)$$

□

Lemma 2.8.50. *Let $a, b \in R \setminus 0$. Then $a = \prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p(a)}$ and $b = \prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p(b)}$. One finds that*

$$\gcd(a,b) = \prod_{[p]_\sim \in \mathscr{P}/\sim} p^{\min(v_p(a), v_p(b))} \text{ and } \operatorname{lcm}(a,b) = \prod_{[p]_\sim \in \mathscr{P}/\sim} p^{\max(v_p(a), v_p(b))},$$

and these are unique up to multiplication by units.

Proof. Let $c \in R$ such that $c \mid a$ and $c \mid b$, then

$$c = \prod_{[p]_\sim \in \mathscr{P}/\sim} p^{v_p(c)}$$

with $v_p(c) \leq v_p(a)$ and $v_p(c) \leq v_p(b)$ by the above lemma, hence $v_p(c) \leq \min(v_p(a), v_p(b))$. Suppose $d \in R$ is a greatest common divisor of a and b, Then $d \mid \gcd(a,b)$ and $\gcd \mid d$, hence $v_p(d) = \max(v_p(a), v_p(b))$. Let $c \in R$ such that $a \mid c$ and $b \mid c$. Then $v_p(a) \leq v_p(c)$ and $v_p(b) \leq v_p(c)$ by the above lemma, hence $\max(v_p(a), v_p(b)) \leq v_p(c)$. Showing that $\operatorname{lcm}(a,b)$ is that unique up to multiplication by a unit is similar to the $\gcd$-case. □

Lemma 2.8.51. *Let R be an integral $r = \prod_1^m p_i^{v_i} \in R$ where $p_1, \ldots, p_m \in R$ are distinct primes and $v_1, \ldots, v_m \geq 1$. Then $\operatorname{rad}(\langle r \rangle) = \left\langle \prod_1^m p_i \right\rangle$.*

Proof. Clearly $\prod_1^m p_i \in \operatorname{rad}(\langle r \rangle)$, since for $n = \max v_i$, $r = \prod_1^m p_i^{v_i} \mid \prod_1^m p_i^n$. Conversely, if $a \in \operatorname{rad}(\langle r \rangle)$, then for some $n \geq 0$, $r \mid a^n$, implying $\prod_1^m p_i \mid a^n$, hence $\prod_1^m p_i \mid a$. □

Lemma 2.8.52. *Let R be a UFD. The prime ideals of R are R, 0 and principal ideals generated by $\langle p \rangle$ for an irreducible element $p \in R$. This means that a proper non-zero prime ideal in a UFD contains no non-trivial prime ideal.*

Proof. This follows from Lemma 2.8.26 and Lemma 2.8.46. A non-trivial prime ideal contained in $\langle p \rangle$ is on the form $\langle q \rangle$. Then $q = ap$ for some $a \in K$. Since q is irreducible a is a unit hence $\langle p \rangle = \langle q \rangle$. □

2.8.6 Principal Ideal Domains

Definition 2.8.53. An ideal $I \subset R$ is *principal.* A domain in which every ideal is principal is called a *principal ideal domain* or a *PID.*

Lemma 2.8.54. *Let R be a PID. The non-trivial maximal ideals of R are those generated by primes.*

Proof. A maximal ideal is generated by some p that is non-zero and a non-unit. Since a maximal ideal is prime we have that p is prime.
Let a prime p be given. Suppose $\langle p \rangle \subset \langle x \rangle$. Then $p = qx$. Then $q \in R^*$ or $x \in R^*$, since p is in particular irreducible. Then $\langle p \rangle = \langle q \rangle$ or $\langle p \rangle = R$, hence $\langle p \rangle$ is maximal. ☐

Lemma 2.8.55. *Let R be a PID. Irreducible elements in R are prime.*

Proof. p be irreducible. Suppose $\langle p \rangle \subset \langle p' \rangle$. Then $p = qp'$ for some q, then $q \in R^*$ or $p' \in R^*$. In the first case $\langle p \rangle = \langle p' \rangle$ and in the second case $\langle p' \rangle = R$. Then $\langle p \rangle$ is maximal, hence p is prime. ☐

Lemma 2.8.56. *A PID is a UFD.*

Proof. Let R be a PID. If we can prove that any non-unit non-zero element in R decomposes into a product of irreducible elements, we are done by Proposition 2.8.46, having the prior lemma in mind. Let $a \in R \setminus 0$ be a non-unit. Since R is Noetherian we can find a maximal ideal $\langle p_1 \rangle \supset \langle a \rangle$. Note that then $\langle p_1 \rangle$ is prime, hence p_1 is prime and that $a = a_1 p_1$ for some a_1. Define a_{n+1} to be an element such that $a_n = a_{n+1} p_n$ for some p_n. We get an ascending chain

$$\langle a \rangle \subset \langle a_1 \rangle \subset \ldots$$

Since R is Noetherian, for some m, $\langle a_m \rangle = \langle a_n \rangle$ for $n \geq m$. Pick m to be the smallest such. If a_m was reducible, then $\langle a_m \rangle \subsetneq \langle a_{m+1} \rangle$. So it follows by induction that $a = a_m \prod_1^m p_i$; a product of irreducible elements.

☐

The following is an immediate result of the prior two lemmas

Corollary 2.8.57. *The non-zero maximal ideals of a PID are those generated by irreducible elements*

2.8.7 Local Rings, Localizations & Field of Fractions

Definition 2.8.58. A ring R is called *local* if it has unique maximal left ideal.

Proposition 2.8.59. *Let R be ring. R is local if and only if $\mathfrak{m} := R \setminus R^*$ is a left ideal.*

Proof. "$\Rightarrow$": Let I be the unique maximal left ideal of R. Then since I is proper, $I \subset \mathfrak{m}$. Note that for every $x \in \mathfrak{m}$, $\langle x \rangle$ is a proper ideal in R, hence $x \in I$. Therefor $\mathfrak{m} = I$, hence $\mathfrak{m}$ is an ideal.
"$\Leftarrow$": Let $I \subsetneq R$ be an ideal. Then every element of I is a non-unit, hence $I \subset \mathfrak{m}$, thus $\mathfrak{m}$ is the unique maximal left ideal in R. □

Definition 2.8.60. Let R be a commutative ring and $X \subset R$ a subset that is a submonoid of $(R, \cdot)$. For $(r,x),(r',x') \in R \times X$ we define a relation that $(r,x) \sim (r',x')$ if $rx' = r'x$. We define $X^{-1}R := R/\sim$ and denote an $(r,x) \in X^{-1}R$ by $\frac{r}{x}$. Hence

$$X^{-1}R = \left\{ \frac{r}{x} : r \in R, x \in X \right\}.$$

This is called the *localization of R with respect to X*. For an $x \in X$, if

$$X := \{x^n : n \geq 0\},$$

we define $R_x := X^{-1}R$. When $X = R \setminus \{0\}$, we define $Q(R) := X^{-1}R$. In this case $Q(R)$ is called the *field of fractions* of R.

Remark 2.8.61. We give some properties of this construction. Note that every $x, y \in X$

$$0y = 0x \Rightarrow \frac{0}{x} = \frac{0}{y}$$

and that for every $r, r' \in R$

$$r = r' \Rightarrow \frac{r}{1} = \frac{r'}{1},$$

thus we may regard R as a subset of $X^{-1}R$ via the map $r \mapsto \frac{r}{1}$. We also have that

$$xy = yx \Rightarrow \frac{x}{x} = \frac{y}{y}.$$

Furthermore,

$$(rx)x = rx^2 \Rightarrow \frac{rx}{x^2} = \frac{r}{x}$$

Lemma 2.8.62. *Let R be a commutative ring, $X \subset R$ a submonoid of R*

Lemma 2.8.63. *Let R be an integral domain and $X \subset R \setminus \{0\}$ a subset that is a submonoid of $(R, \cdot)$. For $\frac{r_1}{x_1}, \frac{r_2}{x_2} \in X^{-1}R$ we define*

$$\frac{r_1}{x_1} + \frac{r_2}{x_2} := \frac{r_1 x_2 + r_2 x_1}{x_1 x_2}$$

and

$$\frac{r_1}{x_1}\frac{r_2}{x_2} := \frac{r_1 r_2}{x_1 x_2}.$$

This makes $(X^{-1}R, +, \cdot)$ a commutative ring containing R as a subring, i.e. the image of the embedding of R in $X^{-1}R$ is a subring isomorphic to R.

Proof. We first need to check that the two operations are well-defined. Let $\frac{r_1}{x_1} = \frac{r'_1}{x'_1} \in X^{-1}R$ and $\frac{r_2}{x_2} = \frac{r'_2}{x'_2} \in X^{-1}R$. Then $r_1 x'_1 = r'_1 x_1$ and $r_2 x'_2 = r'_2 x_2$, which means

$$(r_1 x_2 + r_2 x_1) x'_1 x'_2 = r_1 x'_1 x_2 x'_2 + r_2 x'_2 x_1 x'_1 = r'_1 x_1 x_2 x'_2 + r'_2 x_2 x_1 x'_1 = \left(r'_1 x'_2 + r'_2 x'_1\right) x_1 x_2,$$

implying

$$\frac{r_1}{x_1} + \frac{r_2}{x_2} = \frac{r_1 x_2 + r_2 x_1}{x_1 x_2} = \frac{r'_1 x'_2 + r'_2 x'_1}{x'_1 x'_2} = \frac{r'_1}{x'_1} + \frac{r'_2}{x'_2},$$

hence addition is well-defined. In the same vein

$$r_1 r_2 x'_1 x'_2 = r_1 x'_2 r_2 x'_1 = r'_1 x_2 r'_2 x_1 = r'_1 r'_2 x_1 x_2,$$

implies

$$\frac{r_1}{x_1}\frac{r_2}{x_2} = \frac{r_1 r_2}{x_1 x_2} = \frac{r'_1 r'_2}{x'_1 x'_2} = \frac{r'_1}{x'_1}\frac{r'_2}{x'_2}.$$

We proceed to check the ring axioms. Let, in addition, $\frac{r_3}{x_3} \in X^{-1}R$ be given. Then

$$\begin{aligned}\frac{r_1}{x_1} + \left(\frac{r_2}{x_2} + \frac{r_3}{x_3}\right) &= \frac{r_1}{x_1} + \frac{r_2 x_3 + r_3 x_2}{x_2 x_3} = \frac{r_1 x_2 x_3 + r_2 x_3 x_1 + r_3 x_2 x_1}{x_1 x_2 x_3} = \frac{(r_1 x_2 + r_2 x_1) x_3 + r_3 x_2 x_1}{x_1 x_2 x_3} \\ &= \frac{r_1 x_2 + r_2 x_1}{x_1 x_2} + \frac{r_3}{x_3} = \left(\frac{r_1}{x_1} + \frac{r_2}{x_2}\right) + \frac{r_3}{x_3}.\end{aligned}$$

We define $0 := \frac{0}{1}$, with which we get

$$0 + \frac{r_1}{x_1} = \frac{0 x_1 + r_1 \cdot 1}{1 x_1} = \frac{r_1}{x_1}.$$

One should note that for any $x \in X$ $x \cdot 0 = 1 \cdot 0$, hence

$$\frac{0}{1} = \frac{0}{x}.$$

We define $-\frac{r_1}{x_1} := \frac{-r_1}{x_1}$ with which we get

$$\frac{r_1}{x_1} - \frac{r_1}{x_1} = \frac{r_1 x_1 - r_1 x_1}{x_1 x_1} = \frac{0}{x_1 x_1} = 0.$$

Lastly

$$\frac{r_1}{x_1}+\frac{r_2}{x_2}=\frac{r_1x_2+r_2x_1}{x_1x_2}=\frac{r_2x_1+r_1x_2}{x_2x_1}=\frac{r_2}{x_2}+\frac{r_1}{x_1},$$

hence $(X^{-1}R,+)$ is an additive group. We also have that

$$\frac{r_1}{x_1}\left(\frac{r_2}{x_2}\frac{r_3}{x_3}\right)=\frac{r_1}{x_1}\frac{r_2r_3}{x_2x_3}=\frac{r_1(r_2r_3)}{x_1(x_2x_3)}=\frac{(r_1r_2)r_3}{(x_1x_2)x_3}=\left(\frac{r_1r_2}{x_1x_2}\right)\frac{r_3}{x_3}=\left(\frac{r_1}{x_1}\frac{r_2}{x_2}\right)\frac{r_3}{x_3}.$$

We define $\mathbf{1}:=\frac{1}{1}$. Then

$$1\frac{r_1}{x_1}=\frac{1r_1}{1x_1}=\frac{r_1\cdot 1}{x_1\cdot 1}=\frac{r_1}{x_1}.$$

Furthermore,

$$\begin{aligned}\frac{r_1}{x_1}\left(\frac{r_2}{x_2}+\frac{r_3}{x_3}\right)&=\frac{r_1}{x_1}\frac{r_2x_3+r_3x_2}{x_2x_3}=\frac{r_1r_2x_3+r_1r_3x_2}{x_1x_2x_3}=\frac{r_1r_2x_3x_1+r_1r_3x_2x_1}{x_1x_2x_1x_3}=\frac{r_1r_2}{x_1x_2}+\frac{r_1r_3}{x_1x_3}\\&=\frac{r_1}{x_1}\frac{r_2}{x_2}+\frac{r_1}{x_1}\frac{r_3}{x_3}.\end{aligned}$$

Thus $(X^{-1}R,+,\cdot)$ is a ring. We check that is commutative. Indeed

$$\frac{r_1}{x_1}\frac{r_2}{x_2}=\frac{r_1r_2}{x_1x_2}=\frac{r_2r_1}{x_2x_1}=\frac{r_2}{x_2}\frac{r_1}{x_1}.$$

Let $r,r'\in R$. Then

$$\begin{aligned}&r+r'=\frac{r}{1}+\frac{r'}{1}=\frac{r+r'}{1}\in R,\\&-r=\frac{-r}{1}\in R\\&0=\frac{0}{1}\in R\\&rr'=\frac{r}{1}\frac{r'}{1}=\frac{rr'}{1}\in R\\&1=\frac{1}{1}\in R.\end{aligned}$$

These computations prove that $\operatorname{im} R\hookrightarrow X^{-1}R$ is a subring of $X^{-1}R$ (or that $R\hookrightarrow X^{-1}R$ is a ring homomorphism), hence $\operatorname{im} R\hookrightarrow X^{-1}R\simeq R$. □

Proposition 2.8.64. *Let R be a commutative ring, $\mathfrak{p}\subset R$ a prime ideal and $X:=R\setminus\mathfrak{p}$. Then X is a submonoid of $(R,\cdot)$ and $X^{-1}R$ is a local ring.*

Proof. Let □

Definition 2.8.65. Let R be an integral domain. Let X be a submonoid of $(R\setminus\{0\},\cdot)$. We define the *saturation* of X to be the set

$$\widehat{X}:=\left\{r\in R:\exists r'\in R,r'r\in X\right\}.$$

A submonoid $X\subset R\setminus\{0\}$ is *saturated*, if

$$\widehat{X}=X.$$

Remark 2.8.66. Let $x \in X$, then $1x \in X$, hence $x \in \widehat{X}$. We thus have $X \subset \widehat{X}$. Let $r,s \in \widehat{X}$, then for some $r',s' \in R$, $r'r \in X$ and $s's \in X$. Then $r's'rs \in X$, hence $rs \in \widehat{X}$. Clearly $1 \in \widehat{X}$. Thus $\widehat{X}$ is a submonoid of $R \setminus \{0\}$. The saturation of X is clearly saturated. Let Y be a saturated submonoid of $R \setminus \{0\}$ containing X. Let $r \in \widehat{X}$. Then for some $r' \in R$, $r'r \in X \subset Y$. Thus $\widehat{X} \subset Y$, hence $\widehat{X}$ is the smallest saturated submonoid of $R \setminus \{0\}$ containing X.

Lemma 2.8.67. *Let R be an integral domain and $X \subset R \setminus \{0\}$ a subset that is a submonoid of $(R \setminus \{0\}, \cdot)$. Then X is saturated if and only if for every $x,y \in R$ s.t. $xy \in X$, $x,y \in X$*

Proof. "$\Rightarrow$": Suppose X is saturated. Let $x,y \in R$ s.t. $xy \in X$. Then $y \in \widehat{X} = X$ and since $yx = xy \in X$, $x \in \widehat{X} = X$.
"$\Leftarrow$": Let $r \in \widehat{X}$, then for some $r' \in R$, $r'r \in X$, which by assumption means $r \in X$. $\square$

Lemma 2.8.68. *Let R be an integral domain and $X \subset R \setminus \{0\}$ a subset that is a submonoid of $(R \setminus \{0\}, \cdot)$. Consider the map*

$$\iota : R \hookrightarrow X^{-1}R$$
$$r \mapsto \frac{r}{1}$$

Let $\frac{r}{x} \in X^{-1}R$. Then

$$\frac{r}{x} \in \left(X^{-1}R\right)^* \iff r \in Y := \iota^{-1}\left(\left(X^{-1}R\right)^*\right).$$

Furthermore, $\widehat{X} = Y$. Thus

$$\left(X^{-1}R\right)^* = \left\{\frac{r}{x} \in X^{-1}R : r \in \widehat{X}\right\}.$$

Proof. "$\Rightarrow$": Suppose $\frac{r}{x} \in \left(X^{-1}R\right)^*$. Then for some $\frac{s}{y} \in X^{-1}R$, $\frac{r}{x}\frac{s}{y} = \frac{s}{y}\frac{r}{x} = 1$. From this we get that

$$r\frac{s}{xy} = \frac{rx}{x}\frac{s}{xy} = \frac{r}{x}x\frac{1}{x}\frac{s}{y} = 1,$$

hence $r \in Y$.
"$\Leftarrow$": If $r \in Y$, then $r\frac{s}{y} = 1$ for some $\frac{s}{y} \in X^{-1}R$, hence $\frac{r}{x}\frac{sx}{y} = 1$, implying $\frac{r}{x} \in \left(X^{-1}R\right)^*$.
If $r \in \widehat{X}$. Then for some $r' \in R$, $r'r \in X$. Then

$$r\frac{r'}{r'r} = \frac{r'r}{r'r} = 1 \Rightarrow r \in Y.$$

Let $r \in Y$. Then for some $\frac{s}{y} \in X^{-1}S$, $r\frac{s}{y} = 1$, meaning

$$sr = sr\frac{1}{y}y = r\frac{s}{y}y = y \in X \Rightarrow r \in \widehat{X}.$$

$\square$

Proposition 2.8.69. *Let R be an integral domain. Then $Q(R)$ is the smallest field containing R as a subring.*

Proof. The monoid $(R\setminus\{0\},\cdot)$ is obviously saturated. Hence, by the above lemma,

$$Q(R)^* = \left\{\frac{r}{s} \in Q(R) : r \in R\setminus\{0\}\right\} = Q(R)\setminus\{0\}.$$

This means $Q(R)$ is a field. Let K be a field containing R as a subring. Let $\frac{r}{s} \in Q(R)$. Then $r \in K$ and $\frac{1}{s} = s^{-1} \in K$, hence $\frac{r}{s} = r\frac{1}{s} \in K$. This means $Q(R) \subset K$, hence $Q(R)$ is the smallest field containing R as a subring. $\square$

Remark 2.8.70. From the above we conclude that if K is a field then $K = Q(K)$, and in general $Q(R) = Q(Q(R))$.

Definition 2.8.71. We define the *rational numbers* to be the field $\mathbb{Q} := Q(\mathbb{Z})$.

Lemma 2.8.72. *Let R and S be integral domains. Let $X \subset R\setminus 0$ be a submonoid of $(R,\cdot)$. Let $\sigma : R \to S$ such that $\sigma(X) \subset S\setminus 0$. Then*

$$\begin{aligned}\overline{\sigma} : X^{-1}R &\to \sigma(X)^{-1}S\\ \frac{a}{b} &\mapsto \frac{\sigma(a)}{\sigma(b)}\end{aligned}$$

is unique well-defined ring homomorphism such that $\overline{\sigma}|_R = \sigma$

Proof. By assumption $\sigma(X)$ is a submonoid of $(S,\cdot)$ not containing 0. Let $\frac{a}{b} = \frac{c}{d} \in X^{-1}R$. Then $ad = bc$, hence

$$\sigma(a)\sigma(d) = \sigma(ad) = \sigma(bc) = \sigma(b)\sigma(c) \Rightarrow \overline{\sigma}\left(\frac{a}{b}\right) = \frac{\sigma(a)}{\sigma(b)} = \frac{\sigma(c)}{\sigma(d)} = \overline{\sigma}\left(\frac{c}{d}\right).$$

Let $\frac{a}{b}, \frac{c}{d} \in X^{-1}R$ be arbitrary. Then

$$\overline{\sigma}\left(\frac{a}{b} + \frac{c}{d}\right) = \frac{\sigma(ad+bc)}{\sigma(bd)} = \frac{\sigma(a)\sigma(d)+\sigma(b)\sigma(c)}{\sigma(b)\sigma(d)} = \frac{\sigma(a)}{\sigma(b)} + \frac{\sigma(c)}{\sigma(d)} = \overline{\sigma}\left(\frac{a}{b}\right) + \overline{\sigma}\left(\frac{c}{d}\right),$$

and

$$\overline{\sigma}\left(\frac{a}{b}\frac{c}{d}\right) = \frac{\sigma(ac)}{\sigma(bd)} = \frac{\sigma(a)\sigma(c)}{\sigma(b)\sigma(d)} = \frac{\sigma(a)}{\sigma(b)}\frac{\sigma(c)}{\sigma(d)} = \overline{\sigma}\left(\frac{a}{b}\right)\overline{\sigma}\left(\frac{c}{d}\right).$$

Lastly, let $r \in R$. Then

$$\overline{\sigma}(r) = \overline{\sigma}\left(\frac{r}{1}\right) = \frac{\sigma(r)}{\sigma(1)} = \frac{\sigma(r)}{1} = \sigma(r),$$

hence in particular $\overline{\sigma}(1) = \sigma(1) = 1$. Let $\sigma' : X^{-1}R \mapsto \sigma(X)^{-1}S$ be another homomorphism with the property that $\sigma'|_R = \sigma$. Let $a, b \in R$ with $b \neq 0$. Then

$$\sigma'\left(\frac{1}{b}\right) = \sigma'\left(\frac{b}{1}\right)^{-1} = \sigma'(b)^{-1} = \frac{1}{\sigma(b)}.$$

One then sees that

$$\sigma'\left(\frac{a}{b}\right)=\sigma'(a)\sigma'\left(\frac{1}{b}\right)=\sigma(a)\frac{1}{\sigma(b)}=\frac{\sigma(a)}{\sigma(b)}=\overline{\sigma}\left(\frac{a}{b}\right)\Rightarrow \sigma'=\sigma.$$

□

Lemma 2.8.73. *The collection of pairs (R,X) where R is an integral domain $X\subset R\setminus 0$ a multiplicative submonoid of R with morphisms being ring homomorphisms $\sigma:(R,X)\to(S,Y)$ with $\sigma(X)\subset Y\subset R\setminus 0$, X defines a category.*

Proof. If $\sigma\in\mathrm{Hom}((R,X),(S,Y))$ and $\tau\in\mathrm{Hom}((S,Y),(T,Z))$, then $\sigma(X)\subset Y$ hence

$$\tau\circ\sigma(X)=\tau(\sigma(X))\subset\tau(Y)\subset Z\subset T\setminus 0.$$

Clearly $\mathbb{1}_{(R,X)}:=\mathrm{id}_R\in\mathrm{Hom}((R,X),(R,X))$. □

Proposition 2.8.74. *Call the category described in the above lemma $\mathscr{C}$. The assignment of a pair $((R,X),\sigma)$ in $(Ob(\mathscr{C}),\mathrm{Hom}(\mathscr{C}))$ to $(X^{-1}R,\overline{\sigma})$ in the category of integral domains defines a covariant functor.*

Proof. Let $\tau\in\mathrm{Hom}((S,Y),(T,Z))$, $\sigma\in\mathrm{Hom}((R,X),(S,Y))$ and $\frac{a}{b}\in X^{-1}R$. Then

$$\overline{\tau\circ\sigma}\left(\frac{a}{b}\right)=\frac{(\tau\circ\sigma)(a)}{(\tau\circ\sigma)(b)}=\frac{\tau(\sigma(a))}{\tau(\sigma(b))}=\overline{\tau}\left(\frac{\sigma(a)}{\sigma(b)}\right)=\overline{\tau}\left(\overline{\sigma}\left(\frac{a}{b}\right)\right)=(\overline{\tau}\circ\overline{\sigma})\left(\frac{a}{b}\right),$$

hence $\overline{\tau\circ\sigma}=\overline{\tau}\circ\overline{\sigma}$. Lastly

$$\overline{\mathbb{1}_{(R,X)}}\left(\frac{a}{b}\right)=\frac{a}{b}=\mathrm{id}_{X^{-1}R}\left(\frac{a}{b}\right)=\mathbb{1}_{X^{-1}R}\left(\frac{a}{b}\right).$$

□

Corollary 2.8.75. *Let R and S be integral domains and $X\subset R\setminus 0$ a submonoid. Then $R\overset{\sigma}{\simeq}S\Rightarrow X^{-1}R\simeq\sigma(X)^{-1}R$.*

Definition 2.8.76. Let R be an integral domain and $X\subset R\setminus 0$ a multiplicative submonoid of R. Consider an ideal $I\subset R$. We define the *localization ideal of I in $X^{-1}R$*. To be the set

$$X^{-1}I:=\left\{\frac{a}{x}\in X^{-1}R: a\in I, x\in X\right\}$$

Lemma 2.8.77. *Let R be an integral domain and $X\subset R\setminus 0$ a multiplicative submonoid of R. Let $I\subset R$ be an ideal. Then $(X^{-1}R)I=X^{-1}I$. Consequentially, $X^{-1}I$ is an ideal.*

Proof. Let $\sum_1^n \frac{r_i}{x_i} a_i \in (X^{-1}R)I$ where $\frac{r_i}{x_i} \in X^{-1}R$ and $a_i \in I$. Then

$$\sum_1^n \frac{r_i}{x_i} a_i = \frac{\sum_{i=1}^n \left(\prod_{j\in\{1,\dots,n\},j\neq i} x_j\right) r_i a_i}{\prod_{j=1}^n x_i} \in X^{-1}I.$$

The converse inclusion is trivial. □

Lemma 2.8.78. *Let R be an integral domain and $X \subset R \setminus 0$ a multiplicative submonoid of R. Let $I \subset X^{-1}R$ be an ideal. Then $X^{-1}(I \cap R) = I$.*

Proof. Let $\frac{a}{x} \in X^{-1}(I \cap R)$, where $a \in I$, $x \in X$.Then

$$\frac{a}{x} = \frac{1}{x} a \in I.$$

Conversely, let $\frac{a}{x} \in I$. Then $a = x\frac{a}{x} \in I$, hence $\frac{a}{x} \in X^{-1}(I \cap R)$. □

Lemma 2.8.79. *A local ring R with principal maximal ideal is Noetherian*

2.8.8 Discrete Valuation Rings

Definition 2.8.80. Let R be a non-field integral domain. R is a *discrete valuation ring (DVR)* if it is noetherian and local with the maximal ideal is principal.

Proposition 2.8.81. *Let R be a non-field integral domain. Then R is a DVR if and only if there is an irreducible element $t \in R$ such that for every $z \in R \setminus 0$ there are unique $u \in R^*$ and $n \geq 1$ satisfying $z = ut^n$*

Proof. "$\Rightarrow$": Let $\mathfrak{m} = \langle t \rangle$ be the maximal ideal of R. Then t is prime hence irreducible by the maximality. Let $z \in R \setminus 0$. Then either z is a unit in which case $z \notin \mathfrak{m}$ or z is not a unit hence $z \in \mathfrak{m}$. There is a $u \in R \setminus 0$ with $t \nmid u$ and a maximal n such that $z = ut^n$. Since $u \notin \langle t \rangle$ it is a unit by maximality. Suppose $u' \in R^*$ and $n' \geq 0$ are given such that $ut^n = u't^{n'}$. Then $n = n'$, since otherwise $t \mid u'$, hence $u = u'$.
"$\Leftarrow$": Let $\mathfrak{m} = \langle t \rangle$. Every non-unit is of the form ut^n, where $n \geq 1$, hence $R \setminus R^* = \mathfrak{m}$, hence R is local by Proposition 2.8.59. Let $I \subsetneq R$ be an ideal, then $I \subset \mathfrak{m}$. Then $I = \langle t^r \rangle$, where $r = \min\{n \geq 0 : t^n \in I\rangle$. Indeed if $a \in I$, then $a = ut^n$ for some $n \geq r$, hence $a = ut^{n-r}t^r \in I$. Then R is a PID, and hence Noetherian. □

Remark 2.8.82. We refer to an element such as t as an *uniformizing parameter*.The uniformizing parameters of a DVR are of the form ut where u is a unit and t is a uniformizing parameter. Set $K = Q(R)$. Then every $z \in K \setminus 0$ can be written uniquely on the form $z = ut^n$ for a unit $u \in R$ and an integer n. The integer n is called *the* *order of* z denoted $\mathbf{ord}(z)$. We set $\mathbf{ord}(0) = \infty$. One sees that $R = \mathbf{ord}^{-1}(\mathbb{Z}_{\geq 0} \cup \{\infty\})$ and $\mathfrak{m} = \mathbf{ord}^{-1}(\mathbb{Z}_{\geq 1} \cup \{\infty\})$. The order is independent of uniformizing parameter. Since if t is a uniformizing parameter and $z = ut^n$ for unique $u \in R^*$, $n \in \mathbb{Z}$, then for a unit $s \in R$, $z = \frac{a}{b} = \frac{vst^l}{yst^k} = \frac{v}{y}t^{l-k}$ for suitable units v, y, $l, k \geq 0$, hence by uniqueness $\frac{v}{y} = u$ and $l - k = n$.

Proposition 2.8.83. *The localization of* $\mathbb{Z}$ *with respect to a maximal ideal* $\langle p \rangle$ *(*$p \in \mathbb{Z}$ *is prime),* $\mathbb{Z}_{\langle p \rangle}$ *is a DVR whose quotient field is* $\mathbb{Q}$*.*

Proof. One notes that $\mathbb{Z}_{\langle p \rangle} = \{\frac{a}{n} \in \mathbb{Q} : p \nmid n\}$. Suppose $p = \frac{a}{n}\frac{b}{m}$. Then $mn \mid ab$, WLOG $mn = 1$, hence $m = 1$ and $n = 1$. Then $p = ab$ hence a or b is a unit, meaning p is irreducible in $\mathbb{Z}_{\langle p \rangle}$. For $\frac{a}{n} \in \mathbb{Z}_{\langle p \rangle}$ let $\nu_p\left(\frac{a}{n}\right) = \max(\{n \geq 0 : p^n \mid \frac{a}{n}\})$. One eaily checks that $p \mid \frac{a}{n}$ if and only if $p \mid a$, hence $\nu_p(\frac{a}{n}) = \nu_p(a)$. We thus get that

$$\frac{a}{n} = \frac{q}{n} p^{\nu_p(a)},$$

where $p \nmid q$. Note that $\frac{n}{q}$ is the inverse of $\frac{q}{n}$. The uniqueness of this decomposition follows from p not being being a unit. Proposition 2.8.81 shows that $\mathbb{Z}_{\langle p \rangle}$ is a DVR. Every element in $\mathbb{Q}$ can be written as $\frac{a}{b} = \frac{s}{t}p^{\nu_p(a) - \nu_p(b)} = \frac{s}{t}p^{\mathrm{ord}(\frac{a}{b})}$, where $p \nmid s, t$, hence $\mathbb{Q} \subset Q(\mathbb{Z}_{\langle p \rangle}) \subset \mathbb{Q}$. □

Lemma 2.8.84. *Let* R *be a DVR. Set* $K = Q(R)$*. Let* $\mathfrak{m} = \langle t \rangle$ *be the maximal ideal in* R*. If* $z = \frac{a}{b}t^{\mathrm{ord}(z)} \in K \setminus R$*, then* $z^{-1} \in \mathfrak{m}$*.*

Proof. Note that

$$K \setminus R = \left\{ u \frac{1}{t^n} : u \in R^*, n \geq 0 \right\},$$

hence $z = ut^{-n}$ for suitable $u \in R^*$, $n \geq 0$. Then $z^{-1} = u^{-1}t^n \in \mathfrak{m}$. □

Proposition 2.8.85. *Let* S *be a DVR containing a subring* R *which is also a DVR. Set* $K = Q(R)$ *and suppose* $S \subset K$*. Let* $\mathfrak{m} = \langle t \rangle$ *be the maximal ideal in* R*. If the maximal ideal of* S*,* $\mathfrak{n} = \langle s \rangle$*, contains* $\mathfrak{m}$*, then* $S = R$*.*

Proof. Since $t \in \mathfrak{n}$, $t = us^n$ for some $n \geq 1$. By irreducibility of t, $t = us$. Let $v \in S^*$. Then $v^{-1} \notin \mathfrak{n} \supset \mathfrak{m}$, hence $v \in R$ by the prior lemma. This means $R^* = S^*$. Thus if $x \in S$, $x = vs^n = ut^n$ for some $n \geq 0$, $v \in S^* = R^*$, hence $x \in R$. □

Definition 2.8.86. An *order function* on a field K is a function $\nu : K \to \mathbb{Z} \cup \{\infty\}$, satisfying:

1. for every $a \in K$, $\nu(a) = \infty \iff a = 0$,

2. for every $a, b \in K$, $v(ab) = v(a) + v(b)$,

3. for every $a, b \in K$, $v(a+b) \geq \min(v(a), v(b))$.

Definition 2.8.87. Let K be a field with an order function v. We define the ring induced by v to be the set

$$R_v(K) := R_v := v^{-1}(\mathbb{Z}_{\geq 0} \cup \{\infty\}) = \{r \in K : v(r) \geq 0\}.$$

Define the ideal induced by v to be the set

$$\mathfrak{m}_v := v^{-1}(\mathbb{Z}_{\geq 1} \cup \{\infty\}) = \{r \in R_v : v(r) > 0\} \subset R_v$$

Lemma 2.8.88. *Let K be a field with an order function v. We collect the following facts about R_v and $\mathfrak{m}_v$:*

(i) R_v is subring of K and hence is an integral domain.

(ii) For every $u \in R_v$,

$$u \in R_v^* \iff v(u) = 0.$$

(iii) $\mathfrak{m}_v$ is the unique maximal ideal of R_v, hence R_v is local.

(iv) $R_v = K$ if and only if v is trivial. If v is non-trivial, R_v is not a field.

Proof. (i) Property 2. ensures that R_v is closed under multiplication while property 3. ensures that R_v is closed under addition. $0 \in R_v$ by property 1. Note that $v(1) = v(1 \cdot 1) = v(1) + v(1)$, hence $v(1) = 0$, hence $1 \in R_v$. Let $u \in R_v^*$. Then $0 = v(1) = v(uu^{-1}) = v(u) + v(u^{-1})$, hence $v(u^{-1}) = -v(u)$ and since $v(u) \geq 0$, it follows that $v(u^{-1}) = 0$. We thus in particular find that $v(-1) = 0$, hence for any $r \in R_v$ we have that $v(-r) = v(-1 \cdot r) = v(-1) + v(r) = v(r)$. It thus follows that $-r \in R_v$. We thus get that R_v is a subring of K.
(ii) "$\Rightarrow$": This was already proven in the proof of (i).
"$\Leftarrow$": Let $u \in R_v$ such that $v(u) = 0$. For some $v \in K \setminus 0$, $uv = vu = 1$. Then $0 = v(1) = v(uv) = v(u) + v(v) = v(v)$, hence $v \in R_v$, which means u is a unit in R_v.
(iii) $\mathfrak{m}_v = R_v \setminus R_v^*$, hence it is sufficient to prove that $\mathfrak{m}_v$ is an ideal by Proposition 2.8.59. $\mathfrak{m}_v$ is closed under addition by property 3. Let $r \in R_v$, $x \in \mathfrak{m}_v$. Then $v(rx) = v(r) + v(x) \geq 0 + 1 = 1$, hence $rx \in \mathfrak{m}_v$. (iv) "$\Rightarrow$": Suppose v is not trivial. Then there is some $x \in K \setminus 0$ such that $v(x) \geq 1$, then $v(x^{-1}) \leq 1$, hence $x^{-1} \notin R_v$.
"$\Leftarrow$": Suppose v is trivial. Then for $x \in K$

$$v(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ \infty & \text{otherwise} \end{cases},$$

in any case $\nu(x) \geq 0$, hence $x \in R_\nu$.

$\square$

Theorem 2.8.89. *Let K be a field and ν a non-trivial order function on K. Then $R = \{z \in K : \nu(z) \geq 0\}$ is a DVR with maximal ideal $\mathfrak{m} := \{z \in R : \nu(z) > 0\}$ such that $Q(R_\nu) = K$. Conversely, if R is a DVR with quotient field K, then $\operatorname{ord} : K \to \mathbb{Z} \cup \{\infty\}$ is an order function. We thus obtain a one-to-one correspondence between DVR's and order functions.*

Proof. By the above lemma it is sufficient to check that R_ν is a PID. Let $I \subset R_\nu$ be a non-zero ideal. Let $s \in I$ be given such that $\nu(s) = \min(\nu(I))$. Let $x \in I$. For some $q \in K$, $x = qs$. Since

$$\nu(x) = \nu(q) + \nu(s) \Rightarrow \nu(q) = \nu(x) - \nu(s) \geq 0,$$

it follows that $s \mid x$ in R_ν, hence $I = \langle s \rangle$. Let $k \in K$. Then either $k \in R_\nu$ or $k^{-1} \in R_\nu$. In the first trivially $k \in Q(R_\nu)$. In the second case $k = (k^{-1})^{-1} = \frac{1}{k^{-1}} \in Q(R_\nu)$.
For the second statement, we have for $a \in R$ that $\operatorname{ord}(a) = \infty$ if and only if $a = 0$. Let $a, b \in K$. Then $a = ut^m$ and $b = vt^l$ for unique $u, v \in R^*$ and $l := \operatorname{ord}(b)$, $m := \operatorname{ord}(a)$.First, we get that

$$ab = uvt^{m+l} \Rightarrow \operatorname{ord}(ab) = m + l = \operatorname{ord}(a) + \operatorname{ord}(b),$$

hence ord satisfies property 2. Secondly,

$$a + b = \underbrace{\left(ut^{m-\min(m,l)} + vt^{l-\min(m,l)}\right)}_{q} t^{\min(m,l)}.$$

Note that $q \in R$ since $m - \min(m,l), l - \min(m,l) \geq 0$. We then have that $q = \alpha t^d$ for a unique unit $\alpha \in R^*$ and $d := \operatorname{ord}(q) \geq 0$. This implies that $a + b = \alpha t^{d+\min(l,m)}$, hence

$$\operatorname{ord}(a + b) = d + \min(m,l) \geq \min(m,l) = \min(\operatorname{ord}(a), \operatorname{ord}(b)),$$

proving property 3. $\square$

Remark 2.8.90. One notes from the above proof that the uniformizing parameter for R_ν is $t \in R_\nu$ where $\nu(t) = \min(\nu(\mathfrak{m}_\nu))$.

Lemma 2.8.91. *Let R be a DVR with $K := Q(R)$. If $a_1, \ldots, a_n \in K$ where for some i $\operatorname{ord}(a_i) < \operatorname{ord}(a_j)$ for every $j \neq i$, then $\operatorname{ord}(\sum_1^n a_j) = a_i$ and $\sum_1^n a_j \neq 0$*

Proof. We prove the first statement using induction in $n \geq 2$. Consider first the case $n = 2$. WLOG $m_1 := \mathrm{ord}(a_1) < \mathrm{ord}(a_2) := m_2$. Write $a_1 = u_1 t^{m_1}$ and $a_2 = u_2 t^{m_2}$ for suitable $u_1, u_2 \in R^*$. Then $a_1 + a_2 = (u_1 + u_2 t^{m_2 - m_1}) t^{m_1}$ and since $t \nmid u_1$ and $t \mid u_2 t^{m_2 - m_1}$, it follows that $t \nmid u := u_1 + u_2 t^{m_2 - m_1}$, hence u is a unit in R, meaning $\mathrm{ord}(a_1) = \mathrm{ord}(a_1 + a_2)$.
WLOG $i = n+1$. Note that $\mathrm{ord}(a_n + a_{n+1}) = \mathrm{ord}(a_{n+1})$, hence setting $a'_n = a_n + a_{n+1}$, we get $\mathrm{ord}(a'_n) = \mathrm{ord}(a_{n+1}) < a_j$ for every $j \in \{1, \ldots, n-1\}$. By induction it follows that

$$\mathrm{ord}(\sum_1^{n+1} a_j) = \mathrm{ord}(\sum_1^{n-1} a_j + (a_n + a_{n+1})) = \mathrm{ord}(\sum_1^{n-1} a_i + a'_n) = \mathrm{ord}(a'_n) = \mathrm{ord}(a_{n+1}).$$

Since $\mathrm{ord}(\sum_1^n a_j) < \mathrm{ord}(a_j) \leq \infty$, it follows that $\sum_1^n a_j \neq 0$ by property 1. of order functions. □

Lemma 2.8.92. *Let R be a DVR with maximal ideal $\mathfrak{m}$, $L := Q(R)$ and a subring K that is a field. Suppose that the composition $\sigma : K \hookrightarrow R \twoheadrightarrow L$ is an isomorphism. Then for any $z \in R$ there is a unique $\lambda \in K$ such that $z - \lambda \in \mathfrak{m}$.*

Proof. **The case $z \in \mathfrak{m}$:** Pick $\lambda = 0$. By the prior lemma for any $\mu \in K \setminus 0$ $\mathrm{ord}(z - \mu) = \mathrm{ord}(\mu) = 0$, hence $z - \mu \notin \mathfrak{m}$.
The case $z \notin \mathfrak{m}$: Then z is a unit in R. By a result in the "A First Look a Algebras"-subsubsection σ is a K-algebra isomorphism. For some $\lambda \setminus 0$, $\lambda = \sigma(\lambda) = z$, hence $z - \lambda = 0 \in \mathfrak{m}$. Since σ is injective it follows that λ is unique. □

Proposition 2.8.93. *We assume the same setup as the prior lemma. Let t be a uniformizing parameter of R and $z \in R$. For any $n \geq 0$ there are unique $\lambda_0, \ldots, \lambda_n \in K$, $z_n \in R$ such that*

$$z = \sum_0^n \lambda_i t^i + z_n t^{n+1}.$$

Proof. **Existence:** For the case $n = 0$ the statement follows from Lemma 2.8.92. Assuming the statement is true for some $n \geq 0$, we can write

$$z = \sum_0^n \lambda_i t^i + z_n t^{n+1}.$$

If $z_n \in R^* = K \setminus 0$, pick $\lambda_{n+1} := z_n$ and $z_{n+1} := 0$. Otherwise write $z_n = u t^l$, $l \geq 1$ and pick $\lambda_{n+1} := 0$, $z_{n+1} := u t^{l-1}$.
Uniqueness: Suppose there are $\lambda_0, \ldots, \lambda_n, \mu_1, \ldots, \mu_n \in K$ and $z_n, w_n \in R$ such that

$$\sum_0^n \lambda_i t^i + z_n t^{n+1} = \sum_0^n \mu_i t^i + w_n t^{n+1},$$

Then $\sum_0^n(\lambda_i-\mu_i)t^i+(z_n-w_n)t^{n+1}=0$, hence $\mathbf{ord}((\lambda_i-\mu_i)t^i)=\mathbf{ord}((\lambda_j-\mu_j)t^j)$ for every i,j hence $(\lambda_i-\mu_i)t^i=0$ for every i, implying $\lambda_i=\mu_i$. Similarly one can conclude that $z_n=w_n$. □

Lemma 2.8.94. *We keep the same setup as the above two results. Then* $\dim_K \mathfrak{m}^n/\mathfrak{m}^{n+1}=1$ *for every* $n\geq 0$.

Proof. We use induction in $n\geq 0$. We first prove that

$$\mathfrak{m}^n/\mathfrak{m}^{n+1}=\{\lambda t^n+\mathfrak{m}^{n+1}:\lambda\in K\}.$$

This is clear since any element in $\mathfrak{m}^n$, is equal to λt^n+zt^{n+1} for some $\lambda\in K$, $z\in R$ as by Lemma 2.8.91 $\mathbf{ord}(\sum_0^n\lambda_i t^i)=l$ where l is the smallest index such that $\lambda_l\neq 0$. Hence in the image of $\mathfrak{m}^n/\mathfrak{m}^{n+1}$ such an element is given by $\lambda t^n+zt^{n+1}+\mathfrak{m}^{n+1}=\lambda t^n+\mathfrak{m}^{n+1}$. It follows that

$$\begin{aligned}\sigma:K&\to\mathfrak{m}^n/\mathfrak{m}^{n+1}\\ \lambda&\mapsto\lambda t^n+\mathfrak{m}^{n+1}\end{aligned}$$

is a surjective K-algebra. Suppose $\lambda t^n\in\mathfrak{m}^{n+1}$. Then $\mathbf{ord}(\lambda)>0$, hence $\lambda=0$. We thus conclude that $\mathfrak{m}^n/\mathfrak{m}^{n+1}\simeq K$, meaning $\dim\ \mathfrak{m}^n/\mathfrak{m}^{n+1}=\dim\ K=1$. □

Lemma 2.8.95. *We keep the setup from the prior results. For each* $n\geq 0$, $\dim\ R/\mathfrak{m}^n=n$. *It follows that* $\mathbf{ord}(z)=\dim\ R/\langle z\rangle=\dim\ R/\mathfrak{m}^{\mathbf{ord}(z)}$ *for each* $z\in R$.

Proof. We prove the result by induction in n. The base case is trivial. By Lemma 2.6.25 and Lemma 2.7.4 and the induction hypothesis it follows that

$$\dim\ R/\mathfrak{m}^{n+1}=\dim\ \mathfrak{m}^n/\mathfrak{m}^{n+1}+\dim\ R/\mathfrak{m}^n=n+1,$$

where we also use the prior lemma. We now that $z=\lambda t^{\mathbf{ord}(z)}$, hence $\langle z\rangle=\langle t\rangle^{\mathbf{ord}(z)}=\mathfrak{m}^{\mathbf{ord}(z)}$, hence $\mathbf{ord}(z)=\dim\ R/\mathfrak{m}^{\mathbf{ord}(z)}$. □

2.9 Polynomial Rings & Formal Power Series

In this subsection every ring will be commutative, unless we explicitly declare it to not (necessarily) be the case. Really the base ring for a polynomial ring need not be commutative, but for our purposes we do not need to explore the non-commutative case. By a polynomial in n variables over a ring R, we mean some expression of the form

$$\sum_{v=(v_1,\ldots,v_n)\in\mathbb{N}^n}a_v x_1^{v_1}\cdots x_n^{v_n},$$

where $x_1,\ldots,x_n$ are *variables* and $a_v = 0$ for all but finitely many $a_v \in R$. Thus we want to consider elements of the algebra over R generated by $x_1,\ldots,x_n$, i.e. $R[x_1,\ldots,x_n]$. The term variable is informal, and our goal will be to make the term variable precise. There are some properties that we want these variables to have. For instance we do not want $x_i = x_j$ when $i \neq j$. In general, we want $x_1^{v_1}\cdots x_n^{v_n} \neq x_1^{w_1}\cdots x_n^{w_n}$ whenever $(v_1,\ldots,v_n) \neq (w_1,\ldots,w_n)$. To do this, we first introduce the notion of algebraic (in)dependence

Definition 2.9.1. Let S be an R-algebra. We say a finite sequence of elements $s_1,\ldots,s_n \in Z(S)$ are *algebraically independent* over R if for every finite sequence $(a_v) \in \prod_{v\in\mathbb{N}^n} R$, which is not the sequence $(0) \in \prod_{v\in\mathbb{N}^n} R$, we get that

$$\sum_{v\in\mathbb{N}^n} a_v s_1^{v_1}\cdots s_n^{v_n} \neq 0.$$

If a finite sequence of elements in S are not algebraically independent over R, we say that they are *algebraically dependent.*

One quickly sees that the concept over algebraic independence is really just a special case of linear independence.

Lemma 2.9.2. *Let S be an R-algebra. Then a finite sequence in $s_1,\ldots,s_n \in Z(S)$ is algebraically independent over R if and only if $\{s_1^{v_1}\cdots s_n^{v_n}\}_{(v_1,\ldots,v_n)\in\mathbb{N}^n}$ is linearly independent over R.*

We also want that elements of R can be seen as polynomials. Before proceeding with actually constructing a polynomial ring that does the job we will present the approach that at first might seem fruitful, but will not quite capture the behaviour we desire. That is to define $R[x_1,\ldots,x_n]$ as the set of functions

$$\mathrm{Pol}(R^n,R) := \left\{ f : R^n \to R : \begin{smallmatrix} f(x_1,\ldots,x_n)=\sum_{v\in\mathbb{N}^n} a_v x_1^{v_1}\cdots x_n^{v_n} \text{ for some} \\ \text{finite sequence } \{a_v\}_{v\in\mathbb{N}^n}\subset R \text{ for all } (x_1,\ldots,x_n)\in R^n \end{smallmatrix} \right\},$$

i.e. the set of polynomial functions, which is a subring of $\mathrm{Fun}(R^n,R)$, the set of functions from R^n to R. The main issue is that we can't always distinguish terms of form $x_1^{v_1}\cdots x_n^{v_n}$. For instance, if $\#R < \infty$, then clearly $\#\mathrm{Pol}(R^n,R) < \infty$, but we want the number of distinct terms $x_1^{v_1}\cdots x_n^{v_n}$ to be countably infinite. To be concrete, taking $R := \mathbb{Z}/2\mathbb{Z}$ then $x \mapsto x$ and $x \mapsto x^2$ is the same function, hence $x = x^2 \iff x^2 - x = 0$, thus $\mathrm{Pol}(R^n,R)$ fails to produce the right notion of variable in a lot of cases. Instead we will present an alternative approach.

2.9.1 Defining the Polynomial Ring

In this subsection we give a rigorous construction of the polynomial ring.

Definition 2.9.3. Consider the function

$$|\bullet| : \mathbb{N}^n \to \mathbb{N}$$
$$v = (v_1, \dots, v_n) \mapsto \sum_1^n v_i,$$

For an n-tuple $v \in \mathbb{N}^n$ we will refer to the quantity $|v|$ as the *modulus* of v.

Remark 2.9.4. One easily sees that a sequence $(a_v) \in \bigoplus_{v \in \mathbb{N}^n} R$ for some ring R is finite if and only if $a_v = 0$ whenever $|v| > N$ for some $N \geq 0$

One recalls that $(\mathbb{N}^n, +)$ is a commutative monoid for every $n \geq 1$, where for $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_n)$ in $\mathbb{N}^n$,

$$v + w := (v_1 + w_1, \dots, w_n + v_n).$$

We have the following result.

Lemma 2.9.5. *For every $n \geq 1$, the modulus function is additive.*

Proof. Let $v, w \in \mathbb{N}^n$ with $v = (v_1, \dots, v_n)$ and $w = (w_1, \dots, w_n)$. Then

$$|v + w| = |(v_1 + w_1, \dots, v_n + w_n)| = \sum_1^n (v_i + w_i) = \sum_1^n v_i + \sum_1^n w_i = |v| + |w|.$$

□

Definition 2.9.6. Let R be a ring and n a positive integer. By a *polynomial in n variables over R*, we mean an element (a_v) in the left R-module $\bigoplus_{v \in \mathbb{N}^n} R$. We denote the set of polynomials in n variables over R by $R[\mathbb{N}^n]$, i.e. $R[\mathbb{N}^n] := \bigoplus_{v \in \mathbb{N}^n} R$.

With this definition we already have that $R[\mathbb{N}^n]$ is a left R-module, since it is an R-submodule of $\prod_{v \in \mathbb{N}^n} R$. The set $\{e_v : v \in \mathbb{N}^n\}$ where $e_v = (\delta_{vw}) \in \prod_{w \in \mathbb{N}^n} R$ is a basis of $R[\mathbb{N}^n]$. We now aim to equip $R[\mathbb{N}^n]$ with a suitable multiplication. We do this by adding structure of ring on $\prod_{v \in \mathbb{N}^n} R$ and showing that $R[\mathbb{N}^n]$ is a subring. The set $\prod_{v \in \mathbb{N}^n} R$ with this structure of ring is called *the ring of formal power series in n variables over R*.

Lemma 2.9.7. *We define multiplication on $\prod_{v \in \mathbb{N}^n} R$ by*

$$(a_v)(b_v) = \left(\sum_{v, w \in \mathbb{N}^n : v + w = u} a_v b_v \right) \in \prod_{u \in \mathbb{N}^n} R.$$

This multiplication makes $\prod_{v \in \mathbb{N}^n} R$ a commutative ring. $R[\mathbb{N}^n]$ is a subring of $\prod_{v \in \mathbb{N}^n} R$.

Proof. Let $(a_v),(b_v),(c_v) \in \prod_{v\in\mathbb{N}^n} R$. Then

$$\begin{aligned}
((a_v)(b_v))(c_v) &= \left(\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w\right)(c_v) = \left(\sum_{r,u\in\mathbb{N}^n:r+u=s}\left(\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w\right)c_r\right)\\
&= \left(\sum_{r,v,w\in\mathbb{N}^n:r+v+w=s}(a_v b_w)c_r\right) = \left(\sum_{r,v,w\in\mathbb{N}^n:r+v+w=s} a_v(b_w c_r)\right)\\
&= \left(\sum_{u,v\in\mathbb{N}^n u+v=s} a_v\left(\sum_{r,w\in\mathbb{N}^n:r+w=u} b_w c_r\right)\right) = (a_v)\left(\sum_{r,w\in\mathbb{N}^n:r+w=u} b_w c_r\right)\\
&= (a_v)((b_v)(c_v)).
\end{aligned}$$

Put $\mathbf{0} := (0,\dots,0)\in\mathbb{N}^n$. We then define $1 := e_{\mathbf{0}} = (\delta_{\mathbf{0}v})$. Then

$$1(a_v) = \left(\sum_{v,w\in\mathbb{N}^n:v+w=u}\delta_{\mathbf{0}v}a_w\right) = \left(\sum_{w\in\mathbb{N}^n:w=u} a_w\right) = (a_v).$$

Similarly it is easy to check that $(a_v)1 = (a_v)$. Finally we have that

$$\begin{aligned}
(a_v)(((b_v)+(c_v)) &= \left(\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v(b_w+c_w)\right) = \left(\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w + a_v c_w\right)\\
&= \left(\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w + \sum_{v,w\in\mathbb{N}^n:v+w=u} a_v c_w\right)\\
&= \left(\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w\right) + \left(\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v c_w\right) = (a_v)(b_v)+(a_v)(c_v).
\end{aligned}$$

This means $\prod_{v\in\mathbb{N}^n} R$ is a ring with this multiplication. Note also that

$$(a_v)(b_v) = \left(\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w\right) = \left(\sum_{v,w\in\mathbb{N}^n:v+w=u} b_v a_w\right) = (b_v)(a_v).$$

Hence $\prod_{v\in\mathbb{N}^n} R$ is a commutative ring with this multiplication.
To check that $R[\mathbb{N}^n]$ is a subring of $\prod_{v\in\mathbb{N}^n} R$, we need to check that $1\in R[\mathbb{N}^n]$ and that $R[\mathbb{N}^n]$ is closed under multiplication. Since $\delta_{\mathbf{0},v} = 0$ for every $v\in\mathbb{N}^n\setminus\{\mathbf{0}\}$, it follows that $1\in R[\mathbb{N}^n]$. Let $(a_v),(b_v)\in R[\mathbb{N}^n]$. We note that for some $N,M\geq 0$, $a_v = 0$ for $v\in\mathbb{N}^n$ with $|v|\geq N$ and $b_w = 0$ for $w\in\mathbb{N}^n$ with $|w|\geq M$. Let $u\in\mathbb{N}^n$ with $|u|\geq N+M$. Consider then $v,w\in\mathbb{N}^n$ such that $v+w$. Then using LEMMA?

$$|v|+|w| = |v+w| = |u| \geq N+M \Rightarrow |v|\geq N \text{ or } |w|\geq M \Rightarrow a_v b_w = 0.$$

Thus $\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w = 0$, meaning $(a_v)(b_v)\in R[\mathbb{N}^n]$. ☐

Remark 2.9.8. As a notational trick one often denotes an $(a_v)\in\prod_{v\in\mathbb{N}^n}$ by $\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v$, where $\mathbf{x}^v$ is a short-hand notation for $x_1^{v_1}\cdots x_n^{v_n}$. With this choice of notation the elements of $\prod_{v\in\mathbb{N}^n} R$ are seen to act like some sort of power series in n variables with coefficients in R, where we of course "forget" the notion of convergence. The ring of formal power series in n variables is denoted $R[\![x_1,\dots,x_n]\!]$, but for now we will make no more remarks about this ring and focus on the polynomial ring. We will see that the notation $\sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v$ is actually on the nose in the sense of Remark 2.2.18 for the ring of polynomials.

Definition 2.9.9. Let R be a ring and n a positive integer. Consider a $w\in\mathbb{N}^n$. We define the *monomial* associated with w as the polynomial

$$e_v = (\delta_{vw}) \in R[\mathbb{N}^n].$$

For $i\in\{1,\dots,n\}$ we define *the i'th variable* in $R[\mathbb{N}^n]$ to be monomial associated with the n-tuple of non-negative integers for which the i'th entry is 1 and for which the remaining entries are 0.

Remark 2.9.10. A note on notation: We choose to denote the i'th variable by some letter, say x, subscripted by x_i, i.e. we denote the n variables by $x_1,\dots,x_n$. With this choice of letter we choose denote a monomial associated with a $v\in\mathbb{N}^n$ by $\mathbf{x}^v$. Later on this notation will be motivated. Had we chosen y as our letter we would get variables $y_1,\dots,y_n$ and monomials $\mathbf{y}^v$. This remark is not of a mathematical nature and serves only as an excuse to not explicitly state what is meant by a notation a'la $\mathbf{x}^v$ every time we make use of it.

Lemma 2.9.11. *Any element $f=(a_v)\in R[\mathbb{N}^n]$ can be written uniquely as*

$$\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v,$$

Proof. This is just a matter of book keeping. Since $\{\mathbf{x}^v : v\in\mathbb{N}^n\} = \{e_v : v\in\mathbb{N}^n\}$ is a basis of $R[\mathbb{N}^n] = \bigoplus_{v\in\mathbb{N}^n} R$, it follows that for some finite set $X\subset\mathbb{N}^n$, $a_v\neq 0$ for every $v\in\mathbb{N}^n$. Hence

$$f = \sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v.$$

The uniqueness of this representation follows from $\{\mathbf{x}^v : v\in\mathbb{N}^n\}$ being a basis. $\square$

Remark 2.9.12. A further consequence is that for $f=\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v, g=\sum_{v\in\mathbb{N}^n} b_v\mathbf{x}^v\in R[\mathbb{N}^n]$

$$fg = \sum_{u\in\mathbb{N}^n} c_u\mathbf{x}^u,$$

where $(c_u) = \left(\sum_{v,w\in\mathbb{N}^n : v+w=u} a_v b_w\right)$. Suppose $f=\mathbf{x}^\nu$ and $g=\mathbf{x}^\mu$. Then

$$fg = (c_u) = \left(\sum_{v,w\in\mathbb{N}^n : v+w=u} \delta_{\nu v}\delta_{\mu w}\right).$$

Note that

$$\delta_{\nu v}\delta_{\mu w} = 0 \iff \delta_{\nu v} = 0 \text{ or } \delta_{\mu w} = 0 \iff \nu = v \text{ or } \mu = w,$$

hence $\delta_{\nu v}\delta_{\mu w} = 1$ if $\nu = v$ and $\mu = w$ and 0 else. Thus $c_u = 1$ when $u = \nu + \mu$ and else it is 0. This means $f = (\delta_{v,\nu+\mu}) = \mathbf{x}^{\nu+\mu}$. It follows that $\mathbf{x}^{\mu} = x_1^{\mu_1} \cdots x_n^{\mu_n}$. Any polynomial can thus uniquely be represented as sum

$$\sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v = \sum_{v=(v_1,\ldots,v_n)\in\mathbb{N}^n} a_v x_1^{v_1} \cdots x_n^{v_n}.$$

One useful fact that one should note is that this means that for some $v_1, \ldots, v_m \in \mathbb{N}^n$, we have that the above is equal to

$$\sum_1^n a_{v_i} \mathbf{x}^{v_i}.$$

Another way of representing the above, which also may be useful is that for some N, the above is equal to

$$\sum_{v \in \mathbb{N}^n : |v| \leq N} a_v \mathbf{x}^v$$

We record the fact that R is embedded as a subring in a polynomials in the most natural way

Lemma 2.9.13. *$R1$ is a subring of $R[\![x_1,\ldots,x_n]\!]$ contained in $R[x_1,\ldots,x_n]$, ring isomorphic to R. Furthermore, $R1[x_1,\ldots,x_n] = R[x_1,\ldots,x_n]$. Lastly $x_1,\ldots,x_n$ are algebraically independent over R.*

Proof. We consider the map

$$\sigma : R \to R1$$
$$r \mapsto r1 = r$$

This is clearly a surjective ring homomorphism hence $R1$ is subring of $R[x_1,\ldots,x_n]$ whose inverse is

$$\sigma^{-1} : R1 \to R$$
$$r1 \mapsto r$$

We already know that $R1[x_1,\ldots,x_n]$ is a subring of $R[\mathbb{N}^n]$. Furthermore we have already seen in the above remark that any element in $R[x_1,\ldots,x_n]$ can be written as finite linear combination over R of elements in $\{x_1^{v_1} \cdots x_n^{v_n} : (v_1,\ldots,v_n) \in \mathbb{N}^n\}$. The fact that this set constitutes a basis of $R[x_1,\ldots,x_n]$ over R, means that $x_1,\ldots,x_n$ are algebraically independent over R (cf. Lemma 2.9.2). $\square$

Remark 2.9.14. We have now fully justified the existence of a polynomial ring with the properties described in the introduction to this subsection. In summary, we found that the set of finite sequences in R indexed by elements in $\mathbb{N}^n$ could be endowed with the structure we sought after. From now we we will "forget" the underlying structure.

We collect all the data established about the polynomial ring in the following theorem

Theorem 2.9.15. *The rings $R[\![x_1,\ldots,x_n]\!]$ and*

$$R[x_1,\ldots,x_n] = \left\{ \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v : a_v \in R, a_v = 0 \text{ whenever } |v| > N \text{ for some } N \geq 0 \right\},$$

are rings containing R as a subring. $R[x_1,\ldots,x_n]$ is generated by $x_1,\ldots,x_n$. Furthermore, $x_1,\ldots,x_n$ are algebraically independent over R.

2.9.2 Specializations of Polynomials

Proposition 2.9.16. *Let R and S be commutative rings and consider a ring homomorphism $\sigma : R \to S$. Then σ induces a well-defined ring homomorphism given by*

$$\begin{aligned} \overline{\sigma} : R[x_1,\ldots,x_n] &\to S[x_1,\ldots,x_n] \\ \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v &\mapsto \sum_{v\in\mathbb{N}^n} \sigma(a_v)\mathbf{x}^v \end{aligned}$$

Proof. Suppose $\sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v = \sum_{v\in\mathbb{N}^n} b_v \mathbf{x}^v \in R[\mathbf{x}]$ then $a_v = b_v$ for every $v \in \mathbb{N}^n$ hence $\sigma(a_v) = \sigma(a_v)$ for every $v \in \mathbb{N}^n$, meaning

$$\overline{\sigma}\left(\sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v\right) = \sum_{v\in\mathbb{N}^n} \sigma(a_v)\mathbf{x}^v = \sum_{v\in\mathbb{N}^n} \sigma(b_v)\mathbf{x}^v = \overline{\sigma}\left(\sum_{v\in\mathbb{N}^n} b_v \mathbf{x}^v\right),$$

we thus conclude that $\overline{\sigma}$ is well-defined.
Consider arbitrary $\sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v, \sum_{v\in\mathbb{N}^n} b_v \mathbf{x}^v \in R[\mathbf{x}]$. Then

$$\begin{aligned} \overline{\sigma}\left(\sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v + \sum_{v\in\mathbb{N}^n} b_v \mathbf{x}^v\right) &= \sum_{v\in\mathbb{N}^n} \sigma(a_v + b_v)\mathbf{x}^v = \sum_{v\in\mathbb{N}^n} (\sigma(a_v) + \sigma(b_v))\mathbf{x}^v \\ &= \sum_{v\in\mathbb{N}^n} \sigma(a_v)\mathbf{x}^v + \sum_{v\in\mathbb{N}^n} \sigma(b_v)\mathbf{x}^v = \overline{\sigma}\left(\sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v\right) + \overline{\sigma}\left(\sum_{v\in\mathbb{N}^n} b_v \mathbf{x}^v\right) \end{aligned}$$

hence $\overline{\sigma}$ is additive. It also multiplicative. Indeed,

$$\begin{aligned}\overline{\sigma}\left(\left[\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right]\left[\sum_{v\in\mathbb{N}^n} b_v\mathbf{x}^v\right]\right) &= \sum_{u\in\mathbb{N}^n}\left[\sum_{v,w\in\mathbb{N}^n:v+w=u}\sigma(a_vb_w)\right]\mathbf{x}^u\\ &= \sum_{u\in\mathbb{N}^n}\left[\sum_{v,w\in\mathbb{N}^n:v+w=u}\sigma(a_v)\sigma(b_w)\right]\mathbf{x}^u\\ &= \left[\sum_{v\in\mathbb{N}^n}\sigma(a_v)\mathbf{x}^v\right]\left[\sum_{v\in\mathbb{N}^n}\sigma(b_v)\mathbf{x}^v\right]\\ &= \overline{\sigma}\left(\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right)\overline{\sigma}\left(\sum_{v\in\mathbb{N}^n} b_v\mathbf{x}^v\right).\end{aligned}$$

Lastly, $\overline{\sigma}(1)=\sigma(1)=1$. □

Definition 2.9.17. For a ring extension $R\supset K$ where K is a field, given a ring map $\sigma:R\to K$, we call $\overline{\sigma}$ a *specialization of R in K.*

Lemma 2.9.18. *Let rings R,S,T be given and consider $\sigma\in\operatorname{Hom}(R,S)$, $\tau\in\operatorname{Hom}(S,T)$. Then $\overline{\tau\circ\sigma}=\overline{\tau}\circ\overline{\sigma}$. We also have that $\overline{id_R}=id_{R[x_1,\dots,x_n]}$. In other words $(R,\sigma)\mapsto(R[x_1,\dots,x_n],\overline{\sigma})$ is a covariant functor for every $n\geq 1$.*

Proof. Indeed, for $\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\in R[\mathbf{x}]$

$$\overline{\tau\circ\sigma}\left(\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right)=\sum_{v\in\mathbb{N}^n}\tau(\sigma(a_v))\mathbf{x}^v=\overline{\tau}\left(\sum_{v\in\mathbb{N}^n}\sigma(a_v)\mathbf{x}^v\right)=\overline{\tau}\circ\overline{\sigma}\left(\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right),$$

and lastly

$$\overline{\mathrm{id}_R}\left(\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right)=\sum_{v\in\mathbb{N}^n}\mathrm{id}_R(a_v)\mathbf{x}^v=\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v=\mathrm{id}_{R[\mathbf{x}]}\left(\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right).$$

□

Corollary 2.9.19. *Suppose $R\overset{\sigma}{\simeq}R$. Then $R[x_1,\dots,x_n]\overset{\overline{\sigma}}{\simeq}S[x_1,\dots,x_n]$.*

Proof. An immediate consequence of functoriality. □

Example 2.9.20. It is in general not true that if $R[x_1,\dots,x_n]\simeq S[x_1,\dots,x_n]$, then $R\simeq S$. Reference example.

2.9.3 Degree, Evaluation & Roots

Definition 2.9.21. Let $S\supset R$ be a ring extension. We define *evaluation* to be the map

$$\mathrm{ev}:S^n\times R[x_1,\dots,x_n]$$
$$((s_1,\dots,s_n),f)=\left((s_1,\dots,s_n),\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right)\mapsto\sum_{v\in\mathbb{N}^n} a_v s_1^{v_1}\cdots s_n^{v_n}$$

Let $s_1,\dots,s_n \in S$. We define *evaluation in* $(s_1,\dots,s_n) \in S^n$ as the map

$$\begin{aligned} \mathrm{ev}_{s_1,\dots,s_n} &: R[x_1,\dots,x_n] \to S \\ f = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v &\mapsto \mathrm{ev}((s_1,\dots,s_n), f) = \sum_{v \in \mathbb{N}^n} a_v s_1^{v_1} \cdots s_n^{v_n} \end{aligned}$$

For an $f \in R[x_1,\dots,x_n]$ we define $f(s_1,\dots,s_n) := \mathrm{ev}_{s_1,\dots,s_n}(f)$.

Lemma 2.9.22. *Evaluation is a well-defined map. Moreover, the map* $\mathrm{ev}_{s_1,\dots,s_n}$ *is a well-defined ring homomorphism such* $\mathrm{ev}_{s_1,\dots,s_n}(r) = r$ *for every* $r \in R$*, therefor it is also an* R*-module homomorphism. In other words, evaluation in an element is an* R*-algebra homomorphism.*

Proof. **The map is well-defined:** Let $(s_1,\dots,s_n) := (t_1,\dots,t_n) \in S^n$ and $f = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v$, $g = \sum_{v \in \mathbb{N}^n} b_v \mathbf{x}^v$ be polynomials in $R[x_1,\dots,x_n]$ such that $f = g$. Note that for a polynomial $h := \sum_{v \in \mathbb{N}^n} c_v \mathbf{x}^v$ that

$$\mathrm{ev}((s_1,\dots,s_n), h) = \sum_{v \in \mathbb{N}^n} c_v s_1^{v_1} \cdots s_n^{v_n} = \sum_{v \in \mathbb{N}^n} c_v t_1^{v_1} \cdots t_n^{v_n} = \mathrm{ev}((t_1,\dots,t_n), h).$$

By Theorem 2.9.15 it follows that $a_v = b_v$ for every $v \in \mathbb{N}^n$. Thus in particular

$$\mathrm{ev}(s_1,\dots,s_n), f) = \sum_{v \in \mathbb{N}^n} a_v s_1^{v_1} \cdots s_n^{v_n} = \sum_{v \in \mathbb{N}^n} b_v s_1^{v_1} \cdots s_n^{v_n} = \mathrm{ev}((s_1,\dots,s_n), g) = \mathrm{ev}((t_1,\dots,t_n), g),$$

which means the evaluation map is well defined.

evaluation is a ring homomorphism: Let $f = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v$ and $g = \sum_{v \in \mathbb{N}^n} b_v \mathbf{x}^v$ be polynomials in $R[x_1,\dots,x_n]$ and $s_1,\dots,s_n \in S$. The map is additive

$$\begin{aligned} \mathrm{ev}_{s_1,\dots,s_n}(f+g) &= \sum_{v \in \mathbb{N}^n} (a_v + b_v) s_1^{v_1} \cdots s_n^{v_n} = \sum_{v \in \mathbb{N}^n} a_v s_1^{v_1} \cdots s_n^{v_n} + \sum_{v \in \mathbb{N}^n} b_v s_1^{v_1} \cdots s_n^{v_n} \\ &= \mathrm{ev}_{s_1,\dots,s_n}(f) + \mathrm{ev}_{s_1,\dots,s_n}(g). \end{aligned}$$

The map is multiplicative:

$$\begin{aligned} \mathrm{ev}_{s_1,\dots,s_n}(fg) &= \sum_{v,w \in \mathbb{N}^n} a_v b_w s_1^{v_1+w_1} \cdots s_n^{v_n+w_n} = \sum_{v,w \in \mathbb{N}^n} a_v s_1^{v_1} \cdots s_n^{v_n} b_v s_1^{w_1} \cdots s_n^{w_n} \\ &= \left(\sum_{v \in \mathbb{N}^n} a_v s_1^{v_1} \cdots s_n^{v_n} \right) \left(\sum_{v \in \mathbb{N}^n} b_v s_1^{w_1} \cdots s_n^{w_n} \right) = \mathrm{ev}_{s_1,\dots,s_n}(f) \mathrm{ev}_{s_1,\dots,s_n}(g). \end{aligned}$$

Evaluation fixes R: Let $r \in R$. Then

$$\mathrm{ev}_{s_1,\dots,s_n}(r) = \mathrm{ev}_{s_1,\dots,s_n}\left(r\mathbf{x}^{(0,\dots,0)} \right) = r s_1^0 \cdots s_n^0 = r.$$

□

Remark 2.9.23. Given a commutative R-algebra S and elements $s_1, \ldots, s_n \in S$, we note that the R-algebra generated by these elements, i.e. $R[s_1, \ldots, s_n]$ is given as the image of $\mathrm{ev}_{s_1,\ldots,s_n} : R[x_1, \ldots, x_n] \to S$.

Lemma 2.9.24. *Let R be a ring and $J \subset R[y_1, \ldots, y_m]$ be an ideal. Consider $\overline{f_1} := f_1 + J, \ldots, \overline{f_n} := f_n + J \in R[\mathbf{y}]/J$. Then for each $f \in R[\mathbf{x}]$*

$$\mathrm{ev}_{\overline{f_1},\ldots,\overline{f_n}}(f) = \mathrm{ev}_{f_1,\ldots,f_n}(f) + J$$

Proof. This is a simple matter of using the definition addition and multiplication in the quotient ring. Indeed we can write $f = \sum_1^k a_{v_i} \mathbf{x}^{v_i}$ for suitable distinct $v_1, \ldots, v_k \in \mathbb{N}^n$ and $a_{v_i} \in R$. Then

$$\begin{aligned}\mathrm{ev}_{\overline{f_1},\ldots,\overline{f_n}}(f) &= \sum_1^k a_{v_i} (f_1 + J)^{v_{i1}} \cdots (f_n + J)^{v_{in}} = \sum_1^k a_{v_i} \left(f_1^{v_{i1}} + J\right) \cdots \left(f_n^{v_{in}} + J\right) \\ &= \left[\sum_1^k a_{v_i} f_1^{v_{i1}} \cdots f_n^{v_{in}}\right] + J = \mathrm{ev}_{f_1,\ldots,f_n}(f) + J\end{aligned}$$

□

Proposition 2.9.25. *Let S be a commutative R-algebra. Let $\sigma : R[x_1, \ldots, x_n] \to S$ be an R-algebra homomorphism. Then $\sigma = \mathrm{ev}_{\sigma(x_1),\ldots,\sigma(x_n)}$. Hence any element of $\mathrm{Hom}^{R-alg}(R[x_1, \ldots, x_n], S)$ is uniquely determined by it's behavior on the variables of $R[x_1, \ldots, x_n]$.*

Proof. Let $f = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v \in R[x_1, \ldots, x_n]$. Then using the multiplicativity and additivity of σ we have that

$$\sigma(f) = \sigma\left(\sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v\right) = \sum_{v \in \mathbb{N}^n} \sigma(a_v)\sigma\left(x_1^{v_1} \cdots x_n^{v_n}\right) = \sum_{v \in \mathbb{N}^n} a_v \sigma\left(x_1^{v_1}\right) \cdots \sigma\left(x_n^{v_n}\right) = \mathrm{ev}_{\sigma(x_1),\ldots,\sigma(x_n)}(f)$$

□

Corollary 2.9.26. *Let S be a commutative R-algebra. Then*

$$\mathrm{Hom}^{R-Alg}(R[x_1, \ldots, x_n], S) = \left\{\mathrm{ev}_{s_1,\ldots,s_n} : (s_1, \ldots, s_n) \in S^n\right\}.$$

Corollary 2.9.27. *Let $I \subset R[x_1, \ldots, x_n]$, $J \subset R[y_1, \ldots, y_n]$ be ideals. Then $\mathrm{Hom}^{R-Alg}(R[\mathbf{x}]/I, R[\mathbf{y}]/J)$ is equal to*

$$\left\{f + I \mapsto \mathrm{ev}_{f_1,\ldots,f_n}(f) + J : f_1, \ldots, f_n \in R[\mathbf{y}], \mathrm{ev}_{f_1,\ldots,f_n}(f) = 0 \textit{ for every } f \in I\right\} =: F$$

Proof. It is easy to check that $F \subset \mathrm{Hom}^{K-\mathrm{Alg}}(R[\mathbf{x}]/I, R[\mathbf{y}]/J)$. Indeed, consider such a map σ for given $f_1,\dots,f_n \in R[\mathbf{y}]$. Consider

$$\sigma' : R[\mathbf{x}] \to R[\mathbf{y}]/J, f \mapsto \mathrm{ev}_{f_1+J,\dots,f_n+J}(f) = \mathrm{ev}_{f_1,\dots,f_n}(f) + J,$$

is clearly a ring homomorphism satisfying $\sigma'(r) = r$ for $r \in R$, since it is equal to $\pi \circ \mathrm{ev}_{f_1,\dots,f_n}$, where $\pi : R[\mathbf{y} \to R[y]/J$ is the canonical surjection. Since $\mathrm{ev}_{f_1,\dots,f_n}(f) = 0$ for every $f \in I$, it thus follows that σ is a well-defined ring homomorphism satisfying $\sigma(r) = r$ for every $r \in R$ and hence an R-algebra homomorphism.
Let $\sigma \in \mathrm{Hom}^{R-\mathrm{Alg}}(R[\mathbf{x}]/I, R[\mathbf{y}]/J)$. Then $\sigma \circ \pi \in \mathrm{Hom}^{R-\mathrm{Alg}}(R[\mathbf{x}], R[\mathbf{y}]/J)$, where $\pi : R[\mathbf{x}] \to R[\mathbf{x}]/I$ is the canonical surjection. Hence by the prior corollary, $\sigma \circ \pi = \mathrm{ev}_{f_1+J,\dots,f_n+J}$ for suitable $f_1 + J,\dots,f_n + J \in R[\mathbf{y}]/J$. Let $f + I \in R[\mathbf{x}]/I$. Then by Lemma 2.9.24

$$\sigma(f+I) = \sigma \circ \pi(f) = \mathrm{ev}_{f_1+J,\dots,f_n+J}(f) = \mathrm{ev}_{f_1,\dots,f_n}(f) + J,$$

hence $\sigma \in F$. □

Lemma 2.9.28. *Let $S \supset R$ be a ring extension and $s_1,\dots,s_n$ be algebraically independent over R. When $R[x_1,\dots,x_n]$ denotes the polynomial over R in n variables then $R[s_1,\dots,s_n] \simeq R[\mathbf{x}]$.*

Proof. $\mathrm{ev}_{s_1,\dots,s_n} : R[\mathbf{x}] \to R[s_1,\dots,s_n]$ defines a surjective ring homomorphism. Let $f \in R[\mathbf{x}]$. By the definition of algebraic independence

$$\mathrm{ev}_{s_1,\dots,s_n}(f) = 0 \iff f = 0.$$

. $\mathrm{ev}_{s_1,\dots,s_n}$ is therefor injective, implying $R[s_1,\dots,s_n] \simeq R[\mathbf{x}]$ □

Corollary 2.9.29. *Consider the ring extension $R[x_1,\dots,x_n,y_1,\dots,y_m] \supset R$. Then the subring of $R[\mathbf{x},\mathbf{y}]$ generated by $x_1,\dots,x_n$ is isomorphic to the polynomial ring in n variables.*

Corollary 2.9.30. *Consider the ring extension $R[x_1,\dots,x_n,y_1,\dots,y_m] \supset R$. Then $R[\mathbf{x},\mathbf{y}] = R[\mathbf{x}][\mathbf{y}]$. Furthermore $R[\mathbf{x},\mathbf{y}] \simeq R[z_1,\dots,z_n][w_1,\dots,w_m]$.*

Definition 2.9.31. Let $f \in R[x_1,\dots,x_n] \setminus \{0\}$, we define the *degree* of $f = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v$, denoted $\mathbf{deg}\ f$, as the non-negative integer

$$\max\left\{|v| : v \in \mathbb{N}^n, a_v \neq 0\right\}.$$

Remark 2.9.32. For a polynomial $f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v \in K[x_1,\ldots,x_n]$ with $d := \deg f$ we may write

$$f = \sum_{v\in\mathbb{N}^n:|v|\leq d} a_v \mathbf{x}^v.$$

Definition 2.9.33. Let $f \in R[x_1,\ldots,x_n]\setminus\{0\}$, we define a *leading coefficient* of $f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v$, to be a coefficient $a_v \in R\setminus 0$, where $|v| = \deg f$.

Lemma 2.9.34. *If R is an integral domain, then $R[x_1,\ldots,x_n]$ is an integral domain.*

Proof. We proceed by induction in n. Suppose $n = 1$ and let $f, g \in R[x]\setminus 0$. Then $f = \sum_0^d a_i x^i$, $g = \sum_0^{d'} b_i x^i$, for $d, d' \geq 0$, $a_d \neq 0 \neq b_{d'}$. Then

$$fg = \sum_{k=0}^{d+d'} \left(\sum_{0\leq i\leq d, 0\leq j\leq d' : i+j=k} a_i b_j \right) x^k.$$

Note that for $i \leq d$ and $j \leq d'$, $i + j = k$ if and only if $i = d$ and $j = d'$, hence $\sum_{0\leq i\leq d,0\leq j\leq d':i+j=k} a_i b_j = a_d b_{d'} \neq 0$, using that R is a domain. Since $\{x^i : i \in \mathbb{N}\}$ is linearly independent over R it follows that $fg \neq 0$. Suppose $R[x_1,\ldots,x_n]$ is a domain for some $n \geq 1$. Then by the one variable case $R[x_1,\ldots,x_{n+1}] \simeq (R[x_1,\ldots,x_n])[x_{n+1}]$ is a domain. $\square$

Lemma 2.9.35. *The function*

$$\deg : R[x_1,\ldots,x_n]\setminus 0 \to \mathbb{N}$$
$$f \mapsto \deg f$$

has the following properties

1. *The degree function is sub-additive for pairs of distinct polynomials, i.e.* $\deg f + g \leq \max(\deg f, \deg g)$ *for every* $f, g \in R[\mathbf{x}]\setminus 0$ *with* $f \neq g$.

2. *For every* $f, g \in R[\mathbf{x}]\setminus 0$, $\deg f > \deg g \Rightarrow \deg f + g = \deg f$.

3. *The degree function is sub-multiplicative, i.e.* $\deg fg \leq \deg f + \deg g$ *for every* $f, g \in R[\mathbf{x}]\setminus 0$

4. *Suppose R is an integral domain. Then* $\deg fg = \deg f + \deg g$ *for every* $f, g \in R[\mathbf{x}]\setminus 0$.

Proof. Put $d = \deg f$ and $d' = \deg g$, and write $f = \sum_{v\in\mathbb{N}^n:|v|\leq d} a_v \mathbf{x}^v$, $g = \sum_{v\in\mathbb{N}^n:|v|\leq d'} b_v \mathbf{x}^v$
1. Let $v \in \mathbb{N}^n$ such that $|v| > \max(d, d')$. Then in particular $|v| > d$ and $|v| > d'$, meaning $a_v = 0$ and $b_v = 0$, hence $a_v + b_v = 0$. This means

$$\max\{|v| : v \in \mathbb{N}^n, a_v + b_v = 0\} \leq \max(d, d')$$

2. From 1. we have that $\mathbf{deg}\ f+g \le \max(d,d') = d$, hence it suffices to show that $a_v + b_v \neq 0$ for some $v \in \mathbb{N}^n$ with $|v| = d$. There exists a $v \in \mathbb{N}^n$ with $|v| = d$ and $a_v \neq 0$. Since $|v| = d > d'$, $b_v = 0$ hence $a_v + b_v = a_v \neq 0$.
3. Let $u \in \mathbb{N}^n$ be given such that $|u| > d + d'$. Consider $v \in \mathbb{N}^n$ and $w \in \mathbb{N}^n$ with $v + w = u$. Then $|v| + |w| = |u| > d + d'$, hence $|v| > d$ or $|w| > d'$, implying $a_v = 0$ or $b_w = 0$, hence $\sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w = 0$, implying

$$\mathbf{deg}\ fg = \max\left\{ |u| : \sum_{v,w\in\mathbb{N}^n:v+w=u} a_v b_w \neq 0 \right\} \le d + d' = \mathbf{deg}\ f + \mathbf{deg}\ g.$$

4. Let $f' = \sum_{v\in\mathbb{N}^n:|v|=d} a_v \mathbf{x}^v$ and $g' = \sum_{w\in\mathbb{N}^n:|w|=d'} b_w \mathbf{x}^w$. For some $v,w \in \mathbb{N}^n$ with $|v| = d$ and $|w| = d'$ we have that $a_v \neq 0$ and $b_w \neq 0$, hence $f' \neq 0$ and $g' \neq 0$ implying that $f'g' \neq 0$ by Lemma 2.9.34. Furthermore $\mathbf{deg}\ f'g' = d + d'$. Let $r_f = \sum_{v\in\mathbb{N}^n:|v|<d} a_v \mathbf{x}^v$ and $r_g = \sum_{w\in\mathbb{N}^n:|w|<d'} b_w \mathbf{x}^w$. Note that $\mathbf{deg}\ r_f < d$ and $\mathbf{deg}\ r_g < d'$, hence by 3.

$$\mathbf{deg}\ f'r_g \le d + \mathbf{deg}\ r_g < d + d',$$
$$\mathbf{deg}\ g'r_f \le d' + \mathbf{deg}\ r_f < d + d',$$
$$\mathbf{deg}\ r_f r_g \le \mathbf{deg}\ r_f + \mathbf{deg}\ r_g < d + d'.$$

We thus get that

$$\mathbf{deg}\ f'r_g + g'r_f + r_f r_g \le \max(f'r_g, g'r_f, r_f, r_g) < d + d' = \mathbf{deg}\ f'g'.$$

By 2. we get

$$\mathbf{deg}\ fg = \mathbf{deg}\ (f'+r_f)(g'+r_g) = \mathbf{deg}\ f'g' + (f'r_g + g'r_f + r_f r_g) = \mathbf{deg}\ f'g' = d + d' = \mathbf{deg}\ f + \mathbf{deg}\ g.$$

□

Definition 2.9.36. Let $S \supset R$ be a commutative ring extension. Let $f \in R[x_1, \ldots, x_n]$, we say that $(s_1, \ldots, s_n) \in S^n$ is a *zero (over S)* of f if $f(s_1, \ldots, s_n) = 0$. If $f \in R[x]$ and $s \in S$ is a zero of f we call it a *root (in S)*.

Definition 2.9.37. Let $S \supset R$ be a commutative ring extension. Given a polynomial $f \in R[x_1, \ldots, x_n]$, we denote the set of zeroes over S of f by

$$V_S(f).$$

The above definitions are central to the classical treatment of algebraic geometry, since the geometric objects considered are build from set zeroes of polynomials over a field K.

Proposition 2.9.38. *Let S be an R-algebra. Let $f \in R[x_1,\dots,x_n] \subset S[x_1,\dots,x_n]$ and $(s_1,\dots,s_n) \in S^n$, set $I := \langle x_1 - s_1,\dots,x_n - s_n\rangle \subset S[x_1,\dots,x_n]$. Then $(s_1,\dots,s_n)$ is a zero of f if and only if $f \in I$. We call I the* ***point ideal of*** *$(s_1,\dots,s_n)$.*

Proof. Write $f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v$. "$\Rightarrow$": Suppose $(s_1,\dots,s_n)$ is a zero of f. Then, since $x_i + I = s_i + I$ for each i,

$$\begin{aligned} f + I &= \left(\sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v\right) + I = \sum_{v\in\mathbb{N}^n} a_v (x_1 + I)^{v_1}\cdots(x_n+I)^{v_n} = \sum_{v\in\mathbb{N}^n} a_v (s_1+I)^{v_1}\cdots(s_n+I)^{v_n} \\ &= \left(\sum_{v\in\mathbb{N}^n} a_v s_1^{v_1}\cdots s_n^{v_n}\right) + I = f(s_1,\dots,s_n) + I = 0 + I, \end{aligned}$$

thus $f \in I$

"$\Leftarrow$": Suppose $f \in I$. Then there are $\lambda_1,\dots,\lambda_n \in S[x_1,\dots,x_n]$ such that $f = \sum_1^n \lambda_i \cdot (x_i - s_i)$. It then follows that

$$f(s_1,\dots,s_n) = \sum_1^n \lambda_i(s_1,\dots,s_n)(s_i - s_i) = 0,$$

hence $(s_1,\dots,s_n)$ is a zero of f. □

Corollary 2.9.39. *Let $(r_1,\dots,r_n) \in R^n$. If $\mathrm{ev}_{r_1,\dots,r_n}$ is surjective, then $R[x_1,\dots,x_n]/\langle x_1 - r_1,\dots,x_n - r_n\rangle \simeq R$. Hence if R is a field, then $\langle x_1 - r_1,\dots,x_n - r_n\rangle$ is maximal.*

Corollary 2.9.40. *Let $S \supset R$ be a commutative ring extension and consider $f \in R[x]$ and $a \in S$. Then a is a root of f if and only if $x - a \mid f$ in $S[x]$.*

The following theorem is useful when one wants to eliminate certain variables in a finitely generated R-algebra.

Corollary 2.9.41. *Let $f_1,\dots,f_m, g_1,\dots,g_l \in R[x_1,\dots,x_n]$, where R is a commutative ring. Set $I := \langle g_1,\dots,g_l\rangle \subset R[\mathbf{x}]$ and $J := \langle g_1,\dots,g_l, y_1 - f_1,\dots,y_m - f_m\rangle \subset R[\mathbf{x}, y_1,\dots,y_m]$. Then $R[\mathbf{x},\mathbf{y}]/J \simeq R[\mathbf{x}]/I$*

Proof. Consider the surjective ring homomorphism

$$\begin{aligned} \sigma := \mathrm{ev}_{\mathbf{x},f_1,\dots,f_m} : R[\mathbf{x},\mathbf{y}] &\to R[\mathbf{x}]/I \\ h &\mapsto h(\mathbf{x}, f_1,\dots,f_m) + I \end{aligned}$$

Clearly $J \subset \ker\ \sigma$. Let $h \in \ker\ \sigma$. Then $h(x_1,\dots,x_n,f_1,\dots,f_m) \in I$ and hence is also an element of J. It follows that

$$h + J = h(\mathbf{x}, f_1,\dots,f_m) + J = 0 + J \Rightarrow\ h \in J.$$

We thus see by the isomorphism theorem that

$$R[\mathbf{x},\mathbf{y}]/J \simeq R[\mathbf{x}]/I$$

□

Lemma 2.9.42. *Let $S \supset R$ be an integral domain extension. Let $f,g \in R[x_1,\dots,x_n]$ and $v \in S^n$. Then v is a zero of fg if and only if v is a zero of f or g. In other words $V_S(fg) = V_S(f) \cup V_S(g)$.*

Proof. Since R is an integral domain,

$$0 = (fg)(v) = f(v)g(v) \iff f(v) = 0 \text{ or } g(v) = 0.$$

□

Proposition 2.9.43. *Let R be an integral domain. Consider $f \in R[x] \setminus \{0\}$ with $d :=$ $\deg\ f$. Then there at most d roots of f in R.*

Proof. We proceed by induction in d. Let $d = 1$. If f has no roots we are done. Suppose it does have a root $a \in R$. Then Corollary 2.9.40 tells us that $f = q(x-a)$, for some $q \in R[x]$. Since $f \neq 0$, we have that $q \neq 0$. By Lemma 2.9.35 4. it follows that

$$1 = \deg\ f = \deg\ q(x-a) = \deg\ q + \deg x - a = q + 1 \Rightarrow \deg\ q = 0,$$

hence q is a non-zero constant. It follows from 2.9.42 that f has exactly 1 root.
Now consider a polynomial $f \in R[x]$ of degree $d+1$ for some $d \geq 1$. If f has no roots, we are done. Suppose then that f has a root $a \in R$. Then by Corollary 2.9.40 $(x-a) \mid f$, hence $f = g \cdot (x-a)$ for some $g \in R[x]$, again since $f \neq 0$, we have that $g \neq 0$, by Lemma 2.9.35 4. it follows that

$$d + 1 = \deg\ f = \deg\ g + \deg\ x - a = \deg\ g + 1 \Rightarrow \deg\ g = d$$

it follows by induction hypothesis that g has at most d roots. By Lemma 2.9.42 $V(f) = V(g) \cup V(x-a)$, hence $\#V(f) = \#(V(g) \cup V(x-a)) \leq \#V(g) + \#V(x-a) \leq d+1$ □

Lemma 2.9.44. *Let R be an integral domain. Consider $f \in R[x_1,\dots,x_n] \setminus 0$ and $f_1,\dots,f_n \in R[y_1,\dots,y_m] \setminus 0$ with $d_i := \deg\ f_i$. Then*

$$\deg\ f(f_1,\dots,f_n) \leq \deg\ f(x_1^{d_1},\dots,x_n^{d_n})$$

Proof. Write $f = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v$. Let $v \in \mathbb{N}^n$ such that $a_v \neq 0$ and set $M_v = \mathbf{x}^v$. Then

$$\deg\ M_v(x_1^{d_1}, \ldots, x_n^{d_n}) = \sum_1^n v_i d_i = \sum_1^n \deg\ f_i^{v_i} = \deg f_1^{v_1} \cdots f_n{}^{v_n} = \deg\ M_v(f_1, \ldots, f_n).$$

Note that the map $(v_1, \ldots, v_n) \mapsto (v_1 d_1, \ldots, v_n d_n)$ is injective hence,

$$\deg\ f(f_1, \ldots, f_n) \leq \max_{v \in \mathbb{N}^n : a_v \neq 0} \deg\ M_v(f_1, \ldots, f_n) = \max_{v \in \mathbb{N}^n : a_v \neq 0} \deg\ M_v(x_1^{d_1}, \ldots, x_n^{d_n}) = \deg\ f(x_1^{d_1}, \ldots, x_n^{d_n}).$$

□

Remark 2.9.45. The above result doesn't always hold with equality. take for instance $f = x_1 x_2 - x_3$, take $f_1 = y_1$, $f_2 = y_2$ and $f_3 = -y_1 y_3$. Then $f(f_1, f_2, f_3) = 0$, while $f(x_1, x_2, x_3^2) = x_1 x_2 - x_3^2$.

2.9.4 Some Results about Polynomials that I proper subsubsections for

Lemma 2.9.46. *Let $f = \sum_0^d a_i x^i \in K[x] \setminus 0$ and set $I = \langle f \rangle$. Then $K[x]/I$ is a vector space of dimension d with basis $\{x^i + I : i \in \{0, \ldots, d-1\}\}$.*

Proof. One finds that

$$0 = \left[\sum_0^d a_i x^i \right] + I \Rightarrow x^d = - \sum_0^{d-1} \left(a_d^{-1} a_i x^i + I \right),$$

so $\{x^i + I : i \in \{0, \ldots, d-1\}\}$ generates $K[x]/I$ over K. Suppose $g + I = \left[\sum_0^{d-1} b_i x^i\right] + I = 0$. Then $g \in I$. This means either $g = 0$ or $\deg\ g \geq d$, hence $g = 0$ and $a_i = 0$ for $i \in \{1, \ldots, d-1\}$. So $\{x^i + I : i \in \{0, \ldots, d-1\}\}$ is linearly independent over K and is thus a basis for $K[x]/I$, which means $\dim_K\ K[x]/I = d$. □

2.9.5 Polynomials over Infinite Rings

Proposition 2.9.47. *Let R be an infinite integral domain an $f \in R[x_1, \ldots, x_n]$. Then $f = 0$ if and only if $f(v) = 0$ for every $v \in K^n$.*

Proof. "$\Rightarrow$": This is trivial "$\Leftarrow$":We prove that if $f \neq 0$, then there is a $v \in R^n$ such that $f(v) \neq 0$. We prove this by induction in n.
Base case: Consider first the case $n = 1$. Since $f \neq 0$, the number of roots is bounded by the non-negative integer $\deg\ f$ by Proposition 2.9.43. Then since $\#R = \infty$, there is an $a \in R$ such that $f(a) \neq 0$.
Induction hypothesis: Suppose that there is an $n \geq 1$ s.t. if $h \in R[x_1, \ldots, x_n] \setminus 0$, then

there is a $v \in R^n$ such that $f(v) \neq 0$.

Induction Step: Let $f \in R[x_1,\ldots,x_{n+1}] \setminus 0$. We can write

$$f = \sum_0^d f_i x_{n+1}^i,$$

for some $d \geq 0$ and suitable $f_0,\ldots,f_d \in R[x_1,\ldots,x_n]$ where $f_j \neq 0$ for some $j \in \{0,\ldots,d\}$. By the induction hypothesis, there is a $(v_1,\ldots,v_n) \in R^n$ such that $f_j(v) \neq 0$. Then

$$R[x_{n+1}] \ni f' := f(v_1,\ldots,v_n,x_{n+1}) = \sum_0^d f_i(v_1,\ldots,v_n) x_{n+1}^i \neq 0.$$

By the base case there is a $v_{n+1} \in R$ such that $f'(v_{n+1}) \neq 0$. Hence upon putting $v = (v_1,\ldots,v_n,v_{n+1})$ we get that

$$f(v) = f'(v_{n+1}) \neq 0.$$

$\square$

2.9.6 The Hilbert Basis Theorem

Theorem 2.9.48. *(Hilbert Basis Theorem) Let R be a left/right noetherian ring. Then $R[x]$ is left/right noetherian. Furthermore $R[x_1,\ldots,x_n]$ is left/right noetherian.*

Proof. We prove the contrapositive. Suppose That $R[x]$ is not noetherian, or equivalently by Theorem 2.4.58 suppose there is an ideal $I \subset R[x]$ that is not finitely generated. Let $d_1 = \min\{\deg\ f : f \in I\}$. Let $f_1 \in I$ such that $\deg\ f_1 = d_1$. We then let $I_1 = R[x]f_1$ and recursively define $I_n = \sum_1^n R[x]f_i$, where $f_n \in I \setminus I_{n-1}$ where $\deg\ f_n = d_n = \min\{\deg\ f : f \in I \setminus I_{n-1}\}$. Note that since $I \setminus I_n \supset I \setminus I_{n+1}$, $d_n \leq d_{n+1}$ for each $n \geq 1$. For each n we can write

$$f_n = \sum_{i=0}^{d_n} a_i^{(n)} x^i,$$

for suitable $a_i^{(n)} \in R$. Set $a(n) = a_{d_n}^{(n)}$ We then have an ascending chain $J_1 \subset J_2 \subset \ldots$ in R where $J_n = \sum_1^n Ra(i)$. Suppose for a contradiction that $J_n = J_{n+1}$ for some $n \geq 1$. Then

$$a(n+1) = \sum_1^n b_i a(i),$$

for suitable $b_1,\ldots,b_n \in R$. Put

$$g = f_{n+1} - \underbrace{\sum_{i=1}^n \alpha_i x^{d_{n+1}-d_i} f_i}_{h}.$$

Then $g \in I$, $h \in I_n$ and $f_{n+1} = g + h \in I \setminus I_n$, thus $g \in I \setminus I_n$. However, upon further inspection, we find

$$\begin{aligned}
g &= a(n+1)x^{d_{n+1}} - \sum_{i=1}^{n} \alpha_i x^{d_{n+1}-d_i} \sum_{j=1}^{d_i} a_j^{(i)} x^j + \underbrace{\sum_{i=1}^{d_{n+1}-1} a_i^{(n+1)} x^i}_{r} \\
&= a(n+1)x^{d_{n+1}} - \sum_{i=1}^{n} \alpha_i a(i) x^{d_{n+1}-d_i} x^{d_i} - \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{d_i-1} a_j^{(i)} x^i}_{r'} + r \\
&= a(n+1)x^{d_{n+1}} - \left(\sum_{i=1}^{n} \alpha_i a(i)\right) x^{d_{n+1}} + r' + r = a(n+1)x^{d_{n+1}} - a(n+1)x^{d_{n+1}} + r' + r \\
&= \sum_{i=1}^{n+1} \sum_{j=1}^{d_i-1} a_j^{(i)} x^i.
\end{aligned}$$

Thus $\deg\, g = \max\{d_1 - 1, \ldots, d_{n+1} - 1\} = d_{n+1} - 1 < d_{n+1} = \min\{\deg\, f : f \in I \setminus I_n\}$, leading to a contradiction. This means that $J_1 \subset J_2 \subset \ldots$ is a non-stabilizing ascending chain hence R is not noetherian.

Suppose that R is noetherian. Then by induction $R[x_1, \ldots, x_n] \simeq (R[x_1, \ldots, x_{n-1}])[x_n]$ is noetherian. ☐

Corollary 2.9.49. *Let K be a field. Then $K[x_1, \ldots, x_n]$ is noetherian.*

2.9.7 Polynomials over Fields

Definition 2.9.50. A field K is called algebraically closed if every non-constant $f \in K[x] \setminus 0$ has a root $a \in K$.

Lemma 2.9.51. *Let K be a field. Then $K[x]$ is a PID.*

Proof. The trivial ideals are trivially principal. So consider a non-zero proper ideal $I \subset K[x]$. Let $d := \min\{\deg\ f : f \in I\} \geq 1$. Pick an $f \in I$ of degree d. Let $g \in I$. Then there is a $q, r \in K[x]$ where $r = 0$ or $\deg\ r < \deg\ f$ such that $g = qf + r$. By minimality $r = 0$, hence $f \mid g$, hence $I = \langle f \rangle$ ☐

Lemma 2.9.52. *Let K be a field and $f = \sum_0^d a_i x^i \in K[x]$ an irreducible polynomial. Set $I := \langle f \rangle$. Then $F := K[x]/I$ is a field and $x + I$ is a root of $g := \sum_0^d a_i y^i \in F[y]$.*

Proof. Lemma 2.8.54 and Lemma 2.8.55 shows that I is maximal. Then $K[x]/I$ is a field by Proposition 2.8.14. Secondly,

$$g(x+I) = \sum_0^d a_i (x+I) = \left(\sum_0^d a_i x\right) + I = 0 + I.$$

☐

2.9.8 More on Power Series

Lemma 2.9.53. *Let R be any commutative ring. Then*

$$R[\![x_1,\ldots,x_n,y_1,\ldots,y_m]\!] \simeq R[\![x_1,\ldots,x_n]\!][\![y_1,\ldots,y_m]\!] := (R[\![x_1,\ldots,x_n]\!])[\![y_1,\ldots,y_m]\!].$$

Proof. Consider the map $\sigma : R[\![\mathbf{x},\mathbf{y}]\!] \to R[\![\mathbf{x}]\!][\![\mathbf{y}]\!], \sum_{v\in\mathbb{N}^n,w\in\mathbb{N}^m} a_v \mathbf{x}^v \mathbf{y}^w \mapsto \sum_{w\in\mathbb{N}^n} (\sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v)\mathbf{y}^w$. Note that σ is actually just currying of functions $\mathbb{N}^m \times \mathbb{N}^n \to R$. This is trivially a ring homomorphism. The inverse is given by uncurrying. □

Lemma 2.9.54. *If R is an integral domain, so is $R[\![x_1,\ldots,x_n]\!]$.*

Proof. In the case $n = 1$. Consider $f = \sum_{i\in\mathbb{N}} a_i x^i, g = \sum_{i\in\mathbb{N}} b_i x^i \in R[\![x]\!]$. Let $k,l \geq 0$ be the smallest integers such that $a_i \neq 0$, $b_i \neq 0$. Consider $i,j \in \mathbb{N}$ such that $i+j = k+l$. If $i > k$ then $j < l$ and vice versa, hence

$$\sum_{i,j\in\mathbb{N}:i+j=k+l} a_i b_j = a_k b_l \neq 0.$$

By the prior lemma $R[\![x_1,\ldots,x_{n+1}]\!] \simeq R[\![x_1,\ldots,x_n]\!][\![x_{n+1}]\!]$, which by induction and the case $n = 1$ is an integral domain. □

2.9.9 Formal Power Series & DVRs

Lemma 2.9.55. *Suppose R is an integral domain. Then*

$$R[\![x]\!]^* = \left\{ \sum_{i\in\mathbb{N}} a_i \in R[\![x]\!] : a_i \in R^* \right\}$$

Proof. Let $s = \sum_{i\in\mathbb{N}} a_i x^i$ be an element of the right-hand side. Set $b_0 := a_0^{-1}$ and $b_k := -a_0^{-1} \sum_{j=1}^{k} a_j b_{k-j}$ for $k \geq 1$. Define $t = \sum_{i\in\mathbb{N}^n} b_i x^i$. We prove by induction that $\sum_{j,k\in\mathbb{N}:j+k=i} a_j b_k = 0$ for $i \geq 1$. For $i = 1$ we have that

$$\sum_{j,k\in\mathbb{N}:j+k=1} a_j b_k = 0 = a_0 b_1 + a_1 b_0 = -a_0 a_0^{-1} \sum_{h=1}^{1} a_h b_{1-h} + a_1 a_0^{-1} = -a_1 a_0^{-1} + a_1 a_0^{-1} = 0.$$

Then for $i \geq 0$,

$$\sum_{j,k\in\mathbb{N}:j+k=i+1} a_j b_k = \sum_{j,k\in\mathbb{N}:j+k=i+1} -a_j a_0^{-1} \sum_{h=1}^{k} a_h b_{k-h}$$

□

Lemma 2.9.56. *For an integral domain R, $x \in R[\![x]\!]$ is irreducible.*

Proof. x is a non-zero, non-unit. Suppose $x = ab$ for $a, b \in R[\![x]\!]$. Then $a_0 b_0 = 0$, hence $a_0 = 0$ or $b_0 = 0$. Furthermore, $a_0 b_1 + a_1 b_0 = 1$, hence $a_0 b_1 = 1$ or $a_1 b_0 = 1$, hence $a_0 \in R^*$ or $b_0 \in R^*$, hence either a or b is a unit in $R[\![x]\!]$. We thus have that x is irreducible in $R[\![x]\!]$ □

Proposition 2.9.57. *The ring of power series $K[\![x]\!]$ is a DVR with uniformizing parameter x when K is a field.*

Proof. x is irreducible by the above lemma. Let $t \in K[\![x]\!]$. Put $n := \max(\{k \geq 1 : x^k \mid t\})$. Note for $h \in K[\![x]\!]$, $x \mid h$ if and only if h is not a unit, hence $t = sx^n$, where s a unit. uniqueness of this representation follows from the maximality of n and the irreducibility of x. It thus follows that $K[\![x]\!]$ is a DVR with uniformizing parameter x by Proposition 2.8.81. □

Definition 2.9.58. For an integral domain R, we take $R(\!(x_1, \dots, x_n)\!)$ to mean $Q(R[\![x_1, \dots, x_n]\!])$.

Proposition 2.9.59. *Consider the setup and statement of Proposition 2.8.93. Then to each $z \in R$ there is a unique (possibly infinite) sequence $(\lambda_i) \in \prod_{i \in \mathbb{N}} K$. In other words the map*

$$\begin{aligned} \sigma : R &\to K[\![x]\!] \\ z &\mapsto \sum_{i \in \mathbb{N}} \lambda_i x^i \end{aligned}$$

is a well-defined map. It is furthermore an injective ring K-algebra homomorphism. It extends to a homomorphism of $L = Q(R)$ onto $K(\!(x)\!)$.

Proof. clearly map fixes K. Let $z, w \in R$ be given with associated power series $\sum_{i \in \mathbb{N}} \lambda_i x^i$ resp. $\sum_{i \in \mathbb{N}} \mu_i x^i$. Then for any $n \geq 0$, $z = \sum_0^n \lambda_i t^i + z_n t^{n+1}$, $w = \sum_0^n \mu_i t^i + w_n t^{n+1}$ for suitable unique $z_{n+1}, w_{n+1} \in R$, hence

$$\begin{aligned} z + w = \sum_0^n (\lambda_i + \mu_i) t^i + (w_n + z_n) \Rightarrow \sigma(z + w) &= \sum_{i \in \mathbb{N}} (\lambda_i + \mu_i) x^i \\ &= \sum_{i \in \mathbb{N}^n} \lambda_i x^i + \sum_{i \in \mathbb{N}^n} \mu_i x^i = \sigma(z) + \sigma(w). \end{aligned}$$

Let $n \geq 0$. Then

$$zw = \sum_0^{2n} \left(\sum_{i,j \in \mathbb{N} : i+j=h} \lambda_i \mu_j \right) t^h + \underbrace{\sum_0^n (\lambda_i w_k + \mu_i z_k) t^{i+n+1} + w_n z_n t^{n+2}}_{r}.$$

Since $\operatorname{ord}(r) \geq n+1$, it follows that the n'th coefficient of the formal power series of zw is equal to $\sum_{i+j=n} \lambda_i \mu_i$, hence

$$\sigma(zw) = \sum_{h \in \mathbb{N}} \left(\sum_{i,j \in \mathbb{N} : i+j=h} \lambda_i \mu_j \right) x^h = \sigma(z)\sigma(w).$$

Injectivity follows from the uniqueness of the coefficients in the power series. Hence $\ker\sigma = 0$, hence Lemma 2.8.72 implies that

$$\overline{\sigma} : L \to K(\!(x)\!), \frac{z}{w} \mapsto \frac{\sigma(z)}{\sigma(w)}$$

is the unique extension of σ to a K-algebra homomorphism between the fraction fields of R and $K[\![x]\!]$. □

Remark 2.9.60. The unique formal power series $\sum_{i\in\mathbb{N}} \lambda_i \in K[\![x]\!]$ associated with $z \in R$ is called *the power series expansion of z in terms of t*,

2.9.10 Term Orders & a Polynomial Division Algorithms

Definition 2.9.61. A *term order* is total order $\leq$ on $\mathbb{N}^n$ such that

1. $0 \leq v$ for every $v \in \mathbb{N}^n$,
2. for every $v_1, v_2, v \in \mathbb{N}^n, v_1 \leq v_2 \Rightarrow v_1 + v \leq v_2 + v$.

Example 2.9.62. 1. A simple example of a term order is $\leq$ on $\mathbb{N}$.

2. *The lexicographic term order*, denoted $\leq_{\text{lex}}$, on $\mathbb{N}^n$ for $v = (v_1, \ldots, v_n), w = (w_1, \ldots, w_n) \in \mathbb{N}^n$ is defined by $v \leq_{\text{lex}} w$ if $v = w$ or there is an $i \in \{1, \ldots, n\}$ such that $v_j = w_j$ for $j < i$ and $v_i < w_i$. For example $\left(2, 10^6, 10^{10^6}\right) \leq_{\text{lex}} (3,1,1)$ since $2 < 3$. This is indeed a term order: We first check that it is a total order. By definition it is reflexive. Let $v, w, u \in \mathbb{N}^n$.
Note that if $v \neq w$, then there is a minimal i such that $v_i \neq w_i$, hence either $v <_{\text{lex}} w$ or $w <_{\text{lex}} v$. Hence in general $v \leq_{\text{lex}} w$ or $w \leq_{\text{lex}} v$.
If there is an i such $v_i < w_i$ and $v_j = w_j$ for $j < i$ then $v \neq w$. Hence if $v \leq_{\text{lex}} w$ and $w \leq_{\text{lex}} v$, then necessarily $v = w$.
Suppose $v \leq_{\text{lex}} w$ and $w \leq_{\text{lex}} u$. We check by cases that $v \leq_{\text{lex}} u$.
Case 1: Suppose first $v = w$ and $w = u$. Then $v = u$, implying $v \leq_{\text{lex}} u$.
Case 2: Suppose $v = w$ and that there is an i such that $w_i < u_i$ and $w_j = u_j$ for every $j < i$. Then $v_i = w_i < u_i$ and $v_j = w_j = u_j$ for $j < i$ hence $v \leq_{\text{lex}} u$.
Case 3: Suppose there are $h, i \in \{1, \ldots, n\}$ such that $v_h < w_h$, $w_i < u_i$ and $v_j = w_j$, $w_k = u_k$ for $h < j$, $i < k$. If $h \leq i$, then $v_h < w_h \leq u_h$ and $v_j = w_j = u_j$ for $j < h$. If $i < h$, then $v_i = w_i < u_i$ and $v_j = w_j = u_j$ for $j < i$. In any case $v \leq_{\text{lex}} u$.
Case 4: Suppose there is an i such that $v_i < w_i$ and $v_j = w_j$ for every $j < i$ and $w = u$. Then $v_i < w_i = u_i$ and $v_j = w_j = u_j$ for every $j < i$, hence $v \leq_{\text{lex}} u$.

In conclusion $\leq_{\text{lex}}$ is a total order. Note that $0 \leq v_i$ for every $i \in \{1,\ldots,n\}$. Hence either $0 = v_i$ for every i or there is an i such that $0 < v_i$, meaning $0 \leq_{\text{lex}} v$. Suppose $v \leq_{\text{lex}} w$. If $v = w$ then, $v+u = w+u$, hence $v+u \leq_{\text{lex}} w+u$. Suppose there is an i such that $v_i < w_i$ and $v_j = w_j$ for each $j < i$. Then $v_i + u_i < w_i + u_i$ and $v_j + u_j = w_j + u_j$ for every $j < i$, which implies $v+u \leq_{\text{lex}} w+u$.

For $v \in \mathbb{N}^n$ define

$$v + \mathbb{N}^n = \left\{v + w : w \in \mathbb{N}^n\right\}.$$

Theorem 2.9.63. *(Dickson's Lemma)*
Let $S \subset \mathbb{N}^n$ be non-empty. Then there are vectors $v_1,\ldots,v_m \in S$ such that

$$S \subset \bigcup_1^m \left(v_i + \mathbb{N}^n\right)$$

Proof. We proceed by induction in n. For $n = 1$, S has a minimal element s by the well ordering of the natural numbers, hence any element of S can be written as $s+t$ for some $t \in \mathbb{N}$. Suppose Dickson's lemma is true for some $n \geq 1$. Let S be some non-empty subset of $\mathbb{N}^{n+1}$. Consider the canononical surjection

$$\pi : \mathbb{N}^{n+1} \to \mathbb{N}^n$$

$$(v_1,\ldots,v_n,v_{n+1}) \mapsto (v_1,\ldots,v_n)$$

Consider the set

$$S' := \pi(S) = \left\{(x_1,\ldots,x_n) \in \mathbb{N}^n : (x_1,\ldots,x_n,x_{n+1}) \in S \text{ for some } x_{n+1} \in \mathbb{N}\right\}.$$

By induction there are $s_1 = (s_{11},\ldots,s_{1,n+1}),\ldots,s_m = (s_{m1},\ldots,s_{m,n+1}) \in S$ such that upon defining $s'_i := (s_{i1},\ldots,s_{in}) \in S'$

$$S' \subset \bigcup_1^m \left(s'_i + \mathbb{N}^n\right).$$

Let $s_{\max} = \max_{i\in\{1,\ldots,m\}} s_{i,(n+1)}$. Define

$$S_i := \{v = (v_1,\ldots,v_{n+1}) \in S : v_1 = i\} \quad (i \in \{0,\ldots,s_{\max}\})$$

and put

$$S_{\max} := \{v = (v_1,\ldots,v_{n+1}) \in S : v_{n+1} \geq s_{\max}\}.$$

Note that $S_{\max} \subset \bigcup_1^m \left(s_i + \mathbb{N}^{n+1}\right)$. Indeed, if $x \in S_{\max}$, then $(x_1,\ldots,x_n) \in \bigcup_1^m \left(s'_i + \mathbb{N}^n\right)$. In particular, for some $i \in \{1,\ldots,m\}$ $x = (s_{i1} + v_1,\ldots,s_{in} + v_n) \in s'_i + \mathbb{N}^n$. Since $x_{n+1} \geq s_{\max}$, it follows that $x_{n+1} = s_{i,n+1} + v_{n+1}$, and thus that

$$x = \left(s_{i1} + v_1,\ldots,s_{in} + v_n, s_{i,n+1} + v_{n+1}\right) \in s_i + \mathbb{N}^{n+1} \subset \bigcup_1^m \left(s_j + \mathbb{N}^{n+1}\right).$$

. We furthermore have that $S = S_{\max} \cup \bigcup_0^{s_{\max}-1} S_i$. Again using induction there are $s_1^{(i)}, \ldots, s_{m_i}^{(i)} \in \pi(S_i)$ such that

$$\pi(S_i) \subset \bigcup_{j=1}^{m_i} \left(s_j^{(i)} + \mathbb{N}^n\right)$$

Then

$$S_i \subset \bigcup_{j=1}^{m_i} \left(\left(s_{j1}^{(i)}, \ldots, s_{jn}^{(i)}, i\right) + \mathbb{N}^{n+1}\right).$$

We also have that

$$S_{\max} \subset \bigcup_1^m \left(s_j + \mathbb{N}^{n+1}\right).$$

It thus follows that

$$S \subset \bigcup_1^m \left(s_j + \mathbb{N}^{n+1}\right) \cup \bigcup_{i=0}^{s_{\max}} \bigcup_{j=1}^{m_i} \left(\left(s_{j1}^{(i)}, \ldots, s_{jn}^{(i)}, i\right) + \mathbb{N}^{n+1}\right).$$

□

Corollary 2.9.64. *A term ordering* $\le$ *on* $\mathbb{N}^n$ *is a well-ordering.*

Proof. Let $S \subset \mathbb{N}^n$ be a non-empty subset. By Dickson's lemma there are $s_1, \ldots, s_m \in S$ such that $S \subset \bigcup_1^m (s_i + \mathbb{N}^n)$. Since $\le$ is a total order, we can define $s_{\min} := \min_{i \in \{1,\ldots,m\}} s_i$. Let $s \in S$. For some $j \in \{1, \ldots, m\}$, $s = s_j + v$ for some $v \in \mathbb{N}^n$. Now we have the following implications using properties of term orders,

$$0 \le v \text{ and } s_{\min} \le s_j \Rightarrow s_{\min} \le s_{\min} + v \le s_j + v = s,$$

hence $s_{\min}$ is a least element of S, implying $\le$ is a well-ordering. □

A term order on $\mathbb{N}^n$ defines a total order on the monomials in $R[x_1, \ldots, x_n]$ by defining $x^v \le x^w$ if $v \le w$. This total order will have the property that $x^{v_1} \le x^{v_2} \Rightarrow x^{v_1+v} \le x^{v_2+v}$ and $1 \le x^v$. From this definition we gain a way of comparing polynomials using initials terms with respect to a term order.

Definition 2.9.65. Let $f = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v \in R[x_1, \ldots, x_n] \setminus 0$ and $\le$ a term order on $\mathbb{N}^n$. We define the *initial term of* f *with respect to* $\le$ to be the monomial

$$\text{in}_{\le} f := \max_{v \in \mathbb{N}^n : a_v \neq 0} a_v \mathbf{x}^v.$$

Lemma 2.9.66. *Let* $f, g \in R[x_1, \ldots, x_n] \setminus 0$. *Then one finds that*

1. $(\text{in}_{\le} f + g) \le \max(\text{in}_{\le} f, \text{in}_{\le} g)$
2. *If* $\text{in}_{\le} f < \text{in}_{\le} g$, *then* $(\text{in}_{\le} f + g) = \text{in}_{\le} g$.

3. *If the leading terms of* f *and* g *are equal then* $\text{in}_{\leq}(f-g) < \text{in}_{\leq} f = \text{in}_{\leq} g$.

4. $\text{in}_{\leq} fg \leq (lm_{\leq} f)(lm_{\leq} g)$.

5. *Suppose* R *is an integral domain. Then* $\text{in}_{\leq} fg = (\text{in}_{\leq} f)(\text{in}_{\leq} g)$.

Proof. Write $f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v$ and $g = \sum_{v\in\mathbb{N}^N} b_v \mathbf{x}^v$. Let $w,u \in \mathbb{N}^n$ be given such that $a_v\mathbf{x}^w = \text{in}_{\leq} f$ and $b_w\mathbf{x}^u = \text{in}_{\leq} g$. The proofs here are very similar Lemma 2.9.35 in some respects. 1. Let $v \in \mathbb{N}^n$ be given such that $v > \max(w,u)$. Then $a_v, b_v = 0$ hence $a_v + b_v = 0$. It thus follows that

$$(\text{in}_{\leq} f+g) = \max_{v\in\mathbb{N}^n : a_v+b_v\neq 0} (a_v+b_v)\mathbf{x}^v \leq \max(\text{in}_{\leq} f, \text{in}_{\leq} g).$$

2. When $\text{in}_{\leq} f < \text{in}_{\leq} g$, $(\text{in}_{\leq} f+g) \leq \max(\text{in}_{\leq} f, \text{in}_{\leq} g)$. Note that $a_u + b_u = b_u \neq 0$, hence $(\text{in}_{\leq} f+g) = \text{in}_{\leq} g$.
3. Since leading terms of f and g are equal $w = u$ and $a_w = b_w$. Hence if $v \in \mathbb{N}^n$ is given such that $v \geq w$, we have that $a_v - b_v = 0$, thus it follows that

$$\text{in}_{\leq}(f-g) = \max_{v\in\mathbb{N}^n : a_v-b_v\neq 0} (a_v-b_v)\mathbf{x}^v < \text{in}_{\leq} f.$$

4. & 5. Let $v_1, v_2 \in \mathbb{N}^n$ such that $v_1 + v_2 = w + u$ and $v_1 \neq w$. If $v_1 > w$ or $v_1 < w$. Hence

$$a_{v_1} = 0 \Rightarrow a_{v_1} b_{v_2} = 0.$$

In the other case $v_2 > u$, because otherwise $v_1 + v_2 < u + w$. Hence

$$b_{v_2} = 0 \Rightarrow a_{v_1} b_{v_2} = 0.$$

It thus follows that

$$\sum_{v_1,v_2\in\mathbb{N}^n : v_1+v_2=w+u} a_{v_1} b_{v_2} = a_w b_u.$$

Let $v > u + w$. Let $v_1, v_2 \in \mathbb{N}^n$ such that $v_1 + v_2 = v$. Then $a_{v_1} b_{v_2} = 0$, implying $\sum_{v_1,v_2\in\mathbb{N}^n : v_1+v_2=v} a_{v_1} b_{v_2} = 0$. We thus have that

$$\text{in}_{\leq} fg = \max_{v\in\mathbb{N}^n : \sum_{v_1,v_2\in\mathbb{N}^n : v_1+v_2=v} a_{v_1} b_{v_2} \neq 0} \left[\sum_{v_1,v_2\in\mathbb{N}^n : v_1+v_2=v} a_{v_1} b_{v_2} \right] \mathbf{x}^v = \mathbf{x}^{w+u} = (\text{in}_{\leq} f)(\text{in}_{\leq} g).$$

If R is an integral domain we get that

$$\text{in}_{\leq} fg = \left[\sum_{v_1,v_2\in\mathbb{N}^n : v_1+v_2=w+u} a_{v_1} b_{v_2} \right] \mathbf{x}^{w+u} = \left(a_w \mathbf{x}^w\right)\left(b_u \mathbf{x}^u\right) = (\text{in}_{\leq} f)(\text{in}_{\leq} g).$$

□

The upshot of introducing this tool of bookkeeping is that it allows to do polynomial division. For $\mathbf{x}^v \mid \mathbf{x}^w$ we define $\frac{a\mathbf{x}^w}{b\mathbf{x}^v} := \frac{a}{b}\mathbf{x}^{w-v}$.

Theorem 2.9.67. *Let R be an integral domain. Let $f, f_1, \ldots, f_m \in R[x_1, \ldots, x_n] \setminus 0$. Put $F = \{f_1, \ldots, f_m\}$. Then there are $\lambda_1, \ldots, \lambda_m, f^F \in R[x_1, \ldots, x_n]$ such that*

$$f = \left[\sum_1^m \lambda_i f_i\right] + f^F,$$

and $\mathrm{in}_\leq\ \lambda_i f_i \leq \mathrm{in}_\leq\ f$ for every $i \in \{1, \ldots, m\}$ with $\lambda_i \neq 0$ and $f^F = 0$ or $\mathrm{in}_\leq\ f_i \nmid f^F$ for every i.

Proof. We aim to provide a division algorithm that produces the desired the $\lambda_1, \ldots, \lambda_m, f^F$. Define $\lambda_i^{(0)} := 0$ for every $i \in \{1, \ldots, m\}$, $r^{(0)} = 0$ and $s^{(0)} = f$. We note that

$$f = \left[\sum_1^m \lambda_i^{(0)} f_i\right] + \left(r^{(0)} + s^{(0)}\right).$$

We want to recursively define $\lambda_1^{(j)}, \ldots, \lambda_m^{(j)}, r^{(j)}, s^{(j)} \in R[\mathbf{x}]$ such that

$$f = \left[\sum_1^m \lambda_i^{(j)} f_i\right] + \left((r^{(j)} + s^{(j)}\right) \tag{2}$$

for every j and have that $s^{(N)} = 0$ at some N such that putting $\lambda_i = \lambda_i^{(N)}$, $f^F = s^{(N)}$ these polynomials will have the remaining desired properties. For $j \geq 0$ if $s^{(j)} = 0$ put $N = j$ and terminate, otherwise if there is an $i \in \{1, \ldots, m\}$ such $\mathrm{in}_\leq\ f_i \mid \mathrm{in}_\leq\ s^{(j)}$, and pick the smallest such. We then define

$$\begin{cases} s^{(j+1)} := s^{(j)} - \frac{\mathrm{in}_\leq\ s^{(j)}}{\mathrm{in}_\leq\ f_i} f_i, \\ \lambda_i^{(j+1)} := \lambda_i^{(j)} + \frac{\mathrm{in}_\leq\ s^{(j)}}{\mathrm{in}_\leq\ f_i}, \\ \lambda_k^{(j+1)} = \lambda_k^{(j)} \text{ for } k \neq i, \\ r^{(j+1)} := r^{(j)}. \end{cases} \tag{3}$$

We note that indeed the identity (2) is fulfilled for $j+1$ since it is obtained by adding and subtracting $\frac{\mathrm{in}_\leq\ s^{(j)}}{\mathrm{in}_\leq\ f_i} f_i$. If no such i exists we instead define

$$\begin{cases} r^{(j+1)} := r^{(j)} + \mathrm{in}_\leq\ s^{(j)}, \\ s^{(j+1)} := s^{(j)} - \mathrm{in}_\leq\ s^{(j)}, \\ \lambda_i^{(j+1)} := \lambda_i^{(j)}. \end{cases} \tag{4}$$

Again clearly (2) is still true for $j+1$, since $r^{(j+1)} + s^{(j+1)} = r^{(j)} + s^{(j)}$.
We now show that the above algorithm terminates. Let $j \geq 0$ such that $s^{(j+1)} \neq 0$.

Consider that we land in case (3). We denote $\mathbf{in}_\leq\ s^{(j)} = \alpha_v \mathbf{x}^v$ and $\mathbf{in}_\leq\ f_i = \beta_w \mathbf{x}^w$, where i is the minimal index for which the initial term of f_i divides the initial term of $s^{(j)}$. Then

$$\mathrm{in}_\leq \left(\frac{\mathrm{in}_\leq\ s^{(j)}}{\mathrm{in}_\leq\ f_i} f_i \right) = \frac{a_v}{b_w} \mathbf{x}^{v-w} b_w \mathbf{x}^w = a_v \mathbf{x}^v = \mathrm{in}_\leq\ s^{(j)}$$

we have thus have that

$$\mathrm{in}_\leq\ s^{(j+1)} = \mathrm{in}_\leq \left(s^{(j)} - \frac{\mathrm{in}_\leq\ s^{(j)}}{\mathrm{in}_\leq\ f_i} f_i \right) < \mathrm{in}_\leq\ s^{(j)}.$$

Landing in case (4) we have that

$$\mathrm{in}_\leq\ s^{(j+1)} = \mathrm{in}_\leq \left(s^{(j)} - \mathrm{in}_\leq\ s^{(j)} \right) < \mathrm{in}_\leq\ s^{(j)}.$$

Then sequence of non-zero $s^{(j)}$ is thus a strictly decreasing sequence. Let S denote the set of these elements. Since $s^{(0)} = f \neq 0$, $S \neq \emptyset$. This means S has a minimal element $s^{(N-1)}$, since a term order is a well-ordering. Then $s^{(N)} = 0$, for otherwise $s^{(N)} < s^{(N-1)}$.

As advertised we put $a_i := a_i^{(N)}$ for $i \in \{1,\dots,m\}$ and $f^F := r^{(N)}$. For each $j \geq 0$ for each $i \in \{1,\dots,m\}$ for which $a_i^{(j)} = 0$ and $a_i^{(j+1)} \neq 0$ we have that

$$\mathrm{in}_\leq\ a_i^{(j+1)} f_i = \mathrm{in}_\leq \left(\left(a_i^{(j)} + \frac{\mathrm{in}_\leq\ s^{(j)}}{\mathrm{in}_\leq\ f_i} \right) f_i \right) = \mathrm{in}_\leq \left(\frac{\mathrm{in}_\leq\ s^{(j)}}{\mathrm{in}_\leq\ f_i} f_i \right) = \mathrm{in}_\leq\ s^{(j)} \leq \mathrm{in}_\leq\ f$$

It thus follows by induction in the j for which $a_i^{(j)} \neq 0$ that

$$\begin{cases} \mathrm{in}_\leq\ a_i^{(j+1)} f_i = \mathrm{in}_\leq\ a_i^{(j)} f_i \leq \mathrm{in}_\leq\ f, \\ \text{or} \\ \mathrm{in}_\leq\ a_i^{(j+1)} f_i = \mathrm{in}_\leq \left(\left(a_i^{(j)} + \frac{\mathrm{in}_\leq\ s^{(j)}}{\mathrm{in}_\leq\ f_i} \right) f_i \right) \leq \max\left(a_i^{(j)} f_i, \mathrm{in}_\leq\ s^{(j)} \right) \leq \mathrm{in}_\leq\ f. \end{cases}$$

It thus follos that if $a_i \neq 0$,

$$\mathrm{in}_\leq\ a_i f_i \leq \mathrm{in}_\leq\ f.$$

Note lastly that each $r^{(j)}$ is 0 or a sum of terms not divisible by any $\mathbf{in}_\leq\ f_i$, and hence f^F is either 0 or not divisible by any $\mathbf{in}_\leq\ f_i$. □

2.9.11 Gröbner Bases and Buchbergers Algorithm

For the exploration of Gröbner bases we fix a field K.

Definition 2.9.68. Let $I \subset K[x_1,\dots,x_n]$ be an ideal. Let $\leq$ be a term ordering on $K[\mathbf{x}]$. A finite set of polynomials $G \subset K[\mathbf{x}] \setminus 0$ is called a *Gröbner basis* for I with respect to $\leq$, if $G \subset I$ and for every $f \in I \setminus 0$ there is a $g \in G$ such that $\mathrm{in}_\leq\ g \mid \mathrm{in}_\leq\ f$. A finite set $G = \{f_1,\dots,f_m\} \subset K[\mathbf{x}] \setminus 0$ is called a Gröbner basis with respect to $\leq$ if it is a Gröbner basis for $\langle f_1,\dots,f_m \rangle$ with respect to $\leq$.

Proposition 2.9.69. *Let $G = \{f_1, \dots, f_m\} \subset I$ be a Gröbner basis for an ideal $I \subset K[x_1, \dots, x_n]$ with respect to a term order $\leq$. For $f \in K[\mathbf{x}]$,*

$$f \in I \iff f^G = 0$$

Proof. "$\Leftarrow$": If $f^G = 0$ there are $\lambda_1, \dots, \lambda_m \in K[x_1, \dots, x_n]$ such that $f = \sum_1^m \lambda_i f_i \in I$ using the division algorithm.
"$\Rightarrow$": Suppose $f \in I$. Suppose for a contradiction that $f^G \neq 0$. Using the division algorithm with respect to $\leq$ we obtain $\lambda_1, \dots, \lambda_m \in K[\mathbf{x}]$ such that

$$f = \left[\sum_1^m \lambda_i f_i\right] + f^G \Rightarrow f^G = f - \sum_1^m \lambda_i f_i \in I.$$

Then since G is a Gröbner basis there is some $i \in \{1, \dots, m\}$ such that $\mathrm{in}_{\leq} f_i \mid f^G$, but since $f^G \neq 0$ this is not possible. ☐

As a corollary we obtain that every Gröbner basis of an ideal with respect to some term order will be a generating set for said ideal.

Corollary 2.9.70. *Let $I \subset K[x_1, \dots, x_n]$ be an ideal and $G \subset I$ a Gröbner basis for I with respect to some term order $\leq$. Then $I = \langle G \rangle$.*

Proof. By definition $G \subset I \Rightarrow \langle G \rangle \subset I$. Let $f \in I$. By the above proposition, $f^G = 0$, meaning there are $\lambda_1, \dots, \lambda_m \in K[\mathbf{x}]$ such that

$$f = \sum_1^m \lambda_i f_i \in \langle G \rangle$$

☐

A somewhat curious consequence of the introduction of Gröbner bases is that it provides us with a rather simple way to prove the Hilbert basis theorem over fields. I.e. one can prove that any polynomial ideal over a field has a Gröbner basis, which by the above corollary constitutes a finite generating set.

Corollary 2.9.71. *(Hilbert's basis theorem over fields) Let $I \subset K[x_1, \dots, x_n]$ be a non-zero ideal and $\leq$ a term order. Then there is a Gröbner basis $G = \{f_1, \dots, f_m\} \subset I$ for I with respect to $\leq$, hence $I = \langle f_1, \dots, f_m \rangle$ by the prior corollary.*

Proof. Put $S = \{v \in \mathbb{N}^n : \mathbf{x}^v = \mathrm{in}_{\leq} f \text{ for some } f \in I\}$. Clearly $S \neq \emptyset$, hence by Dickson's lemma we may find $v_1, \dots, v_m \in S$ such that

$$S \subset \bigcup_1^m \left(v_i + \mathbb{N}^n\right)$$

Let $f_i \in I$ be given such that $\mathrm{in}_\leq f_i = \mathbf{x}^{v_i}$ for $i \in \{1,\dots,m\}$ and put $G = \{f_1,\dots,f_m\} \subset I$. Let $f \in I \setminus 0$, and pick $v \in \mathbb{N}^n$ such that $a_v\mathbf{x}^v = \mathrm{in}_\leq f$. Since, $v \in S$, $v = v_j + w$ for some $j \in \{1,\dots,m\}$ and $w \in \mathbb{N}^n$. Then

$$\mathrm{in}_\leq f = a_v\mathbf{x}^v = a_v\mathbf{x}^{v_j+w} = \left(a_v\mathbf{x}^w\right)\mathbf{x}^{v_j} = \left(a_v\mathbf{x}^w\right)\mathrm{in}_\leq f_j \Rightarrow \mathrm{in}_\leq f_j \mid \mathrm{in}_\leq f.$$

This verifies that G is a Gröbner basis for I. □

The machinery of Gröbner bases provides a way to perform the division algorithm with respect to a term order in a fashion that ensures uniqueness of remainders and the indifference of the order of the divisor polynomials.

Theorem 2.9.72. *Let $\leq$ be a term order on $K[x_1,\dots,x_n]$ and $G = \{f_1,\dots,f_m\} \subset K[\mathbf{x}]\setminus 0$ a Gröbner basis with respect to $\leq$. Let $f \in K[\mathbf{x}]\setminus 0$. Then any polynomial r satisfying the properties of f^G obtained from the division algorithm of f by $f_1,\dots,f_m$ is equal to f^G. Furthermore, the remainder outputted by the division algorithm remains unchanged after a permutation of $f_1,\dots,f_m$.*

Proof. Let $\lambda_1,\dots,\lambda_m \in K[\mathbf{x}]$ such that

$$f = \left[\sum_1^m \lambda_i f_i\right] + f^G.$$

Suppose there is an $r \in K[\mathbf{x}]$ with $r = 0$ or r is not divisible by the initial term of any f_i such that there are $\lambda'_1,\dots,\lambda'_m \in K[\mathbf{x}]$ satisfying

$$f = \left[\sum_1^m \lambda'_i f_i\right] + r.$$

Then

$$f^G - r = \sum_1^m (\lambda_i - \lambda'_i) f_i \in I.$$

Suppose for a contradiction $f^G - r \neq 0$. Then there is a $j \in \{1,\dots,m\}$ such that $\mathrm{in}_\leq f_j \mid \mathrm{in}_\leq \left(f^G - r\right)$ implying $\mathrm{in}_\leq f_j \mid \mathrm{in}_\leq f^G$ or $\mathrm{in}_\leq f_j \mid \mathrm{in}_\leq r$ leading to a contradiction. It follows that $f^G = r$.
Let $\omega \in \mathscr{S}(m)$ be a permutation. Let $\lambda'_1,\dots,\lambda'_m,(f^G)' \in K[\mathbf{x}]$ be the outcome of the division with respect to $\leq$ of f with $f_{\omega(1)},\dots,f_{\omega(m)}$. Then by uniqueness of the remainder $f^G = (f^G)'$. □

We have now to some extend motivated the usefulness of Gröbner bases (even though we are yet to see the most impressive applications!). However, as a computational tool, they are unimpressive if there is no way to compute. The introduction of S-polynomials and Buchberger's S-criterion will lead us to Buchberger's algorithm for computing Gröbner bases.

Definition 2.9.73. Let $f \in K[x_1,\dots,x_n]$ and $F=\{f_1,\dots,f_m\}\subset K[\mathbf{x}]\setminus 0$. We say that f *reduces to zero modulo* $\boldsymbol{F}$ if there are $\lambda_1,\dots,\lambda_m \in K[\mathbf{x}]$ such that

$$f=\sum_1^m \lambda_i f_i$$

and $\mathbf{in}_\le \lambda_i f_i \le \mathbf{in}_\le f$ for $i \in \{1,\dots,m\}$ with $a_i f_i \neq 0$. This will be denoted $f \to_F 0$.

Note that this definition does not depend on a term order, however this definition leads us to the following reformulation of Proposition 2.9.69. Before formulating this consider the following lemmas

Lemma 2.9.74. *Let* $F=\{f_1,\dots,f_m\}\subset K[\mathbf{x}]\setminus 0$ *and let* $f \in I := \langle F\rangle$ *be non-zero with initial term* $a_v\mathbf{x}^v$*. Consider* $\lambda_1,\dots,\lambda_m \in K[\mathbf{x}]$ *such that*

$$f=\sum_1^m \lambda_i f_i.$$

For each $i \in \{1,\dots,m\}$ *where* $\lambda_i \neq 0$*, pick* $v_i, w_i \in \mathbb{N}^n$ *such that* $b_i\mathbf{x}^{v_i} = \mathbf{in}_\le \lambda_i$ *and* $c_i\mathbf{x}^{w_i} = \mathbf{in}_\le f_i$*. set*

$$\kappa = \max\ \left\{v_i + w_i \in \mathbb{N}^n : i \in \{1,\dots,m\}, \lambda_i \neq 0\right\}.$$

Then $v \le \kappa$ *and the following statements hold*

1. $v=\kappa \iff \mathbf{in}_\le f_i\lambda_i \le \mathbf{in}_\le f$ *for every* $i\in\{1,\dots,m\}$ *such that* $\lambda_i \neq 0$.

2. $v=\kappa \Rightarrow \mathbf{in}_\le f_i \mid \mathbf{in}_\le f$ *for some* $i \in \{1,\dots,m\}$

Proof. Forgetting briefly that we assumed $f=\sum_1^m \lambda_i f_i$, if $v>\kappa$, then

$$\mathbf{in}_\le f > \max_{i\in\{1,\dots,m\}:\lambda_i\neq 0} \mathbf{in}_\le \lambda_i f_i = \mathbf{in}_\le\left(\sum_1^m \lambda_i f_i\right) \Rightarrow f \neq \sum_1^m \lambda_i f_i.$$

It thus follows that since we assumed $f=\sum_1^m \lambda_i f_i$, we get the bound $v\le\kappa$.
1. For every $i\in\{1,\dots,m\}$ such that $\lambda_i f_i \neq 0$ we have that

$$b_i c_i \mathbf{x}^{v_i+w_i} = \mathbf{in}_\le \lambda_i \le \mathbf{in}_\le f = a_v\mathbf{x}^v \iff \mathbf{x}^{v_i+w_i} \le \mathbf{x}^v \iff v_i + w_i \le v$$

hence

$$\kappa = v \iff \mathbf{in}_\le \lambda_i f_i \le .$$

2. Suppose $v=\kappa$. Then $v = v_i + w_i$and hence $b_i c_i = a_v$ for some $i \in \{1,\dots,m\}$. WLOG we may then write

$$\mathbf{in}_\le f = a_v\mathbf{x}^v = \left[\sum_1^l b_i c_i\right]\mathbf{x}^{v_1}\mathbf{x}^{w_1} = \left(\left[\sum_1^l b_i \frac{c_i}{c_1}\right]\mathbf{x}^{v_1}\right) c_1\mathbf{x}^{w_1} = \left(\left[\sum_1^l b_i \frac{c_i}{c_1}\right]\mathbf{x}^{v_1}\right)\mathbf{in}_\le f_i \Rightarrow \mathbf{in}_\le f_i \mid \mathbf{in}_\le f.$$

for some $l \le m$.

$\square$

Lemma 2.9.75. *Let $F = \{f_1, \dots, f_m\} \subset K[x_1, \dots, x_n] \setminus 0$ and set $I = \langle F \rangle$. The following statements hold*

1. *If $f \to_F 0$ for every $f \in I$ then F is a Gröbner basis.*

2. *If F is a Gröbner basis then for $f \in I \setminus 0$,*

$$f^F = 0 \iff f \to_F 0$$

Proof. 1. Let $f \in I \setminus 0$. Then there are $\lambda_1, \dots, \lambda_m \in K[\mathbf{x}]$ such that

$$f = \sum_1^m \lambda_i f_i,$$

and $\text{in}_\leq \lambda_i f_i \leq \text{in}_\leq f$ for every $i \in \{1, \dots, m\}$ with $\lambda_i \neq 0$. Then by the above lemma

$$\kappa := \max \; \{\text{in}_\leq \lambda_i f_i : i \in \{1, \dots, m\}, \lambda_i \neq 0\} = \text{in}_\leq f$$

and by the same lemma we then have that $\text{in}_\leq f_i \mid \text{in}_\leq f$ for some $i \in \{1, \dots, m\}$, hence F is a Gröbner basis.
2. "$\Rightarrow$": If $f^F = 0$ then there are $\lambda_1, \dots, \lambda_i \in K[\mathbf{x}]$ such that

$$f = \sum_1^m \lambda_i f_i$$

and $\text{in}_\leq \lambda_i f_i \leq \text{in}_\leq f$ for $i \in \{1, \dots, m\}$ with $\lambda_i \neq 0$ hence by definition $f \to_F 0$.
"$\Leftarrow$": This follows from Proposition 2.9.69. □

We now introduce S-polynomials

Definition 2.9.76. Let $f, g \in K[x_1, \dots, x_n] \setminus 0$. Pick $w \in \mathbb{N}^n$ such that $\mathbf{x}^w = \text{lcm}(\text{in}_\leq f, \text{in}_\leq g)$. We define *the S-polynomial* or *the syzygy* of f and g to be

$$S(f, g) := \frac{\mathbf{x}^w}{\text{in}_\leq f} f - \frac{\mathbf{x}^w}{\text{in}_\leq g} g.$$

Remark 2.9.77. Recall that $w = \left(\max\left(w_1^f, w_1^g\right), \dots, \max\left(w_n^f, w_n^g\right)\right)$, where $a_{w^f} \mathbf{x}^{w^f} = \text{in}_\leq f$ and $b_{w^g} \mathbf{x}^{w^g} = \text{in}_\leq g$.

Note this simple fact about S-polynomials

Lemma 2.9.78. *Let $f, g \in K[x_1, \dots, x_n] \setminus 0$. Pick $w \in \mathbb{N}^n$ such that $\mathbf{x}^w = \text{lcm}(\text{in}_\leq f, \text{in}_\leq g)$. Then $\text{in}_\leq S(f, g) < \mathbf{x}^w$. In other words, the initial term of $\frac{\mathbf{x}^w}{\text{in}_\leq f} f$ cancels with the initial term of $-\frac{\mathbf{x}^w}{\text{in}_\leq g} g$.*

Proof. Indeed, note that

$$\begin{aligned}\mathrm{in}_{\le}\ \frac{\mathbf{x}^w}{\mathrm{in}_{\le}\ f} f - \mathrm{in}_{\le}\ \frac{\mathbf{x}^w}{\mathrm{in}_{\le}\ g} g &= \left(\mathrm{in}_{\le}\ \frac{\mathbf{x}^w}{\mathrm{in}_{\le}\ f}\right)(\mathrm{in}_{\le}\ f) - \left(\mathrm{in}_{\le}\ \frac{\mathbf{x}^w}{\mathrm{in}_{\le}\ g}\right)(\mathrm{in}_{\le}\ g)\\ &= \frac{\mathbf{x}^w}{\mathrm{in}_{\le}\ f}\mathrm{in}_{\le}\ f - \frac{\mathbf{x}^w}{\mathrm{in}_{\le}\ g}\mathrm{in}_{\le}\ g = \mathbf{x}^w - \mathbf{x}^w = 0\end{aligned}$$

hence the result follows from Lemma 2.9.66 3. □

These polynomials will be make the criterion for checking that a generating set is a Gröbner basis that Lemma 2.9.75 1. provides more practical. To be precise, we can reduce this criterion to just check that a finite set of S-polynomials reduce to zero modulo F.

Lemma 2.9.79. *Let* $F = \{f_1, \dots, f_m\} \subset K[x_1, \dots, x_n] \setminus 0$ *and let* $1 \le l \le m$ *and* $\sum_1^m \lambda_i f_i \in \langle F \rangle$ *with* $b_i \mathbf{x}^{v_i} = \mathrm{in}_{\le} \lambda_i$, $c_i \mathbf{x}^{w_i} = \mathrm{in}_{\le}\ f_i$ *be given such that*

$$b_i c_i x^{v_i + w_i} = \mathrm{in}_{\le}\ \lambda_i f_i = \kappa := \max_{j \in \{1,\dots,m\}} \mathrm{in}_{\le}\ \lambda_j f_i$$

for every $i \in \{1, \dots, l\}$ *and* $\sum_1^l b_i c_i \ne 0$. *Define*

$$\mu_{\lambda_1,\dots,\lambda_m,F} := \sum_1^m (\mathrm{in}_{\le}\ \lambda_i) f_i.$$

Then $\mu_{\lambda_1,\dots,\lambda_m,F} \in I := \langle S(f_1,f_2), S(f_2,f_3), \dots, S(f_{l-1},f_l)\rangle$, *and* $\mathrm{in}_{\le}\ \mu_{\lambda_1,\dots,\lambda_m,F} < \kappa$.

Proof. Put $g_i := \mathbf{x}^{v_i} \frac{f_i}{c_i}$ for $i \in \{1, \dots, l\}$. Then

$$\begin{aligned}\mu_{\lambda_1,\dots,\lambda_m,F} &= \sum_1^l b_i c_i \left(\mathbf{x}^{v_i + w_i} + \dots\right) = \sum_1^l b_i c_i g_i\\ &= \left[\sum_{j=1}^{l-1} \left[\sum_{i=1}^{j} b_i c_i\right](g_j - g_{j+1})\right] + \underbrace{\left[\sum_1^l b_i c_i\right]}_{=0} g_l.\end{aligned}$$

Put $x^{u_{ij}} := \mathrm{lcm}(\mathbf{x}^{w_i}, \mathbf{x}^{w_j})$ for $i,j \in \{1, \dots, l\}$ and note that

$$g_i - g_j = \frac{\mathbf{x}^{v_i}}{c_i} f_i - \frac{\mathbf{x}^{v_j}}{c_j} f_j = \frac{\mathbf{x}^{v_i + w_i}}{c_i \mathbf{x}^{w_i}} f_i - \frac{\mathbf{x}^{v_j + w_j}}{c_j \mathbf{x}^{w_j}} f_j \overset{(*)}{=} \mathbf{x}^{\xi_{ij}} \left(\frac{\mathbf{x}^{u_{ij}}}{\mathrm{in}_{\le}\ f_i} f_i - \frac{\mathbf{x}^{u_{ij}}}{\mathrm{in}_{\le}\ f_j} f_j\right) = \mathbf{x}^{\xi_{ij}} S(f_i, f_j),$$

where we in step $(*)$ use that $u_{ij} < w_i + v_i = w_j + v_j$ to see that

$$v_i + w_i = v_j + w_j \Rightarrow \underbrace{v_i + w_i - u_{ij}}_{\xi_{ij}} = v_j + w_j - u_{ij}.$$

Upon setting $\xi_i := \xi_{i,i+1}$ we find

$$\mu_{\lambda_1,\dots,\lambda_l,F} = \sum_1^{l-1} \mathbf{x}^{\xi_i} S(f_i, f_{i+1}) \in I.$$

Set $u_i = u_{i(i+1)}$. Then additionally we have that

$$\mathrm{in}_{\leq}\ \mu_{\lambda_1,\dots,\lambda_m,F} = \max_{i\in\{1,\dots,l-1\}} x^{\xi_i}\mathrm{in}_{\leq}\ S(f_i,f_{i+1}) < \max_{i\in\{1,\dots,l-1\}} \mathbf{x}^{\xi_i+u_i} = \max_{i\in\{1,\dots,l-1\}} \mathbf{x}^{v_i+w_i} = \kappa.$$

$\square$

Theorem 2.9.80. *Let $F = \{f_1,\dots,f_m\} \subset K[x_1,\dots,x_n]\setminus 0$. If $S(f_i,f_j) \to_F 0$ for every $i,j \in \{1,\dots,m\}$, then $f \to_F 0$ for every $f \in I := \langle F\rangle$ meaning F is a Gröbner basis (cf. Lemma 2.9.75).*

Proof. Let $f = \sum_1^m \lambda_i f_i \in I$. If $\mathrm{in}_{\leq}\ \lambda_i f_i \leq \mathrm{in}_{\leq}\ f$ for every i, we are done.
Suppose this is not the case. We aim to re-express f as an element of I. We do this via a **right-hand side initial term reduction** (this is non-standard terminology), which we will describe now. WLOG we may assume, adopting the notation from Lemma 2.9.79, that

$$b_i c_i x^{v_i+w_i} = \mathrm{in}_{\leq}\ \lambda_i f_i = \kappa.$$

Then we have that $\mathrm{in}_{\leq}\ f < \kappa$ by Lemma 2.9.74, hence necessarily $\sum_1^l b_i c_i = 0$, which implies $\mu_{\lambda_1,\dots,\lambda_m,F} \in \langle S(f_1,f_2), S(f_2,f_3),\dots,S(f_{l-1},f_l)\rangle$ and $\mathrm{in}_{\leq}\ \mu_{\lambda_1,\dots,\lambda_m} < \kappa$. By assumption there are $\psi_1^{(i)},\dots,\psi_m^{(i)} \in K[\mathbf{x}]$ such that $S(f_i,f_{i+1}) = \sum_{j=1}^m \psi_j^{(i)} f_j$ with $\mathrm{in}_{\leq}\ \psi_j^{(i)} f_j \leq \mathrm{in}_{\leq}\ S(f_i,f_{i+1})$ for every $i \in \{1,\dots,l-1\}$ and $j \in \{1,\dots,m\}$. This means that

$$\mu_{\lambda_1,\dots,\lambda_m,F} = \sum_{j=1}^m \underbrace{\left[\sum_{i=1}^{l-1} \mathbf{x}^{\xi_i}\psi_j^{(i)}\right]}_{\chi_j} f_j$$

with

$$\mathrm{in}_{\leq}\ \chi_j f_j = \max_{i\in\{1,\dots,l-1\}} \mathbf{x}^{\xi_i}\left(\mathrm{in}_{\leq}\ \psi_j^{(i)}\right)\left(\mathrm{in}_{\leq}\ f_j\right) \leq \max_{i\in\{1,\dots,l-1\}} \mathbf{x}^{\xi_i}\mathrm{in}_{\leq}\ S(f_i,f_{i+1}) = \mathrm{in}_{\leq}\ \mu_{\lambda_1,\dots,\lambda_m,F} < \kappa.$$

Now note that

$$f = \mu_{\lambda_1,\dots,\lambda_m,F} + \sum_1^l (\lambda_i - \mathrm{in}_{\leq}\ \lambda_i) f_i + \sum_{l+1}^m \lambda_i f_i,$$

and that every term on the right-hand side of the above expression is strictly smaller then κ. We now obtain another expression for f: Upon putting $\delta_j = 1$ if $j \leq l$ and $\delta_j = 0$ otherwise we have

$$f = \sum_{j=1}^m \underbrace{\left(\chi_j + \lambda_j - \delta_j \mathrm{in}_{\leq}\ \lambda_j\right)}_{\lambda_j'} f_j.$$

This is exactly the right-hand side initial term reduction we wanted to describe. If

$$\mathbf{in}_{\le} f = \kappa' := \max_{i \in \{1,\ldots,m\}} \lambda'_i f_i,$$

we have $\mathbf{in}_{\le} \lambda'_i f_i \le \mathbf{in}_{\le} f$ for every i and we are done. Otherwise we perform another right-hand side initial term reduction. Note that $\kappa' < \kappa$. If we follow the algorithm of terminating if the right-hand side expression leads to concluding $f \to_F 0$ or otherwise performing a right-hand side reduction we see that by the well-ordering of term orders we can only perform a finite number of iterations of right-hand side reductions, hence this algorithm will have to terminate. We thus conclude that $f \to_F 0$. □

from this theorem we readily collect Buchberger's criterion for checking that a generating set is a Gröbner basis

Corollary 2.9.81. *(Buchberger's S-criterion) Let $F \subset K[x_1,\ldots,x_n] \setminus 0$. Then F is a Gröbner basis if and only if $S(f_i,f_j) \to_F 0$ or equivalently $S(f_i,f_j)^F = 0$ (cf. 2.9.75) for every $i,j \in \{1,\ldots,m\}$.*

Proof. If F is a Gröbner basis then $S(f_i,f_j)^F = 0$ for every $i,j \in \{1,\ldots,m\}$ since $S(f_i,f_j) \in I := \langle F \rangle$ by Proposition 2.9.69, hence $S(f_i,f_j) \to_F 0$ by Lemma 2.9.75. If conversely $S(f_i,f_j) \to_F 0$ for every i and j, it follows from Theorem 2.9.80 that F is a Gröbner basis and hence that $S(f_i,f_j)^F = 0$ by Proposition 2.9.69. □

This leads to Buchberger's algorithm for finding a Gröbner basis for an ideal $I = \langle f_1,\ldots,f_m \rangle \subset K[x_1,\ldots,x_n] \setminus 0$, which we will discuss in the following remark

Remark 2.9.82. (Buchberger's algorithm) We now describe an algorithm for computing a Gröbner basis given an arbitrary generating set for an ideal I. Let $F_0 = \left\{f_1^{(0)},\ldots,f_{m(0)}^{(0)}\right\} \subset I$ where $m(0) \ge 1$ be a generating set for I. For $i \ge 0$ if $S\left((f_j^{(i)},f_k^{(i)}\right)^{F_i} = 0$ for every $f_j^{(i)},f_k^{(i)} \in F_i = \left\{f_1^{(i)},\ldots,f_{m(i)}^{(i)}\right\}$ put $F_{i+1} = F_i$ or simply terminate, for then F_i is a Gröbner basis. Otherwise if there are $f_j^{(i)},f_k^{(i)} \in F_i$ such that $S\left(f_j^{(i)},f_k^{(i)}\right)^{F_i} \neq 0$ put

$$F_{i+1} = F_i \cup \left\{S\left(f_j^{(i)},f_k^{(i)}\right)^{F_i}\right\}.$$

Note that $S\left(f_j^{(i)},f_k^{(i)}\right)^F = \sum_1^m \lambda_i f_i - S\left(f_j^{(i)},f_k^{(i)}\right) \in I$, hence $\langle F_{i+1} \rangle = I$. The claim is that the ascending chain $F_0 \subset F_1 \subset \ldots$ will in fact stabilize, hence there we produce a Gröbner basis at some point. We check this claim in the next theorem.

Lemma 2.9.83. *Let $\{t_i\}_{i \ge 0} \subset K[x_1,\ldots,x_n]$ be some sequence of which an element is either a term or 0, i.e. $t_i = a_i \mathbf{x}^{v_i}$ for some $a_i \in K$, $v_i \in \mathbb{N}^n$. Then for some $N \ge 0$ for every $i \ge N$, $t_j \mid t_i$ for some $j < N$*

Proof. If $a_i = 0$ for every $i \geq 0$ then the statement is trivial. Suppose this is not the case and put $S = \{v_i : i \geq 0, a_i \neq 0\} \subset \mathbb{N}^n$, then by Dickson's lemma there are $v_{i(1)}, \ldots, v_{i(k)} \in S$ such that

$$S \subset \bigcup_{j=1}^{k} \left(v_{i(j)} + \mathbb{N}^n\right).$$

Set $N = \max_{j \in \{1,\ldots,k\}} i(j)$ and let $i \geq N$. Then $v_i = v_{i(j)} + w$ for some $j \in \{1,\ldots,k\}$, $w \in \mathbb{N}^n$, hence

$$t_i = a_i \mathbf{x}^{v_i} = a_i \mathbf{x}^{v_{i(j)}+w} = \left(a_{i(j)} \mathbf{x}^{v_{i(j)}}\right)\left(\frac{a_i}{a_{i(j)}} \mathbf{x}^w\right) = t_{i(j)} \left(\frac{a_i}{a_{i(j)}} \mathbf{x}^w\right) \Rightarrow t_{i(j)} \mid t_i$$

□

Theorem 2.9.84. *Buchberger's algorithm terminates and outputs a Gröbner basis.*

Proof. Buchberger's gives rise to an infinite sequence of polynomials in the following way: start with the initial elements of $F_0 = \{f_1, \ldots, f_m\}$. For $i \geq m+1$ if F_{i-m} is the union of F_{i-m-1} and $\left\{S(f_j, f_k)^{F_{i-m-1}}\right\}$ for some $j,k \in \{1,\ldots,i-1\}$ then put $f_i = S(f_j,f_k)^{F_{i-m-1}}$ otherwise put $f_i = 0$. We then put $t_i = \mathrm{in}_{\leq} f_i$ if $f_i \neq 0$ or $t_i = 0$ otherwise for every $i \geq 0$. By the above lemma there is an $N \geq 0$ such that for every $l \geq N$, $t_h \mid t_l$ for some $h < N$. For each $i \geq m$, if $f_i = S(f_j,f_k)^{F_{i-m-1}}$, then any term of f_i is not divisible by $\mathrm{in}_{\leq} f_q$ for any $q \in \{1,\ldots,i-1\}$. Consider then $l \geq \max(m,N)$ if $t_h \mid t_l$ for $h < N$, then t_l cannot be a term of some $S(f_j,f_k)^{F_{l-m-1}}$, hence $t_l = 0$, implying $f_l = 0$ and hence that F_{l-m} satisfies Buchberger's criterion. □

The below proposition will give an easy criterion for checking whether two polynomials in a generating reduce modulo said generating set.

Proposition 2.9.85. *Let* $\leq$ *be a term order on* $K[x_1,\ldots,x_n]$*. Let* $f,g \in K[\mathbf{x}] \setminus 0$*. Suppose* $\gcd(f,g) = 1$*, then* $S(f,g) \to_{\{f,g\}} 0$*.*

Proof. By assumption,

$$\mathrm{lcm}(\mathrm{in}_{\leq} f, \mathrm{in}_{\leq} g) = (\mathrm{in}_{\leq} f)(\mathrm{in}_{\leq} g) = \mathrm{in}_{\leq} fg.$$

Put $r = f - \mathrm{in}_{\leq} f$ and $s = g - \mathrm{in}_{\leq} g$. Then $\mathrm{in}_{\leq} r < \mathrm{in}_{\leq} f$ or $\mathrm{in}_{\leq} r = 0$ and $\mathrm{in}_{\leq} s < \mathrm{in}_{\leq} g$ or $\mathrm{in}_{\leq} s = 0$. Then

$$S(f,g) = (\mathrm{in}_{\leq} g)f - (\mathrm{in}_{\leq} f)g = (g-s)f - (f-r)g = rg - sf.$$

Suppose $r = 0$, then $S(f,g) = -sf$, implying $\mathrm{in}_{\leq} f \leq \mathrm{in}_{\leq} S(f,g)$ and hence $S(f,g) \to_{\{f,g\}} 0$. Suppose $\mathrm{in}_{\leq} r < \mathrm{in}_{\leq} f$. Suppose for a contradiction

$$(\mathrm{in}_{\leq} r)(\mathrm{in}_{\leq} g) = (\mathrm{in}_{\leq} s)(\mathrm{in}_{\leq} f).$$

Then $\mathrm{in}_\leq f \mid \mathrm{in}_\leq r$ (since $\mathrm{in}_\leq g \nmid \mathrm{in}_\leq f$), but then $\mathrm{in}_\leq f \leq \mathrm{in}_\leq r$ leading to a contradiction. We then find that

$$\mathrm{in}_\leq S(f,g) = \mathrm{in}_\leq (rg - sf) = \max(\mathrm{in}_\leq rg, \mathrm{in}_\leq sf)$$

hence by Lemma 2.9.74, $S(f,g) \to_{\{f,g\}} 0$. □

Definition 2.9.86. A Gröbner basis $G = \{f_1,\ldots,f_m\} \subset K[x_1,\ldots,x_n] \setminus 0$ is called *minimal*, if

1. $\mathrm{in}_\leq f_i \nmid \mathrm{in}_\leq f_j$ for every $i,j \in \{1,\ldots,m\}$ with $i \neq j$.

2. $\mathrm{in}_\leq f_i = \mathbf{x}^{v_i}$ for some $v_i \in \mathbb{N}^n$ for every $i \in \{1,\ldots,m\}$.

G is *reduced* if it is minimal and if every term in f_i is not divisible by $\mathrm{in}_\leq f_j$ for every $i,j \in \{1,\ldots,m\}$ with $i \neq j$.

Remark 2.9.87. We describe an algorithm for computing a minimal Gröbner basis for a non-zero ideal $I \subset K[x_1,\ldots,x_n]$. Let $G = \{f_1,\ldots,f_m\} \subset I \setminus 0$ be a Gröbner basis. Let a_i be the leading coefficient for f_i for each i. Then define

$$G_0 := \{a_1^{-1} f_1, \ldots, a_m^{-1} f_m\}$$

For $k \geq 0$, if there are some $f,g \in G_k \setminus 0$ with $f \neq g$ such that that $\mathrm{in}_\leq g \mid \mathrm{in}_\leq f$, define $G_{k+1} := G_k \setminus \{f\}$, otherwise terminate.

Lemma 2.9.88. *Every ideal $0 \neq I \subset K[x_1,\ldots,x_n]$ has a minimal Gröbner basis.*

Proof. We prove that the algorithm described above always produces a minimal Gröbner basis. Let $G = \{f_1,\ldots,f_m\} \subset I \setminus 0$ be a Gröbner basis for I. We prove the statement by induction in m. For $m = 1$, $G = \{g\}$ for some $g \in I \setminus 0$. Let $a \in K \setminus 0$ be the leading coefficient of g. Then $G_0 = \{a^{-1} g\}$ defines a minimal Gröbner basis. Suppose the algorithm always terminates with a minimal Gröbner basis with mm elements for some $m \geq 1$. Let $G = \{f_1,\ldots,f_{m+1}\}$ be a Gröbner basis and assume WLOG that the coefficient of the polynomials in G are all 1, i.e. that $G_0 = G$. Then if there are no $i,j \in \{1,\ldots,m+1\}$ with $i \neq j$ such that $\mathrm{in}_\leq f_j \mid \mathrm{in}_\leq f_i$, G_0 we terminate and indeed G_0 is a minimal Gröbner basis. Otherwise we put $G_1 = G_0 \setminus \{f_i\}$. Let $f \in I$, then for some $l \in \{1,\ldots,m+1\}$, $\mathrm{in}_\leq f_l \mid \mathrm{in}_\leq f$. If $l = i$, then $\mathrm{in}_\leq f_j \mid \mathrm{in}_\leq f$, hence in any case the initial term of f is divisible by the initial term of some polynomial in G_1, implying G_1 is a Gröbner basis. Since G_1 has m elements it follows by induction that G_k is a minimal Gröbner basis. □

Proposition 2.9.89. *Every ideal $0 \neq I \subset K[x_1,\ldots,x_n]$ has a unique reduced Gröbner basis.*

Proof. **Uniqueness:** Consider two reduced Gröbner bases $G = \{f_1,\ldots,f_m\}$ and $G' = \{f'_1,\ldots,f'_{m'}\}$. We first check that the cardinality of these Gröbner bases match. Let $i \in \{1,\ldots,m\}$, then for some $\tau(i) \in \{1,\ldots,m'\}$, $\mathrm{in}_\leq f'_{\tau(i)} \mid \mathrm{in}_\leq f_i$. Let $j \in \{1,\ldots,m'\}$ then for some $\omega(j) \in \{1,\ldots,m\}$, $\mathrm{in}_\leq f_{\omega(j)}$ $\mathrm{in}_\leq f'_j$. Then we have that

$$\mathrm{in}_\leq f_{\omega(\tau(i))} \mid \mathrm{in}_\leq f'_{\tau(i)} \text{ and } \mathrm{in}_\leq f'_{\tau(i)} \mid \mathrm{in}_\leq f_i \Rightarrow \mathrm{in}_\leq f_{\omega(\tau(i))} \mid \mathrm{in}_\leq f_i$$

by minimality of the Gröbner bases $i = \omega(\tau(i))$. A similar argument shows that $\tau(\omega(j))$ for every $j \in \{1,\ldots,m'\}$, thus τ is a bijection, implying $m = m'$. We proceed by checking that the sets are equal. Note that the above argument also shows that $\mathrm{in}_\leq f'_{\tau(i)} = \mathrm{in}_\leq f_i$ for every $i \in \{1,\ldots,m\}$ since the coefficient of every initial term is 1. Let $i \in \{1,\ldots,m\}$. Since $\mathrm{in}_\leq f_i = \mathrm{in}_\leq f'_{\tau(i)}$, either $f_i = f_{\tau(i)}$ or $\mathrm{in}_\leq \left(f_i - f'_{\tau(i)}\right) < \mathrm{in}_\leq f_i$. We shall that the second case implies $f_i = f_{\tau(i)}$ as well. In this case no term in $f_i - f'_{\tau(i)}$ is divisible by $\mathrm{in}_\leq f_i$. Any term in $f_i - f'_{\tau(i)}$ is a term in $f'_{\tau(i)}$ subtracted from f_i, where at least one of these terms in non-zero. Then by the Gröbner bases being reduced, $\mathrm{in}_\leq f_j$ does not divide such a term for any $j \in \{1,\ldots,m\} \setminus \{i\}$. Then $f_i - f'_{\tau(i)} = \left(f_i - f'_{\tau(i)}\right)^G$, but since $f_i - g_i \in I$, this must imply that $f_i = f'_{\tau(i)}$ by Proposition 2.9.69.
Existence: By the prior lemma there is a minimal for Gröbner basis $G = \{f_1,\ldots,f_m\}$ for I. Define $g_i := f_i^{\{g_1,\ldots,g_{i-1},f_i,\ldots,f_m\}\setminus\{f_i\}}$ for every $i \in \{1,\ldots,m\}$. Since $\mathrm{in}_\leq f_i$ is divisible by any $\mathrm{in}_\leq f_j$ for $j \in \{i+1,\ldots,m-1\}$ we see that g_i is of the form $\mathrm{in}_\leq f_i + \ldots$. Thus if $f \in I$, there is some g_j such that $\mathrm{in}_\leq g_j = \mathrm{in}_\leq f_j \mid \mathrm{in}_\leq f$, meaning that each set $\{g_1,\ldots,g_{\{}k-1\}, f_k,\ldots,f_m\}$ and in particular $G' := \{g_1,\ldots,g_m\}$ is a Gröbner basis. The g_i's being residues following the division algorithm by a set of polynomials with initial terms coming from $\{f_1,\ldots,f_m\} \setminus \{f_i\}$ implies that no term in g_i is divisible by any $\mathrm{in}_\leq f_j$ for $j \neq i$, thus G' is a reduced Gröbner basis. □

Remark 2.9.90. Note that the existence proof above is of an algorithmic nature. I.e. given an ideal, use Buchberger's algorithm to produce a Gröbner basis, then use the already presented algorithm for producing a minimal gröbner basis, then apply the division algorithmic in the way we described above to produce the elements of the reduced Gröbner basis.

Theorem 2.9.91. *Let G be a Gröbner basis for an ideal $I \subset K[x_1,\ldots,x_n]$ with respect to the lexicographic term order with $x_1 < \cdots < x_n$. Then $G \cap K[x_1,\ldots,x_i] \subset K[x_1,\ldots,x_i]$ is a Gröbner basis for the ideal $I \cap K[x_1,\ldots,x_i] \subset K[x_1,\ldots,x_i]$ with respect to the lexicographic term order with $x_1 < \cdots < x_i$ for every $i \in \{1,\ldots,i\}$.*

Proof. Let $G' = G \cap K[x_1, \dots, x_i]$. Let $f \in I' = I \cap K[x_1, \dots, i]$. For some $g \in G$, $\text{in}_{\leq} g \mid \text{in}_{\leq} f$. Let $t = a\mathbf{x}^v$ be a term of g and write $b\mathbf{x}^w = \text{in}_{\leq} f$. Then $a\mathbf{x}^v \leq b\mathbf{x}^w$, or in other words $v \leq w$. Let $u \in \mathbb{N}^n$. If $u_j \neq 0$ for $j \in \{i+1, \dots, m\}$, then $w <_{\text{lex}} u$. Hence we conclude that $v_j = 0$ for every $j \in \{i+1, \dots, m\}$, implying $t \in K[x_1, \dots, x_i]$ and ultimately that $g \in K[x_1, \dots, x_i]$. □

The above theorem is a great tool for computing solutions to complicated polynomial equations.

2.9.12 Polynomials over UFD's

We aim to prove that polynomials over unique factorization domains are unique factorization domains. For this reason we fix a UFD R (unless something else is explicitly stated).

Lemma 2.9.92. *Let R be an integral domain. If $p \in R$ is prime then $p \in R[x_1]$ is prime. Therefor if $p \in R$ is prime, then $p \in R[x_1, \dots, x_n]$ is prime.*

Proof. $p \neq 0$ and $p \notin R[x]^* = R^*$. Let $f = \sum_0^k a_i x^i, g = \sum_0^h b_i x^i \in R[x]$ such that $p \nmid f, g$. Set $s := \max\ \{i \in \{1, \dots, k\} : p \nmid a_i\}$ and $t := \max\ \{i \in \{1, \dots, h\} : o \nmid b_i\}$. Note that if $i > s$ or $j > r$, then $p \mid a_i$ or $p \mid b_j$. This means that since $i + j = s + t$ implies $i \geq s$ or $j \geq t$, one finds that p divides every term in $\sum_{i=1}^h \sum_{j=1}^k \sum_{i+j=s+t} a_i b_j$ other than $a_s b_t$, hence

$$p \nmid \sum_{i=1,\dots,h, j=1,\dots,k i+j,=s+t} a_i b_j \Rightarrow p \nmid fg.$$

□

Definition 2.9.93. A polynomial $f = \sum_0^n a_i x^i \in R[x]$ is said to be *primitive* if $\gcd(a_0, \dots, a_n) = 1$.

Lemma 2.9.94. *If $f, g \in R[x]$ is primitive, then fg is primitive. This property extends to multivariable polynomials by induction.*

Proof. Suppose fg is not primitive then the greatest common divisor of the coefficients of fg is divisible by some prime $p \in R$. This means $p \mid fg$, hence $p \mid f$ or $p \mid g$ by Lemma 2.9.92, hence the p divides all of the coefficients f or g, hence f primitive or g is primitive. □

Lemma 2.9.95. *Let $f, g \in R[x]$. If f is primitive and $f \mid g$ in $Q(R)[x]$, then $f \mid g$ in $R[x]$.*

Proof. By assumption we can find an $h \in Q(R)[x]$ such that $g = hf$. If $h = 0$, then $g = 0$ and we are done. Suppose $h \neq 0$. Then for some $c \in R \setminus 0$, $ch \in R[x]$. For some $d \in R \setminus 0$ and primitive $h' \in R[x]$, $ch = dh'$, implying $cg = chf = dh'f$. Note $h'f$ is primitive by the prior lemma, hence d is the greatest common divisor for the coefficients of $dh'f$. This implies d is the greatest common divisor of the coefficients of cg. Since $c \mid cg$, it follows that $c \mid d$, hence $\frac{d}{c} \in R$. This implies that

$$g = \frac{d}{c} h' f \in R[x],$$

which means $f \mid g$ in $R[x]$. □

In the following lemma we classify all the irreducible polynomials in $R[x]$.

Lemma 2.9.96. *1. Let $f \in R[x]$ be primitive. If f is irreducible in $Q(R)[x]$, then f is prime in $R[x]$.*

2. Any non-zero, non-unit element in $R[x]$ can be written as a product of irreducible elements.

3. The irreducible elements in $R[x]$ are the primes in R and the polynomials described in 1.

Proof. 1. Let $a, b \in R[x]$ such that $f \mid ab$ in $R[x]$. Using Result that shows $K[x]$ is UFD we get that f is prime by Proposition 2.8.46 in $K[x]$, hence $f \mid a$ or $f \mid b$ in $K[x]$, hence by Lemma 2.9.95 $f \mid a$ or $f \mid b$ in $R[x]$.
2. Let $f \in R[x]$ be a non-zero, non-unit element. If $\deg f = 0$ it is an element in R, which is a UFD, hence f has a factorization into irreducibles. If $\deg f > 0$, then there primes/irreducibles $g_1, \ldots, g_m \in Q(R)[x]$ such that $f = \prod_1^m g_i$. For suitable $a_1, \ldots, a_m \in K$ and primitive $f_1, \ldots, f_m \in R[x]$ such that $g_i = a_i f_i$ for each $i \in \{1, \ldots, m\}$. Since g_i is irreducible in $Q(R)[x]$ for each i, so is f_i in $Q(R)[x]$. Set $a := \prod_1^m a_i$. Then $f = a \prod_1^m f_i$, hence $\prod_1^m f_i \mid f$ in $Q(R)[x]$, and hence also in $R[x]$ by Lemma 2.9.95. This means $a \in R$. If a is a unit, set $f_1' = af_1$. Then we get a factorization into irreducibles,

$$f = f_1' \prod_2^m f_i.$$

If a is not a unit, we write $a = \prod_1^l p_i$ for primes/irreducibles in R, getting a factorization intio irreducibles

$$f = \left(\prod_1^l p_i \right) \left(\prod_1^m f_i \right).$$

3. We have already established that the two types of elements in question are irreducible in $R[x]$ (cf. 1 and Lemma 2.9.92). Let f be irreducible in $R[x]$. If $\deg f = 0$, then $f \in R$, hence f is prime as R is a UFD. If $\deg f > 0$, then we saw in 2. that f can be written as a product of primes in R and primitive polynomials in $R[x]$ irreducible in $Q(R)[x]$. Writing $f = af_1 \cdots f_m$ for $f_1, \ldots, f_m$ polynomials being of the second type of irreducibles. We see that $m = 1$ and a is a unit for otherwise this would contradict the irreducibility of f. ☐

Theorem 2.9.97. *$R[x]$ is a UFD. By induction $R[x_1, \ldots, x_n]$ is a UFD.*

Proof. Every non-zero non-unit element is a product irreducible elements in $R[x]$ by Lemma 2.9.96 2. Let f be an irreducible element in $R[x]$. Then either f is a prime in R or is primitive in $R[x]$ and irreducible in $Q(R)[x]$ by Lemma 2.9.96 3. In the second case f is also prime by Lemma 2.9.96 1. It follows by Proposition 2.8.46 that $R[x]$ is a UFD. Since $R[x_1, \ldots, x_n] \simeq (R[x_1, \ldots, x_{n-1}])[x_n]$ it follows by induction that $R[x_1, \ldots, x_n]$ is a UFD. ☐

Proposition 2.9.98. *Let $f \in R[x]$ such that f is monic and $\deg f \in \{2, 3\}$. Then f is irreducible if and only if f has not roots.*

Proof. "$\Rightarrow$": Suppose f has a root $\alpha \in R$. Then $f \in \langle x - \alpha \rangle$, hence $f = (x - \alpha)g$ for some $g \in R[x]$. In either case of $\deg f = 2$ or $\deg f = 3$, we have that $\deg g \geq 1$, implying g is not a unit. This means f is reducible.
"$\Leftarrow$": Suppose f is reducible. Since the irreducible non-constant polynomials in $R[x]$ are monic, there is, In any case of f having degree 2 or 3, a polynomial $g = (x - \alpha) \in R[x]$, such that

$$f = gh,$$

for some $h \in R[x]$. It hence follows that f has a root α. ☐

Corollary 2.9.99. *The polynomial $x^2 - a \in R[x]$ is irreducible if and only if a is not a square.*

2.9.13 Eisenstein's Criterion

2.9.14 Homogeneous Polynomials

Definition 2.9.100. A polynomial $f \in R[x_1, \ldots, x_n]$ is *homogeneous of degree d*, if there is a $d \geq 0$ such that $f = \sum_{v \in \mathbb{N}^n : |v| = d} a_v \mathbf{x}^v$.

Remark 2.9.101. Note that if $f \neq 0$ then f is homogeneous if and only if every non-zero term is equal to d, hence $\deg f = d$. Note also that 0 is homogeneous of degree d for every $d \geq 0$.
One readily verifies that the set of degree d homogeneous polynomials in $R[\mathbf{x}]$ is an R-module, which we will denote $V_R(d,n)$. The set $\{\mathbf{x}^v \in R[x_1,\ldots,x_n] : |v| = d\}$ forms a basis for $V_R(d,n)$.

Lemma 2.9.102. *Let $f,g \in R[x_1,\ldots,x_n]$.*

1. *If f,g are homogeneous of degree respectively d and e, then fg is homogeneous of degree $d+e$. Suppose R is an integral domain and $f,g \neq 0$. Then if fg is homogeneous, so is f and g.*

2. *If f,g are homogeneous of degree d, so is $f+g$.*

Proof. 1. Write $f = \sum_{v \in \mathbb{N}^n : |v| = d} a_v \mathbf{x}^v$ and $g = \sum_{w \in \mathbb{N}^n : |w| = e} b_w \mathbf{x}^w$. Then

$$fg = \sum_{v \in \mathbb{N}^n : |v| = d} \sum_{w \in \mathbb{N}^n : |w| = e} a_v b_w \mathbf{x}^{v+w} = \sum_{u \in \mathbb{N}^n : |u| = d+e} \left[\sum_{v,w \in \mathbb{N}^n : v+w = u} a_v b_w \right] \mathbf{x}^u.$$

Suppose f is not homogeneous. Since f is not homogeneous the set $\{l \geq 0 :$ there is a $u \in \mathbb{N}^n$ with $|u| =$ has at least 2 elements. Let m be the minimum of this set and M the maximum. Note that $m \neq M$. Pick $u_m, u_M \in \mathbb{N}^n$ such that $|u_m| = m$ and $|u_M| = M$. Pick $b_{q_k}\mathbf{x}^{q_k}$, $b_{q_K}\mathbf{x}^{q_K}$ to be respectively a lowest degree term and a highest degree term of g. Then $\sum_{v,w \in \mathbb{N}^n : v+w = u_k + u_m} a_v b_w = a_{u_m} b_{u_k} \neq 0$ and $\sum_{v,w \in \mathbb{N}^n : v+w = u_K + u_M} a_v b_w = a_{u_M} b_{u_K} \neq 0$. fg therefor has two non-zero monomial terms of different degree and therefor is not homogeneous.
2. We get that

$$f + g = \sum_{v \in \mathbb{N}^n : |v| = d} a_v \mathbf{x}^v + \sum_{v \in \mathbb{N}^n : |v| = d} b_v \mathbf{x}^v = \sum_{v \in \mathbb{N}^n : |v| = d} (a_v + b_v) \mathbf{x}^v.$$

□

Definition 2.9.103. Let $f \in R[x_1,\ldots,x_n]$. We define *the dehomogenization of f at x_i* to be the polynomial

$$f_{*,i} := f(x_1,\ldots,x_{i-1},1,x_{i+1},\ldots,x_n).$$

Remark 2.9.104. We define $f_* := f_{*,n}$

Lemma 2.9.105. *Let $f \in R[x_1,\dots,x_n]$. There are unique homogeneous polynomials $f_0,\dots,f_d \in R[\mathbf{x}]$ where* $\deg\ f_i = i$ *for $f_i \neq 0$ such that*

$$f = \sum_1^d f_i.$$

If $f \neq 0$, then $f_d \neq 0$

Proof. If $f = 0$ the statement is trivial. So suppose $f \neq 0$. Set $d = \deg\ f$ and write $f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v$. Set $f_i = \sum_{v\in\mathbb{N}^n:|v|=i} a_v \mathbf{x}^v$. Then clearly $f = \sum_1^d f_i$. For some $v \in \mathbb{N}^n$ with $|v| = d$, $a_v \neq 0$, hence $f_d \neq 0$. Uniqueness follows from uniqueness of the monomial representation of a polynomial. □

Corollary 2.9.106. $R_d[x_1,\dots,x_n] = \sum_0^d V_R(n,i)$. *This sum is direct.*

Definition 2.9.107. Let $f \in R[x_1,\dots,x_n]$, write $f = \sum_0^d f_i$ (cf. the above lemma). Then *the homogenization of f* is the polynomial

$$f^* := \sum_0^d x_{n+1}^{d-i} f_i$$

Remark 2.9.108. An alternative definition is that $f^* = x_{n+1}^d f\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right)$ (in $Q(R[x_{n+1}])[x_1,\dots,x_n]$ for instance). Indeed,

$$\begin{aligned}
x_{n+1}^d f\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right) &= x_{n+1}^d \sum_{v\in\mathbb{N}^n} a_v \frac{x_1^{v_1}}{x_{n+1}^{v_1}} \cdots \frac{x_n^{v_n}}{x_{n+1}^{v_n}} \\
&= x_{n+1}^d \sum_0^d \sum_{v\in\mathbb{N}^n:|v|=i} a_v \frac{x_1^{v_1}\cdots x_n^{v_n}}{x_{n+1}^{v_1+\dots+v_n}} \\
&= \sum_0^d \sum_{v\in\mathbb{N}^n:|v|=i} x_{n+1}^{d-i} a_v \mathbf{x}^v = \sum_0^d x_{n+1}^{d-i} f_i.
\end{aligned}$$

Note also $x_{n+1}^{d-i} f_i$ is homogeneous of degree $d-i+i = d$, hence $f^* = x_{n+1}^d f_0 + x_{n+1}^{d-1} f_1 + \dots + f_d$ is homogeneous of degree d. Note that $f^* = 0$ if and only if $f = 0$. Indeed if $f \neq 0$, Then $x_{n+1}^{d-i} f_i \neq 0$, furthermore any monomial in $x_{n+1}^{d-i} f_i$ is different from any monomial in $x_{n+1}^{d-j} f_j$ since their x_{n+1}-degree is $d-i$ resp. $d-j$.

We observe the following facts about homogenization and de-homogenization.

Proposition 2.9.109. *Let $f,g \in R[x_1,\dots,x_n]$ and $F \in R[x_1,\dots,x_{n+1}]$ be homogeneous.*

1. $x_{n+1}^{\deg\ f + \deg\ g - \deg(f+g)}(f+g)^* = x_{n+1}^{\deg\ g} f^* + x_{n+1}^{\deg\ g} g^*$. *Suppose additionally that R is an integral domain. Then $(fg)^* = f^* g^*$.*

2. $(fg)_{*,i} = f_{*,i} g_{*,i}$ *&* $(f+g)_{*,i} = F_{*,i} + G_{*,i}$.

3. $(f^*)_* = f$.

4. Let $r := \max\left(\left\{j \geq 0 : x_{n+1}^j \mid F\right\}\right)$. Then $x_{n+1}^r (F_*)^* = F$

Proof. 1. For the first identity one finds that

$$
\begin{aligned}
x_{n+1}^{\deg\ f+\deg\ g-\deg(f+g)}(f+g)^* &= x_{n+1}^{\deg\ f+\deg\ g-\deg(f+g)+\deg(f+g)}\left[f\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right)+g\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right)\right]\\
&= x_{n+1}^{\deg\ g}\left[x_{n+1}^{\deg\ f} f\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right)\right]+x_{n+1}^{\deg\ f}\left[x_{n+1}^{\deg\ g} g\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right)\right]\\
&= x_{n+1}^{\deg\ g} f^* + x_{n+1}^{\deg f} g^*.
\end{aligned}
$$

For the second see that

$$
\begin{aligned}
(fg)^* &= x_{n+1}^{\deg\ f+\deg\ g} f\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right) g\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right)\\
&= x_{n+1}^{\deg\ f} f\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right) x_{n+1}^{\deg\ g} g\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right) = f^* g^*.
\end{aligned}
$$

2. This follows from evaluation being a ring homomorphism.

3. Indeed

$$
(f^*)_* = \left(\sum_0^d x_{n+1}^{d-i} f_i\right)_* = \sum_0^d f_i = f.
$$

4. Write $F = \sum_{v\in\mathbb{N}^{n+1}} a_v \mathbf{x}^v$. Note that if $x_{n+1}^j \mid F$, then $F = x_{n+1}^j Q$ for some $Q \in R[x_1,\dots,x_{n+1}]$. By Lemma 2.9.102 1. Q is homogeneous of degree $d-j$ where $d := \deg\ F$. Then x_{n+1}^j divides every term F. Set

$$
d' = \deg\ F(x_1,\dots,x_n,1) = \max_{v\in\mathbb{N}^{n+1}:a_v\neq 0} \sum_1^n v_i = \max_{v\in\mathbb{N}^{n+1}:a_v\neq 0} d - v_{n+1}.
$$

For some $w \in \mathbb{N}^{n+1}$ with $a_w \neq 0$, $w_{n+1} = r$. Then $d' = d - w_{n+1} = d - r$, since if there were a $u \in \mathbb{N}^{n+1}$ with $a_u \neq 0$ and $u_{n+1} < r$ then $x_{n+1}^r \nmid a_u x_{n+1}^{u_{n+1}}$. As F is homogeneous $d = \sum_1^{n+1} v_i$ for every $v \in \mathbb{N}^n$ with $a_v \neq 0$, hence $d' = -r + \sum_1^{n+1} v_i$. We get

$$
\begin{aligned}
x_{n+1}^r (F_*)^* &= \sum_{v\in\mathbb{N}^n} a_v x_1^{v_1}\cdots x_n^{v_n} x_{n+1}^{d'+r-\sum_0^n v_i}\\
&= \sum_{v\in\mathbb{N}^n} a_v x_1^{v_1}\cdots x_n^{v_n} x_{n+1}^{\sum_1^{n+1} v_i - r + r - \sum_0^n v_i}\\
&= \sum_{v\in\mathbb{N}^n} a_v x_1^{v_1}\cdots x_n^{v_n} x_{n+1}^{v_{n+1}} = F.
\end{aligned}
$$

□

Corollary 2.9.110. *Suppose R is an integral and consider $F \in R[x_1,\dots,x_{n+1}]$ a homogeneous polynomial. Then a factorization of F determines a factorization of F_* up to a factor x_{n+1}^r. If $R = K$ an algebraically closed field and $F \in K[x,y]$ then F factors into a product of linear factors.*

Proof. Indeed, $F = x_{n+1}^r Q$ for some homogeneous $Q \in R[x_1,\dots,x_n,x_{n+1}]$. Then suppose $Q = q_1 \cdots q_l$ for some $q_1,\dots,q_l \in R[x_1,\dots,x_n,x_{n+1}]$. Then $F_* = q_{1*} \cdots q_{l*}$. Conversely, if $F_* = q_1 \cdots q_l$, then $F = x_{n+1}^* q_1^* \cdots q_l^*$.
For the second statement we can again write $F = y^r Q$ for some homogeneous $Q \in K[x,y]$. Then $Q_* = a\prod_1^l (x-a_i)^{r_i}$ for some $a,a_1,\dots,a_l \in K$ where $a \neq 0$. Then $F = y^r(F_*)^* = ay^r\prod_1^l (x-a_iy)^{r_i}$. □

Proposition 2.9.111. *Let R be an integral domain. Consider $f,g \in R[x_1,\dots,x_n]$ homogeneous of degree d respectively degree $d+1$ with $\gcd(f,g)=1$. Then $f+g$ is irreducible.*

Proof. We proof that if $f+g$ is reducible, then f and g has common factor. Let $a,b \in R[\mathbf{x}]$ with $\deg\ a, \deg\ b > 1$ such that $f+g = ab$. We can write $a = \sum_m^M a_m$ and $b = \sum_l^L b_j$ where $m,l > 1$ and $a_m,\dots,a_M,b_l,\dots,b_L \in R[\mathbf{x}]$ are homogeneous with degree being the index such that $a_m,a_M,b_l,b_L \neq 0$. Note that $d = \deg\ a_m b_l = m+l$. Note also that $d+1 = \deg\ a_M b_L = L+M$, hence (WLOG) $L = l+1$ and $M = m$. We thus find that $ab = a_m b_l + a_m b_{l+1}$. Then $f = a_m b_l$ and $g = a_m b_{l+1}$, hence f,g has a common factor a_m. □

Definition 2.9.112. Let $I \subset R[x_1,\dots,x_{n+1}$ and $J \subset R[x_1,\dots,x_n]$. We define the *dehomogenization of I at* $i \in \{1,\dots,n+1\}$ and the *homogenization of J* to be

$$I_{*,i} = \{f_{*,i} : f \in I\} \text{ resp. } J^* := \langle \{f^* : f \in J\}\rangle \subset R[x_1,\dots,x_{n+1}].$$

We furthermore define $I_* := I_{*,n+1}$

Lemma 2.9.113. *$I_{*,i}$ is an ideal. $I = \langle f_1,\dots,f_m\rangle$, then $I_{*,i} = \langle (f_1)_{*,i},\dots,(f_m)_{*,i}\rangle$. If $J = \langle f\rangle$, then $J^* = \langle f^*\rangle$.*

Proof. The first statement is a trivial consequence of evaluation being a ring homomorphism. The second statement is a matter of checking the definition. □

Example 2.9.114. Consider $I := \langle y-x^2, z-x^3\rangle R[x,y,z]$, Note that $f := z-xy = z-x^3-x(y-x^2) \in I$, hence $f^* = zw - xy \in I^*$, however $f^* \notin J := \langle (y-x^2)^*, (z-x^3)^*\rangle = \langle yw-x^2, zw^2\rangle$, since any term containing z in a polynomial in J has w-degree ≥ 2 or y-degree ≥ 1 or x-degree ≥ 2, therefor no polynomial in J can contain the term zw. We thus see that for a finitely generated ideal $I = \langle f_1,\dots,f_m\rangle$, while trivially $\langle f_1^*,\dots,f_m^*\rangle \subset I^*$, it is not necessarily the case that this holds with equality.

Lemma 2.9.115. *For every $n \geq 1$, $m \geq 1$*

$$\sum_0^m \binom{k+n}{n} = \binom{m+n+1}{n+1}$$

Proof. One readily verifies the $m = 1$ case. By induction we get that

$$\sum_0^{d+1} \binom{k+n}{n} = \binom{m+n+1}{n+1} + \binom{m+n+1}{n} = \binom{(m+1)+n+1}{n+1}.$$

$\square$

Lemma 2.9.116. *For every $n \geq 1$, $d \geq 0$, the set*

$$\Delta_{n,d} := \left\{ v \in \mathbb{N}^n : \sum_1^n v_i = d \right\}.$$

is of size $\binom{d+n-1}{n-1}$.

Proof. Fix $n \geq 1$. In the case $d = 0$, then clearly $\Delta_{n,0} = \{\mathbf{0}\}$, hence $\#\Delta = 1 = \binom{0+n-1}{n-1}$. Suppose the statement is true for some $d \geq 0$. Then for $n = 1$, we see that $\#\Delta_{1,d+1} = 1 = \binom{d+1+(1-1)}{0}$. So for arbitrary $n \geq 1$,

$$\Delta_{n+1,d+1} = \bigcup_0^{d+1} \Delta_j,$$

where

$$\Delta_j := \left\{ v \in \mathbb{N}^{n+1} : v_n = j, \sum_0^{n+1} v_i = d+1 \right\}.$$

Note that these are pairwise disjoint sets and that each for j, Δ_j is in bijection with

$$\left\{ v \in \mathbb{N}^n : \sum_1^n v_i = d+1-j \right\},$$

hence by induction $\#\Delta_j = \binom{d+1-j+n-1}{n-1}$ for $j = 0, \ldots, d+1$. We thus have that

$$\#\Delta_{n+1,d+1} = \sum_0^{d+1} \#\Delta_j = \sum_0^{d+1} \binom{j+n-1}{n-1} = \binom{d+n+1}{n} = \binom{(d+1)+(n+1)-1}{(n+1)-1}.$$

$\square$

Proposition 2.9.117. *For a field K, the dimension of $V_K(d,n)$ is $\binom{d+n-1}{n-1}$.*

Proof. With the notation of the above lemma $\{\mathbf{x}^v \in K[x_1, \ldots, x_n] : v \in \Delta_{n,d}\}$ forms a basis of $V_K(d,n)$, hence by said lemma

$$\dim_K \ V_K(d,n) = \#\Delta_{n,d} = \binom{d+n-1}{n-1}$$

$\square$

Example 2.9.118. $\dim \ V_K(d,1) = 1$, $\dim \ V_K(d,2) = d+1$, $\dim \ V_K(d,3) = \binom{d+2}{2} = \frac{(d+2)(d+1)}{2}$

Proposition 2.9.119. *For each $n \geq 1$, $d \geq 0$,*

$$\dim\ K_{\leq d}[x_1,\ldots,x_n] = \binom{d+n}{d}$$

Proof. One readily verifies that $K_{\leq d}[x_1,\ldots,x_n] = \sum_0^d V_K(d,n)$, hence by Proposition 2.9.117 and Lemma 2.9.115.

$$\dim\ K_{\leq d}[x_1,\ldots,x_n] = \sum_0^d \dim\ V_K(d,n) = \sum_0^d \#\Delta_{n,j} = \sum_0^d \binom{j+n-1}{n-1} = \binom{d+n}{n} = \frac{(d+n)!}{d!n!} = \binom{d+n}{d}.$$

□

Example 2.9.120. We in particular get for $d \geq 1$ that $\dim\ K_{\leq d-1}[x,y] = \binom{d+1}{d-1} = \frac{(d+1)!}{(d+1-d+1)!(d-1)!} = \frac{d(d+1)}{2} = \sum_1^d i$.

Definition 2.9.121. Let K be any field and $f \in K[x_1,\ldots,x_n] \setminus 0$. A point $[v] \in \mathbb{P}^n$ is said to be a *zero of* f if $f(\lambda v) = 0$ for every $\lambda \in K \setminus 0$. We thus write $f([v]) = 0$.

Remark 2.9.122. For a fixed $[v] \in \mathbb{P}^n$ we thus get a well-defined evaluation function on the space of polynomials for which $[v]$ is a zero, mapping to 0. If f is homogeneous of degree d, then if $v \in K \setminus 0$ is a zero of f, $[v]$ is a zero of f. Indeed, for any non-zero $\lambda \in K \setminus 0$ and an $s = (s_1,\ldots,s_{n+1}) \in S^{n+1}$ where $S \supset K$ is a K-algebra. Then

$$f(\lambda s) = \sum_{v \in \mathbb{N}^n} a_v \lambda^{v_1} s_1^{v_1} \cdots \lambda^{v_{n+1}} s_{n+1}^{v_{n+1}} = \lambda^{\sum_1^{n+1} v_i} \sum_{v \in \mathbb{N}^n} a_v s_1^{v_1} \cdots s_{n+1}^{v_{n+1}} = \lambda^d f(s).$$

In particular, if v is a zero of f, then

$$f(\lambda v) = \lambda^d f(v) = 0 \Rightarrow f([v]) = 0.$$

Note that if $[v] \in \mathbb{P}^n$ is a zero of f,g then $(f+g)([v]) = f([v]) + g([v]) = 0$ and $(fg)([v]) = f([v])g([v]) = 0$, hence $[v]$ is a zero of $f+g$ and fg.

Lemma 2.9.123. *Let K be an infinite field. Consider $f = \sum_0^d f_i \in K[x_1,\ldots,x_{n+1}]$ where f_i is homogeneous of degree i. Let $[v] \in \mathbb{P}^n$ be a zero of f. Then $[v]$ is a zero of f_i for each i.*

Proof. Fix $v \in [v]$ and consider

$$g := f(tv_1,\ldots,tv_{n+1}) = \sum_0^f t^i f_i(v) \in K[t].$$

Then $g(\lambda) = 0$ for every $\lambda \in K \setminus 0$, hence $g = 0$. This implies that $f_i(tv_1,\ldots,tv_{n+1}) = t^i f_i(v) = 0$ for each i, meaning $f_i(\lambda v) = 0$ for each $\lambda \in K \setminus 0$. We thus conclude that $f_i([v]) = 0$. □

Definition 2.9.124. Let R be any commutative ring. An ideal $I \subset R[x_1,\dots,x_n]$ is called *homogeneous* if for every $f = \sum_0^d f_i \in R[\mathbf{x}]$ where f_i is homogeneous of degree i, then $f_i \in I$.

Lemma 2.9.125. *For a commutative ring R and a finitely generated $I \subset R[x_1,\dots,x_n]$, I is a homogeneous if and only if I is finitely generated by a finite set of homogeneous polynomials.*

Proof. "$\Rightarrow$": Write $I = \langle f_1,\dots,f_m\rangle$ and $f_i = \sum_0^{d_i} f_{ij}$ with $f_i j$ being homogeneous of degree j. Then $I = \langle\{f_{ij}\}\rangle$. Indeed $I \subset \langle\{f_{ij}\}\rangle$ obviously and $\langle\{f_{ij}\}\rangle \subset I$ since $f_{ij} \in I$ by the assumption that I is homogeneous.
"$\Leftarrow$": Suppose $I = \langle F_1,\dots,F_m\rangle$ for homogeneous F_i of degree d_i. Write $f = \sum_0^d f_i \in I$ with f_i homogeneous of degree i. Write also $f = \sum_1^m \lambda_i F_i$ for $\lambda_1,\dots,\lambda_m \in K[\mathbf{x}]$. If we consider $\lambda_i = \sum_0^{\delta_i} \lambda_{ij}$ with λ_{ij} homogeneous of degree j. Then $f_d = \sum_1^m \lambda_{i,d-d_i} F_i \in I$. By induction in number of non-zero homogeneous f_i, it follows that $f_i \in I$, hence I is homogeneous. □

Remark 2.9.126. Note that it's very clear that an ideal is homogeneous if and only if it is (not necessarily finitely) generated by a set of homogeneous polynomials. We therefor note that the homogenization of an ideal is a homogeneous ideal.

Lemma 2.9.127. *Let $I \subset R[x_1,\dots,x_n]$ be a homogeneous ideal. Then I is prime if and only if $fg \in I$ implies $f \in I$ or $g \in I$ for every form $f,g \in R[\mathbf{x}]$.*

Proof. "$\Rightarrow$": This follows from the definition of prime ideals.
"$\Leftarrow$": Let $\lambda,\mu \in R[\mathbf{x}]$ such that $\lambda\mu \in I$. Write $\lambda = \sum_0^d \lambda_i$ and $\mu = \sum_0^e \mu_i$. Then $\lambda\mu = \sum_{i,j} \lambda_i \mu_j$. Since I is homogeneous $\lambda_d \mu_e \in I$, hence $\lambda_d \in I$ for $\mu_e \in I$. Suppose $\lambda_d \in I$. Then $(\lambda - \lambda_d)\mu \in I$. By induction in the degree of $\lambda\mu$ it follows that $\lambda \in I$ or $\mu \in I$. □

Lemma 2.9.128. *If $I \subset R[x_1,\dots,x_n]$ is prime, then $I^* \subset R[x_1,\dots,x_{n+1}]$ is prime*

Proof. Let $a,b \in R[x_1,\dots,x_{n+1}]$ such that $ab \in I^*$. Then $a_* b_* = (ab)_* \in I$, hence $a_* \in I$ or $b_* \in I$. WLOG $a_* \in I$. Then $(a_*)^* \in I^*$, meaning for a suitable $r \geq 0$, $a = x_{n+1}^r (a_*)^* \in I^*$. □

Proposition 2.9.129. *If $I \subset R[x_1,\dots,x_n]$ is homogeneous, then $\mathbf{rad}(I)$ is homogeneous.*

Proof. Let $f = \sum_0^d f_i \in \mathbf{rad}(I)$. We must prove that $f_i \in \mathbf{rad}(I)$. Note that $f^n = f_d^n + r \in I$ where $\deg r < dn$. Then $f_d^n \in I$, hence $f_d \in \mathbf{rad}(I)$. We thus have that $f - f_d \in \mathbf{rad}(I)$ by induction degree it follows that $f_i \in \mathbf{rad}(I)$ for the remaining i. □

Lemma 2.9.130. *If $\{I_\alpha\}_{\alpha\in A}$ is a family of homogeneous ideals in $R[x_1,\dots,x_n]$, then so is $\sum_{\alpha\in A} I_\alpha$ and $\bigcap_\alpha I_\alpha$.*

Proof. Indeed, for $(f_\alpha)\in\bigoplus_{\alpha\in A} I_\alpha$, we may for some $d\geq 0$ write $(f_\alpha)=(\sum_0^d f_{\alpha,i})$, where $f_{\alpha,i}$ is homogeneous of degree i for each i and α. Then

$$\sum_{\alpha\in A} f_\alpha = \sum_0^d \sum_{\alpha\in A} f_{\alpha,i}.$$

Note that since each I_α is homogeneous $f_{\alpha,i}\in I_\alpha$. Then $\sum_{\alpha\in A} f_{\alpha,i}\in\sum_{\alpha\in A} I_\alpha$, which means $\sum_{\alpha\in A} I_\alpha$ is homogeneous.
Consider $f=\sum_0^d f_i\in\bigcap_{\alpha\in A} I_\alpha$. Then for each $\alpha\in A$, $f\in I_\alpha$, hence $f_i\in I_\alpha$. We thus have that $f_i\in\bigcap_{\alpha\in A} I_\alpha$, which means $\bigcap_{\alpha\in A} I_\alpha$ is homogeneous. □

Lemma 2.9.131. *Let $I,J\subset R[x_1,\dots,x_n]$ be homogeneous ideals. Then IJ is homogeneous.*

Proof. Let $f=\sum_0^d f_i\in I$ and $g=\sum_0^e g_j\in J$. Then each $f_i\in I$ and each $g_j\in J$, meaning $f_ig_j\in IJ$ for each $0\leq i\leq d$ and $0\leq j\leq e$. Then since $fg=\sum_{k=0}^{d+e}\sum_{i,j:i+j=k} f_ig_j$ where $\sum_{i,j:i+j=k} f_ig_j$ is a homogeneous polynomial of degree k for each $0\leq k\leq d+e$ we get that IJ is homogeneous. □

Definition 2.9.132. Let $I\subset R[x_1,\dots,x_n]$ be a homogeneous ideal. An element $\alpha\in R[\mathbf{x}]/I$ is called *homogeneous of degree* d if there is a homogeneous polynomial of degree d, $f\in R[\mathbf{x}]$, such that $\alpha=f+I$.

Lemma 2.9.133. *Let $I\subset R[x_1,\dots,x_n]$ be a homogeneous polynomial. Let $\alpha\in R[\mathbf{x}]/I$. Then for some unique $d\geq 0$, there are unique $\alpha_i\in R[\mathbf{x}]/I$, $i\in\{0,\dots,d\}$ homogeneous of degree i such that*

$$\alpha=\sum_0^d \alpha_i.$$

Proof. **Existence:** Let $f+I\in R[\mathbf{x}]$, then $f=\sum_0^d f_i$ where f_i is homogeneous of degree i for each i. Then we are done picking $\alpha_i := f_i+I$. **Uniqueness:** Suppose we are given two such representations $\sum_0^d(f_i+I)=\sum_0^d(g_i+I)$ (we can always let d be the largest of the two degrees obtained from each respective representation and then set undefined forms to be equal to 0). Then f_i-g_i is a form of degree i for each i, hence $f_i-g_i\in I$ using the fact that I is homogeneous. Consequently $f_i+I=g_i+I$ for each i. □

Remark 2.9.134. Consider $V_R(d,n,I)=\{\alpha\in R[\mathbf{x}]/I : \alpha \text{ homogeneous of degree } d\}$ is an R-submodule of $R[\mathbf{x}]/I$ finitely generated by $\{\mathbf{x}^v+I : |v|=d\}$. In particular

$V_K(d,n,I)$ is a finite dimensional vector space for fields K. In general, it takes some work to say anything about the dimension of this vector space, below we give an example in the case $n=3$ and $d>n$ This is exercise 4.10.

Example 2.9.135.

Lemma 2.9.136. *Let $f \in K[x_1,\dots,x_n]$ be a non-zero form of degree d. Then f_* is non-zero.*

Proof. $0 \neq f = x_{n+1}^r (f_*)^*$ for some $r \geq 0$, hence $(f_*)^* \neq 0$, so by Remark 2.9.108, $f_* \neq 0$. □

2.9.15 Multi- and Bihomogeneous Polynomials

Definition 2.9.137. Let $\{x_{ij} : 1 \leq i \leq m, 1 \leq j \leq n_i\}$ be algebraically independent variables over a ring R. A polynomial in $R[\mathbf{x}]$ is called an m*-homogeneous polynomial* or an m*-form* of m*-degree* $(d_1,\dots,d_m) \in \mathbb{Z}^m_{\geq 0}$, if it is form of degree d_i when seen as an element in $R[x_{kj} : k \neq i][x_{i1},\dots,x_{in_i}]$ for each i. When $m=2$ an 2-form is called a *bihomogeneous polynomial* or a *biform* of *bidegree* (d_1,d_2).

Lemma 2.9.138. *When the same notation as above a polynomial $f \in K[\mathbf{x}]\backslash 0$ of degree d has unique decomposition*

$$f = \sum_{(i_1,\dots,i_m):\sum_1^m i_j \leq d} f_{i_1,\dots,i_m},$$

where $f_{i_1,\dots,i_m}$ is an m-form of m-degree $(i_1,\dots,i_m)$ such that for some $(i_1,\dots,i_m)$ with $\sum_1^m i_j = d$, $f_{i_1,\dots,i_m} \neq 0$.

Proof. Write

$$f = \sum_{(i_1,\dots,i_m):\sum_1^m i_j \leq d} \underbrace{\sum_{\substack{\mathbf{v}\in\prod_1^m \mathbb{N}^{n_k}:\\ \forall k, \sum v_{kj} = i_k}} a_{\mathbf{v}} \mathbf{x_1}^{v_1} \cdots \mathbf{x_m}^{v_m}}_{f_{i_1,\dots,i_m}},$$

each monomial in $f_{i_1,\dots,i_m}$ is a homogeneous of degree i_k in $R[x_{kj} : k \neq i][x_{i1},\dots,x_{in_i}]$ for each k, hence $f_{i_1,\dots,i_m}$ is m-homogeneous of m-degree $(i_1,\dots,i_m)$. Uniqueness follows from $\{\mathbf{x}^{\mathbf{v}}\}$ being linearly independent over R. □

Definition 2.9.139. For an $f \in R[\mathbf{x_1},\dots,\mathbf{x_m}]$ we say that $([v_1],\dots,[v_m]) \in \mathbb{P}^{n_1} \times \mathbb{P}^{n_m}$ is *a zero of* f if for every $(\lambda_1,\dots,\lambda_m) \in K^m$, $\lambda_i \neq 0$, $f(\lambda_1 v_1,\dots,\lambda_m v_m) = 0$.

Remark 2.9.140. If f were an m-form of m-degree $(d_1,\dots,d_m)$ note that $f(\lambda_1 v_1,\dots,\lambda_m v_m) = \left(\prod_1^m \lambda_i^{d_i}\right) f(v_1,\dots,v_m)$. Hence if $(v_1,\dots,v_m)$ is a zero of f, then so is $([v_1],\dots,[v_m])$.

Definition 2.9.141. An ideal $I \subset R[\mathbf{x_1},\ldots,\mathbf{x_m}]$ is called m-homogeneous if for each $f = \sum_{i_1,\ldots,i_m} f_{i_1,\ldots,i_m} \in I$, $f_{i_1,\ldots,i_m} \in I$

Remark 2.9.142. The above is equivalent to I being a homogeneous ideal in $R[\mathbf{x_1},\ldots,\widehat{\mathbf{x_k}},\ldots,\mathbf{x_m}]$ for each k. Any result proven about homogeneous ideals therefor naturally generalizes to m-homogeneous ideals.

2.9.16 Differentiation of Polynomials

Definition 2.9.143. We define *differentiation (with respect to x)* in polynomial ring $R[x]$ as the R-module map

$$D_x : R[x] \to R[x]$$

mapping 1 to 0 and x^n to x^{n-1} for $n \geq 1$.For a polynomial $f \in R[x]$, we call the polynomial $D_x f$ the *derivative of f (with respect to x)*, which we may denote by f'.

Remark 2.9.144. One easily checks that $D_x fg = (D_x f)g + f(D_x g)$, i.e. that is satisfies the Leibniz rule. It also satisfies the chain rule, i.e.

$$D_x(f(g)) = (D_x g) \cdot (D_x f)(g).$$

By definition **over an integral domain of characteristic 0**, $\deg f' = \deg f - 1$ when $\deg f \geq 1$. In positive characteristic this may not be the case. For instance, when $\operatorname{char} R = p > 0$, we get that $D_x x^p = p x^{p-1} = 0$. We can also come up with pathological examples over commutative rings of characteristic 0. Indeed take $R = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}$ with the usual structure of product ring. Then $D_x(1,0)x^2 = (2,2)(1,0)x = (0,2)(1,0)x = 0$.

Definition 2.9.145. In a polynomial ring $R[x_1,\ldots,x_n]$, *the partial derivative of f with respect to x_i* is the polynomial $D_{x_i} f$, where

$$D_{x_i} : R[x_1,\ldots,x_n] = R[x_1,\ldots,x_{i-1},\widehat{x_i},x_{i+1},x_n][x_i] \to R[x_1,\ldots,x_n] = R[x_1,\ldots,x_{i-1},\widehat{x_i},x_{i+1},x_n][x_i],$$

' is differentiation with respect to x_i. We sometimes denote it by $\frac{\partial f}{\partial x_i} := \frac{\partial}{\partial x_i} f := f_{x_i}$.

Lemma 2.9.146. *Let T be translation of one variable, x_1 say, by some element $a \in K$. Then $D_{x_i} T = T D_{x_i}$. Hence in general translation commutes with partial derivatives.*

Proof. Indeed,

$$(D_{x_1} T) f = D_{x_1} f(x_1 + a, x_2, \ldots, x_n) = (D_{x_1}(x_1 + a)) \cdot (D_{x_1} f)(x_1 + a, x_2, \ldots, x_n) = (T D_{x_1}) f.$$

(Here we implicitly use that differentiation commutes with permutation of variables).

□

Lemma 2.9.147. *(Euler's theorem) Let R be an integral domain, $f \in R[x_1,\ldots,x_n]\setminus 0$ be homogeneous of degree $d>0$. Then*

$$\sum_1^n x_i \frac{\partial f}{\partial x_i} = df.$$

Proof. Let $a_v\mathbf{x}^v$ be a term of f. Then

$$\frac{\partial}{\partial x_i} a_v \mathbf{x}^v = \begin{cases} v_i a_v \mathbf{x}^{v-e_i} & \text{if } v_i > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_1^n x_i \frac{\partial}{\partial x_i} a_v \mathbf{x}^v = \sum_{i:v_i>0} v_i a_v x_i \mathbf{x}^{v-e_i} = \left(\sum_{i:v_i>0} v_i\right) a_v x^v = d a_v \mathbf{x}^v.$$

It thus follows that

$$\sum_1^n x_i \frac{\partial f}{\partial x_i} = \sum_1^n x_i \sum_{v\in\mathbb{N}^n} \frac{\partial}{\partial x_i} a_v \mathbf{x}^v = \sum_{v\in\mathbb{N}^n} \sum_1^n x_i \frac{\partial}{\partial x_i} a_v \mathbf{x}^v = d \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v = df.$$

□

We provide a lemma which alleviate some ugly cases in characteristic $p>0$ cases

Lemma 2.9.148. *Let R be an integral domain of characteristic $p>0$. Let $f \in R[x_1,\ldots,x_n]$ be a non-constant polynomial such that $f_{x_i} = 0$ for each i. Then $f = g(x_1^p,\ldots,x_n^p)$ for some g.*

Proof. For any $k \in \mathbb{Z}$, let k_R denote the image of k in R. Consider the case where $n=1$. Write $f = \sum_0^d a_i x^i$. Then $0 = f_x = \sum_1^d i_R a_i x^{i-1}$. For each i, we then get that $i_R a_i = 0$, hence $a_i = 0$ or $i_R = 0$. If $i<p$, then $i_R \neq 0$, hence $a_i = 0$. We furthermore have that if $a_i \neq 0$ then $p \mid i$ It thus follows that $f = \sum_1^k a_{jp} x^{jp}$, where k is chosen such that $d = kp$. Hence picking $g = \sum_1^k a_{jp} x^j$ we are done. We prove the general case by induction using the fact that $\operatorname{char} R[x_1,\ldots,x_n] = p$, $R[x_1,\ldots,x_{n+1}] \simeq R[x_1,\ldots,x_n][x_{n+1}]$ in conjunction with the validity of the 1-variable case. □

2.10 Ring Extensions and Algebras over Rings

We proceed with considerations of the theory of algebras over rings in conjunction with some of the theory modules already developed. We could have noted this earlier, but it is more relevant to note it now.

2.10.1 Finitely Generated Ring Extensions

Definition 2.10.1. A ring extension $S \supset R$ is said to be *module-finite (over R)* if S is finitely generated as an R-module.

Definition 2.10.2. A ring extension $S \supset R$ is a said to be *finitely generated R-algebra* or *ring-finite* if $S = R[s_1, \dots, s_n]$ for suitable $s_1, \dots, s_n \in S$.

Proposition 2.10.3. *Let $S \supset R$ be a ring-finite ring extension. Then*

$$S \simeq R[x_1, \dots, x_n]/I,$$

for some $n \geq 1$ and some ideal $I \subset R[x_1, \dots, x_n]$.

Proof. For suitable $s_1, \dots, s_n \in S$, $S = R[s_1, \dots, s_n]$, hence $\mathrm{ev}_{s_1,\dots,s_n} : R[x_1, \dots, x_n] \to S$ is a surjective R-algebra homomorphism. Then by the first isomorphism theorem $S \simeq K[\mathbf{x}]/\ker\ \mathrm{ev}_{s_1,\dots,s_n}$. □

Proposition 2.10.4. *Let $S \supset R$ is a ring extension that is also a finitely generated R-module. Then S is a finitely generated R-algebra.*

Proof. For suitable $s_1, \dots, s_n \in S$, $S = \sum_1^n R s_i$. We prove that $S = R[s_1, \dots, s_n]$. We already know that $S \supset R[s_1, \dots, s_n]$. Let $s \in S$. Then $s = \sum_1^n r_i s_i$ for suitable $r_1, \dots, r_n \in R$, hence $s \in R[s_1, \dots, s_n]$. □

Example 2.10.5. The converse implication of the above proposition is clearly not true. Consider for instance $R[x_1, \dots, x_n] \supset R$. given $f_1, \dots, f_m \in R[\mathbf{x}]$. If these are all 0, clearly $R[\mathbf{x}] \supsetneq \sum_1^m R f_i$. Otherwise putting $D = \max_{i \in \{1,\dots,m\}} \{\deg\ f_i\}$, we see that for $r_1, \dots, r_m \in R$,

$$\deg \sum_1^m r_i f_i \leq D < D + 1 = \deg\ x_1^{D+1},$$

hence $x_1^{D+1} \notin \sum_1^m R f_i$, hence $R[\mathbf{x}] \supsetneq \sum_1^m R f_i$.

Definition 2.10.6. A ring extension $L \supset K$ is called a *field extension (over K)* if both L and K are fields.

Let $S \supset R$ be a ring extension where S is an integral domain. For $s_1, \dots, s_n \in S$, $R[s_1, \dots, s_n]$ is also an integral domain. We denote the fraction field of $R[s_1, \dots, s_n]$ by $R(s_1, \dots, s_n)$

Definition 2.10.7. A field extension $L \supset K$ is said to be *finite*, if there exist $a_1, \dots, a_n$ such that $L = K(a_1, \dots, a_n)$.

Lemma 2.10.8. *Let K be a field. Consider $K[x_1,\dots,x_n]$ as a subring of $K[\mathbf{x},y_1,\dots,y_m]$ Then $K(x_1,\dots,x_n)(y_1,\dots,y_m) = K(x_1,\dots,x_n,y_1,\dots,y_m)$.*

Proof. Clearly $K(\mathbf{x}) \subset K(\mathbf{x},\mathbf{y})$ and one easily verifies that this is a subfield of $K(\mathbf{x},\mathbf{y})$. To be very precise, this means $K(\mathbf{x},\mathbf{y}) \supset K(\mathbf{x})$ is a ring extension. Hence $K(\mathbf{x},\mathbf{y}) \supset K(\mathbf{x})[\mathbf{y}]$. This means $K(\mathbf{x},\mathbf{y}) = Q(K(\mathbf{x},\mathbf{y})) \supset Q(K(\mathbf{x})[\mathbf{y}]) = K(\mathbf{x})(\mathbf{y})$. □

Remark 2.10.9. Of course this statement What?

Lemma 2.10.10. *Consider $L := K(x_1,\dots,x_n)$ and $R := K\left[\frac{a_1}{b_1},\dots,\frac{a_m}{b_m}\right]$ for $\frac{a_1}{b_1},\dots,\frac{a_m}{b_m} \in K(\mathbf{x})$. Then there is a $b \in K[\mathbf{x}]$ such that $b^d z \in K[\mathbf{x}]$ for every $z \in R$ for some $d \geq 0$.*

Proof. If $\operatorname{lcm}(b_1,\dots,b_m) = 1$ the statement is trivial. Set $b := \operatorname{lcm}(b_1,\dots,b_m)$ and assume $\deg b > 0$. Let $z \in R$. If $z = 0$, then $b^0 z \in K[\mathbf{x}]$. Suppose $z \neq 0$. For some $f \in K[y_1,\dots,y_m] \setminus 0$, $z = f\left(\frac{a_1}{b_1},\dots,\frac{a_m}{b_m}\right)$. Set $d := \deg f$. Let $v \in \mathbb{N}^m$ with $|v| \leq d$. Then

$$\prod_1^m b_i^{v_i} \mid \prod_1^m b^{v_i} = b^{|v|} \text{ and } b^{|v|} \mid b^d \Rightarrow \prod_1^m b_i^{v_i} \mid b^d \Rightarrow b^d \prod_1^m \left(\frac{a_i}{b_i}\right)^{v_i} \in K[\mathbf{x}] \Rightarrow b^d z \in K[\mathbf{x}].$$

□

Proposition 2.10.11. *Consider $L := K(x_1,\dots,x_n) \supset K$. Then L is a finite field extension over K, but not a finitely generated K-algebra.*

Proof. The first statement is obvious as $K(x_1,\dots,x_n)$ is finitely generated as field extension over K by $x_1,\dots,x_n \in K(x_1,\dots,x_n)$. To prove the second statement, let $\frac{a_1}{b_1},\dots,\frac{a_m}{b_m} \in K(\mathbf{x})$. Set $b := \operatorname{lcm}(b_1,\dots,b_m)$. Then there is a $c \in K[\mathbf{x}]$ such that $c \nmid b^d$ for any $d \geq 0$, since There are infinitely many irreducible pol. over K and $K[\mathbf{x}]$ is a UFD, hence $b^d \frac{1}{c} \notin K[x]$ for any $d \geq 0$. By lemma 2.10.10, it follows that $\frac{1}{c} \notin K\left[\frac{a_1}{b_1},\dots,\frac{a_m}{b_m}\right]$, hence $K(\mathbf{x}) \supsetneq K\left[\frac{a_1}{b_1},\dots,\frac{a_m}{b_m}\right]$. This means $K(\mathbf{x})$ is not a finitely generated K-algebra. □

Lemma 2.10.12. *These finiteness conditions are transitive, i.e. the following three statements are true:*

1. *Let $T \supset S, S \supset R$ be module-finite. Then $T \supset R$ is module-finite.*
2. *Let $T \supset S$, $S \supset R$ be ring-finite. Then $T \supset R$ is ring-finite.*
3. *Let $M \supset L$, $L \supset K$ be finite field extensions. Then $M \supset K$ is a finite field extension.*

Proof. 1. We can find $s_1,\dots,s_n \in S$ such that $S = \sum_1^n Rs_i$ and $t_1,\dots,t_m \in T$ such that $T = \sum_1^m St_i$. Let $t \in T$. Then there are $a_1,\dots,a_m \in S$ such that $t = \sum_1^m a_i t_i$. For each i, $a_i = \sum_{j=1}^n b_{ij} s_j$ for suitable $b_{ij} \in R$, hence

$$t = \sum_1^m \sum_1^n b_{ij} t_i s_j \in \sum_1^n \sum_1^m R t_i s_j.$$

Hence T is finitely generated as an R-module by the elements of $\{t_i s_j : i \in \{1,\dots,m\}, j \in \{1,\dots,n\}\}$.
2. We can find $s_1,\dots,s_n \in S$ such that $S = R[s_1,\dots,s_n]$ and $t_1,\dots,t_m \in T$ such that $T = S[t_1,\dots,t_m]$. Let $t \in T$. Then there are $a_v \in S$ such that

$$t = \sum_{v \in \mathbb{N}^m} a_v t_1^{v_1} \cdots t_m^{v_m}.$$

For each $v \in \mathbb{N}^m$,

$$a_v = \sum_{w \in \mathbb{N}^n} b_{vw} s_1^{w_1} \cdots s_n^{w_n} a_v$$

for suitable $b_{vw} \in R$, hence

$$t = \sum_{v \in \mathbb{N}^m} \sum_{w \in \mathbb{N}^n} b_{vw} t_1^{v_1} \cdots t_m^{v_m} s_1^{w_1} \cdots s_n^{w_n} \in R[s_1,\dots,s_n,t_1,\dots,t_m].$$

Hence T is finitely generated as an R-algebra by the elements of $s_1,\dots,s_n,t_1,\dots,t_m \in T$.
3. There are $\alpha_1,\dots,\alpha_m \in M$ such that $M = L(\alpha_1,\dots,\alpha_m)$ and $\beta_1,\dots,\beta_n \in L$ such that $L = K(\beta_1,\dots,\beta_n)$. Then

$$M = L(\alpha_1,\dots,\alpha_m) = K(\beta_1,\dots,\beta_n)(\alpha_1,\dots,\alpha_m) = K(\alpha_1,\dots,\alpha_m,\beta_1,\dots,\beta_n).$$

□

2.10.2 Integral- & Algebraic Extensions

Definition 2.10.13. Let $S \supset R$ be a ring extension. An element of $s \in S$ is said to be *integral (over R)* if it is algebraically dependent over R, i.e. there is a monic $f \in R[x] \setminus 0$ such that $f(s) = 0$.
Let $L \supset K$ be a field extension. An element of L is *algebraic (over K)* if it is integral over K. An element that is not algebraic over K is transcendental over K.

Remark 2.10.14. In the case R is a field, consider $d = \min(\{k > 0 : \exists f \in R[x] \setminus 0, f(s) = 0\})$. Let f_s be a polynomial of degree d vanishing on s. Consider another polynomial $g \in R[x] \setminus 0$ vanishing on s. We can write $g = qf_s + r$, where $r = 0$ or $\deg\ r < d$. Then

$$r(s) = g(s) - q(s) f_s(s) = 0.$$

By minimality $r = 0$, hence $f_s \mid g$. There f_s is the unique non-zero polynomial vanishing on s of minimal degree. We call this polynomial the *defining polynomial of s over K*. Note that $\ker \operatorname{ev}_s = \langle f_s \rangle$, hence $R[s] \simeq R[x]/\langle f_s \rangle$. We refer to $\deg f_s$ as *the degree of s over R*. We can extend this result to the case where R is a UFD. By the same argument as before $f_s \mid g$ in $Q(R)[x]$, hence since f is primitive, $f_s \mid g$ in $R[x]$ by Lemma 2.9.95

Remark 2.10.15. Note that if $a_1, \ldots, a_n \in L \supset K$ are algebraic over K ($L \supset K$ is a field extension), then

$$K[a_1, \ldots, a_n] = K(a_1, \ldots, a_n).$$

Indeed, in the case $n = 1$, $K[a] \simeq K[x]/\langle f_a \rangle$. Then $K[a]$ is a subfield of $K(a)$, and since $K(a)$ is the smallest subfield containing $K[a]$, $K[a] = K(a)$. By induction $K[a_1, \ldots, a_n] = K(a_1, \ldots, a_n)$, so it is sufficient to prove that $K(a_1, \ldots, a_n)[a_{n+1}] = K(a_1, \ldots, a_n)(a_{n+1})$. Since a_{n+1} is algebraic over K it is algebraic over $K(a_1, \ldots, a_n)$, hence by the base case, $K(a_1, \ldots, a_n)[a_{n+1}] = K(a_1, \ldots, a_n)(a_{n+1})$.

Example 2.10.16. Let R be an integral domain. Then $Q(R) \supset R$ is integral. Indeed consider $\alpha = \frac{a}{b} \in Q(R)$. Then α vanishes on $bx - a \in R[x] \setminus 0$

Lemma 2.10.17. *Let $S \supset R$ be ring extension where S is an integral domain. Furthermore, let $s \in S$. The following are equivalent*

1. *s is integral over R.*
2. *$R[s] \supset R$ is module-finite.*
3. *There is a subring of S containing $R[s]$, R' say, which is finitely generated as an R-module.*

Proof. "1. $\Rightarrow$ 2.": We may find $a_0, \ldots, a_{n-1} \in R$ such that

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

It follows that $s^n \in \sum_1^{n-1} Rs^i$. By a simple induction argument it follows that $s^{n+j} \in \sum_1^{n-1} Rs^i$ for every $j \geq 0$. Let $\sum_1^m b_j s^j \in R[s]$. Then by the considerations prior to this, $\sum_1^m b_j s^j \in \sum_1^{n-1} Rs^i$, hence $R[s] = \sum_1^{n-1} Rs^i$.
"2. $\Rightarrow$ 3.": Putting $R' = R[s]$ we have such a subring of S.
"3. $\Rightarrow$ 1.": We can write $R' = \sum_1^n a_i t_i$ for suitable $t_1, \ldots, t_n \in R[s] \setminus 0$. Then

$$st_i = \sum_1^n a_{ij} t_i$$

for suitable $a_{ij} \in R$. Note then that

$$s\begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix} = (a_{ij})\begin{pmatrix} t_1 \\ \vdots \\ t_n \end{pmatrix}.$$

This implies that s is a root of the characteristic polynomial

$$\det\left(x\mathbb{1} - (a_{ij})\right) \in R[x] \setminus 0$$

which is monicref.. It thus follows that s is integral over R. ☐

Lemma 2.10.18. *Consider a tower of ring extensions $T \supset S \supset R$ where T is a domain. Suppose T is integral over S and S is integral over R, then T is integral over R.*

Proof. Let $t \in T$, then there is a monic $f = x^n + \sum_0^{n-1} a_i x^i \in S[x]\setminus 0$ such that $f(t) = 0$. By the above lemma $R' := R[a_1,\ldots,a_n,t] = R[a_1,\ldots,a_n][t] \supset R[a_1,\ldots,a_n]$ is module-finite. By the above lemma we also find that $R[a_1] \supset R$ is module-finite. Recursive usage of the above and the transitivity of module-finiteness thus implies that $R[a_1,\ldots,a_{n-1}] \supset R$ is module-finite and hence that $R' = R[a_1,\ldots,a_{n-1},t] \supset R$ is module-finite. We have thus found a subring of T containing $R[t]$, which is module-finite over R. Hence using the above lemma yet another time it follows that t is integral over R, hence T is integral over R. ☐

Proposition 2.10.19. *Let $S \supset R$ be an integral domain and a ring extension. Then*

$$\left\{s \in S : s \textit{ is integral over } R\right\}$$

is a subring S.

Proof. Let $a, b \in S$ be integral over R. We repeatably use Lemma 2.10.17. Note that b is integral over $R[a]$ hence $R[a,b] \subset R$ is module-finite. Since $a+b, ab \in R[a,b]$, it follows that $R[a,b]$ is a ring contained in S, containing $R[a+b]$ and $R[ab]$ that is module-finite over R, meaning $a+b$ and ab are integral over R. ☐

Lemma 2.10.20. *If $S \supset R$ is module-finite then $S \supset R$ is integral.*

Proof. Let $s \in S$. The ring S is a subring of S containing $R[s]$ which is finitely generated as an R-module. It thus follows by Lemma 2.10.17 that s is integral over R. Hence $S \supset R$ is integral. ☐

Lemma 2.10.21. *Let $S \supset R$ be a ring extension with S an integral domain. Then $S \supset R$ is module-finite if and only if $S = R[s_1,\ldots,s_n]$ where $s_i \in S$ is integral over R.*

Proof. "$\Rightarrow$": This follows from Lemma 2.10.20 "$\Leftarrow$": By assumption there are $s_1,\ldots,s_n \in S$ such that $S = R[s_1,\ldots,s_n]$. Since $s_i \in S$ is integral over R, it follows by Lemma 2.10.17 that $R[s_1] \supset R$ is module-finite and by induction $S = R[s_1,\ldots,s_n] \supset R$ is module-finite. □

Lemma 2.10.22. *Let $L \supset K$ be a field extension with K algebraically closed.*

1. *Every $f \in K[x] \setminus 0$ with $n := \deg f > 0$ has exactly n roots all in K.*

2. *If $a \in L$ is algebraic over K, then $a \in K$.*

3. *If $L \supset K$ is module-finite, then $L = K$.*

Proof. 1. We proceed by induction in n. For $n = 1$, f has a root $a \in K$, and since f has exactly one root, this is the only root.
Suppose now f has degree $n+1$, then f has a root $a \in K$, hence $f = (x-a)g$ for a polynomial $g \in K[x] \setminus 0$ with $\deg g = n$. By induction g has exactly n roots in K. Using that K is an integral domain it follows that f has $n+1$ roots.
2. If $a \in L$ is algebraic over L, then there is a polynomial $f \in K[x] \setminus 0$ such that a is a root f. Since $V(f) \subset K$ by 1. it follows that $a \in K$.
3. If $L \supset K$ is module-finite, then it is algebraic by Lemma 2.10.20. Let $a \in L$. Then a is algebraic over K, hence $a \in K$ by 2. We thus get that $L = K$. □

Lemma 2.10.23. *Let K be a field and set $L := Q(K[x]) = K(x)$. Then*

1. *$a \in L$ is integral then $a \in K[x]$.*

2. *There is no $f \in K[x] \setminus 0$ such that for every $a \in L$, $F^n a$ is integral over $K[x]$ for some $n > 0$.*

Proof. 1. We may write $a = \frac{f}{g}$ for $f, g \in K[x]$ with $g \neq 0$ and $\gcd(f,g) = 1$. We can then find $a_0,\ldots,a_{n-1} \in K[x]$ such that

$$\frac{f^n}{g^n} + \sum_0^{n-1} a_i \frac{f^i}{g^i} = 0 \Rightarrow f^n = \sum_0^{n-1} a_i g^{n-i} f^i = g \sum_0^{n-1} a_i g^{n-(i+1)} f^i,$$

hence $g \mid f^n$, meaning $g \mid f$, hence $g \in K \setminus 0$. This implies that $a = \frac{f}{g} = g^{-1} f \in K[x]$.
2. By Proposition 2.10.11 $K(x) \supsetneq R := K[z_1,\ldots,z_m]$ for any $z_1,\ldots,z_m \in K(x)$. Recall that we proved this by showing that for any $f \in K[x] \setminus 0$ there is some $c \in K[x]$ such that $f^d \frac{1}{c} \notin K[x]$ for any $d > 0$. By 1. this implies that $f^d \frac{1}{c}$ is not integral over K for a. □

Proposition 2.10.24. *Let $L \supset K$ be a field extension. The set*

$$\{a \in L : a \text{ is integral}/K\}$$

is a subfield of L.

Proof. We already know that it is a subring by Proposition 2.10.19. Let $a \in L \setminus 0$ be integral over K. Then there are $a_0, \ldots, a_{n-1} \in K$ such that

$$0 = a^n + \sum_0^{n-1} a_i a^i,$$

where choosing n minimal implies, $a_0 \neq 0$, hence $a\left(a^{n-1}(-a_0)^{-1}\sum_1^{n-1} a_i a^i\right) = 1$, implying a is a unit. □

Proposition 2.10.25. *Let $L \supset K$ be a module-finite field extension. Consider a subring R of L containing K as a subring. Then R is a field.*

Proof. By Lemma 2.10.20, $L \supset K$ is algebraic. Let $r \in R \setminus 0$. Then r has a multiplicative inverse $r^{-1} \in L \setminus 0$. We can thus find $a_0, a_1, \ldots, a_{n-1} \in K$ such that

$$r^{-n} + \sum_0^{n-1} a_i r^{-i} = 0,$$

This implies that

$$r^{-1} = r^{n-1} r^{-n} = -\sum_0^{n-1} a_i r^{n-1} r^{-i} = -\sum_0^{n-1} a_i r^{n-1-i} \in R,$$

hence R is a subfield of L. □

Theorem 2.10.26. *Let $L \supset K$ be a ring-finite field extension generated by $a_1, \ldots, a_n \in L$. Then $L \supset K$ is module finite and hence also algebraic.*

Proof. We use induction in n. For $n = 1$, suppose $L = K[a]$ for some $a \in L$. Consider $\rho = \mathrm{ev}_a : K[x] \to L$. Since $K[x]$ is a PID add result!, we find that $\ker\ \rho = \langle g \rangle$ for some $g \in K[x]$. Then $K[x]/\langle g \rangle \simeq K[a] = L$. We claim that $g \neq 0$
Proof of the claim: Suppose $K[x] \simeq K[a]$, then $K(x) \simeq K(a)$, but then L is not ring-finite over K, by Proposition 2.10.11 leading to a contradiction.
So we may WLOG assume g is monic. Then a is algebraic over K.
Assume the statement is true for some $n \geq 1$. Suppose $L = K[a_1, \ldots, a_n]$. Set $K' = K(a_1)$. Then by induction $L = K'[a_2, \ldots, a_{n+1}]$ is algebraic. Suppose a_1 is algebraic over K. Then $K' \supset K$ is algebraic, hence $L \supset K$ is algebraic. Suppose for

a contradiction that a_1 is not algebraic over K. We note that this implies that $K[a_1] \overset{\sigma}{\simeq} K[x]$ add reference!. We have identities

$$a_i^{n_i} + \sum_{j=0}^{n_i-1} \alpha_{ij} a_i^j = 0,$$

for each $i \geq 2$ for suitable $n_i \geq 1$, $\alpha_{ij} \in K'$. Let $\alpha \in K[a_1]$ be the common denominator of the α_{ij}. Let $M \geq \max_{i \in \{2,\ldots,n+1\}} n_i$. Then

$$(\alpha a_i)^M + \sum_{j=0}^{n_i-1} \alpha^{M-j} \alpha_{ij} (\alpha a_i)^j = 0,$$

hence $\alpha^M a_i$ is integral over $K[a_1]$. Let $z = \sum_{v \in \mathbb{N}^{n+1}} c_v a_1^{v_1} \cdots a_{n+1}^{v_{n+1}} \in L$. Then taking $N \geq 0$ sufficiently large we get that $\alpha^N z$ is integral over $K[a_1]$ by Proposition 2.10.19. However taking $z \in K(a_1)$, this implies that $\sigma(z) \in K(x)$ is a polynomial such that $\sigma(\alpha_1)^N \sigma(z)$ is integral over $K[x]$, leading to a contradiction by Lemma 2.10.23. □

Corollary 2.10.27. *Let $L \supset K$ be a field extension where K is algebraically closed. Suppose also that there is a surjective K-algebra homomorphism from $K[x_1,\ldots,x_n]$ to L for some $n > 0$. Then $K = L$.*

Proof. Let $\sigma : K[x_1,\ldots,x_n] \to L$ be a surjective K-algebra map. By Corollary 2.9.26 there are $a_1,\ldots,a_n \in L$ such that $\sigma = \mathrm{ev}_{a_1,\ldots,a_n}$. It thus follows that $L = \sigma(K[\mathbf{x}]) = \mathrm{ev}_{a_1,\ldots,a_n}(K[\mathbf{x}]) = K[a_1,\ldots,a_n]$, hence by the above theorem L is module-finite over K. It follows from Lemma 2.10.22 that $L = K$. □

Corollary 2.10.28. *Let K be algebraically closed and $I \subset K[x_1,\ldots,x_n]$ be a maximal ideal. Then $K[x_1,\ldots,x_n]/I = K$ (thinking about K as the canonical embedding of K in $K[\mathbf{x}]/I$). This implies $I = \langle x_1 - a_1,\ldots,x_n - a_n\rangle$.*

Proof. The quotient map $\pi : K[\mathbf{x}] \to K[\mathbf{x}]/I, f \mapsto f + I$ is a canonically a surjective K-algebra homomorphism. It follows by the above corollary that $K[\mathbf{x}]/I = K$. Then for each i, $x_i + I = a_i + I$ for some $a_i \in K$, this means $J := \langle x_1 - a_1,\ldots,x_n - a_n\rangle \subset I$. J is maximal by Corollary 2.9.39, hence $J = I$. □

2.10.3 Field Extensions

Lemma 2.10.29. *Let $L \supset K$ be a field extension and $a \in L$. Then a is module finite if and only if $\dim_K K[a] < \infty$. If a is algebraic, let d denote $\deg f_a$. Then $\{1,a,\ldots,a^{d-1}\}$ is a basis for $K[a]$.*

Proof. The first statement follows immediately from Proposition 2.10.17. Set $I = \langle f_a \rangle$. By Lemma 2.9.46, $\{1+I, x+I, \dots, x^{d-1}+I\}$ is a basis of $K[x]/I \overset{\overline{\mathrm{ev}_a}}{\simeq} K[a]$, since ev_a is a K-algebra homomorphism, it is in particular a K-linear map, hence $\{1, a, \dots, a^{d-1}\}$ is a basis of $K[a]$ $\square$

Definition 2.10.30. If $L \supset K$ is a module finite field extension we define define $[L : K] := \dim_K L$ to be *the degree of L over K*

Lemma 2.10.31. *Let $L \supset K$ be a field extension. Then $L \supset K$ is module finite if and only if $L \supset K$ is a finite field extension generated by some $a_1, \dots, a_n \in L$ that are algebraic over K.*

Proof. "$\Rightarrow$": By Lemma 2.10.21 $L = K[a_1, \dots, a_n]$ for suitable $a_1, \dots, a_n \in L$ that are algebraic over K, hence $L \subset K[a_1, \dots, a_n] \subset K(a_1, \dots, a_n) \subset L$, hence $L = K(a_1, \dots, a_n)$. "$\Leftarrow$": Suppose $L = K(a_1, \dots, a_n)$ for some $a_i \in L$. If $n = 1$, then $L = K(a_1) = K[a_1]$ which is module finite over K due to Proposition 2.10.17. Set $L' = K(a_1, \dots, a_n)$ which by induction is module finite over K. Note that a_{n+1} is algebraic over L', hence applying Proposition 2.10.17, $L \supset L'$ is module finite. By the transitive property of module finite extensions, it follows that $L \supset K$ is module finite. $\square$

Lemma 2.10.32. *Let $L \supset K$ be a field extension and $f \in K[x]$ irreducible. Suppose there is an $a \in L$ such that $f(a) = 0$. Then $L \simeq K[x]/I$ where $I := \langle f \rangle$.*

Proof. Consider the K-algebra map

$$\sigma := \mathrm{ev}_a : K[x] \to L$$

This induces an isomorphism

$$\overline{\sigma} : K[x]/\ker\ \sigma \simeq \mathrm{im}\ \sigma$$

$$\mu + \ker\ \sigma \mapsto \mu(a)$$

Since $K[x]$ is a PID, $\ker\ \sigma = \langle f' \rangle$ for some f' and since $f \in \ker\ \sigma$ it follows that $f' \mid f$, hence $\langle f \rangle = \langle f' \rangle$ by the irreducibility of f. Note that $K[x]/I$ is a field (cf. Lemma 2.9.52). Let $z = \frac{g(a)}{h(a)} \in K(a)$. Then since $h(a) \neq 0$, $f \nmid h$, hence $h + I \neq 0$. Then

$$z = \frac{g(a)}{h(a)} = \frac{\sigma(g)}{\sigma(h)} = \frac{\overline{\sigma}(g+I)}{\overline{\sigma}(h+I)} = \overline{\sigma}\left((g+I)(h+I)^{-1}\right) \in \mathrm{im}\ \overline{\sigma} = \mathrm{im}\ \sigma.$$

$\square$

Lemma 2.10.33. *Let $L \supset K$ be a field extension and $f \in K[x]$ an irreducible, monic polynomial. Suppose there is an $a \in L$ such that $f(a) = 0$. Set $L' := K[x]/I \simeq K(a)$, where $I := \langle f \rangle$*

1. *Suppose there is a $g \in k[x]$ that also vanishes on a. Then $f \mid g$.*

2. *identifying K canonically with a subfield of L' and $K(a)$ with L' we find $f = (y-(x+I))f_1$ for some $f_1 \in L[y]$.*

Proof. 1. From the proof of the last lemma we learned that $\mathbf{ev}_a(g) = 0$ if and only if $g \in I$, hence $f \mid g$.
2. Since $x+I$ is a zero of f in L the result follows. □

Theorem 2.10.34. *(Existence Theorem for Splitting Fields) Let K be a field and $f \in K[x]$. There is a field L extending K such that f can be written as a product of linear polynomials over L*

Proof. When $\mathbf{deg}\ f = 1$ the statement is trivial by taking $K = L$. If f is of of degree $d+1$ for some $d \geq 1$, pick a monic irreducible factor of f, g say. Then over $L' = K[x]/\langle g\rangle$, $g = (y-(x+I))g_1$ for some $g_1 \in L'[y]$. The $f = qg = qg_1(y-(x+I))$ for some $q \in L'[y]$ and the result follows by induction in the degree. □

Definition 2.10.35. The above L is called *the splitting field of f over K*.

Lemma 2.10.36. *Let K be a characteristic 0 field and $f \in K[x]$ irreducible monic. Let L be the splitting field of f over K and write $f = \prod_1^d (x-\alpha_i)$ for suitable $\alpha_i \in L$.*

Proof. Suppose L is a field extension over K, and suppose there is an $\alpha \in L$ such that $(x-\alpha^2) \mid f$. Then $g := Df$ also has α as a root. Then $g \nmid f$, hence by Lemma 2.10.33 f cannot be irreducible. In particular if L was the splitting field, and f has a multiple linear factor, then f is not irreducible. □

2.10.4 Theorem of the Primitive Element

Theorem 2.10.37. *Let K be a characteristic 0 field and $L \supset K$ a module-finite extension. Then there is a $c \in L$ such that $L = K(c)$.*

Proof. Suppose $L = K(a,b)$. Then there are monic irreducible polynomials $f,g \in K[x]$ such that $f(a) = 0$ and $g(b) = 0$. Let S be the splitting field of f and g. Write $f = (x-a)\prod_1^l (x-\alpha_i)$ and $g = (x-b)\prod_1^k (x-\beta_i)$. We may pick $\lambda \neq 0$ such that $c := \lambda a + b \neq \lambda\alpha_i + \beta_j$ for any i,j, since $V_{ij} := V(\alpha_i t + \beta_j - (at+b))$ can have at most finitely many points. So pick any $\lambda \notin \{0\} \cup \bigcup V_{ij}$. Set $K' := K(c)$ and $h := g(c-\lambda x) \in K'[x]$. Note that $h(a) = g(b) = 0$ and $h(\alpha_i) = g(\lambda a + b - \lambda\alpha_i)$ and since $\lambda a + b - \lambda\alpha_i \neq \beta_j$ for any j, $h(\alpha_i) \neq 0$. Then $\mathbf{gcd}(f,h) = x-a \in K'[x]$, implying $a \in K'$, and so $b = c - \lambda a \in K'$. In conclusion, $L = K(c)$.

Suppose $L = K(a_1,\dots,a_{n+1})$ for some $n \geq 1$. By induction there are $\lambda_1,\dots,\lambda_n \in K \setminus 0$, so that upon defining $c = \sum_1^n \lambda_i a_i$, $K(a_1,\dots,a_n) = K(c)$, hence $L = K(c,a) = K(c')$ by the first case. $\square$

2.10.5 Transcendence Degree & Transcendence Bases

Definition 2.10.38. Let $L \supset K$ be a field extension. We say that L has *transcendence degree d over K* if there is a set $X = \{a_1,\dots,a_d\} \subset L$ such that X is algebraically independent over K and every other set $Y \subset L$ with more than n elements is algebraically dependent over K. We define

$$\mathbf{trdeg}_K\ L := \mathbf{trdeg}\ L = d.$$

If there is not such d we write $\mathbf{trdeg}_K\ L = \infty$.
A finite field extension over K of transcendence degree n is called *an algebraic function field (over K) in n variables*

Remark 2.10.39. If $\delta < d$ is a positive integer such that there are $b_1,\dots,b_\delta$ that are algebraically independent, then $a_1,\dots,a_\delta$ are algebraically independent over K, hence the transcendence degree of L over K is unique.

Definition 2.10.40. Let $L \supset K$ be a field extension. A set $X = \{a_1,\dots,a_d\} \subset L$ is a *transcendence basis of L over K* if X is algebraically independent over K and $K(a_1,\dots,a_d) \supset K$ is algebraic.

Remark 2.10.41. When $L \supset K$ is algebraic, then $\emptyset$ is a transcendence basis of L over K.

Lemma 2.10.42. *Let $L \supset K$ be a field extension and $X = \{a_1,\dots,a_d\} \subset L$ be algebraically dependent. Consider an element $a \in L$. $X \cup \{a\}$ is algebraically dependent over K if and only if a is algebraic over $K(a_1,\dots,a_d)$. Therefor $a_1,\dots,a_d \in L$ forms a transcendence basis of L over K if and only if $a_1,\dots,a_d$ are algebraically independent over K and every $a \in L$ is algebraic over $K(a_1,\dots,a_d)$.*

Proof. "$\Rightarrow$": Let $f \in K[x_1,\dots,x_{d+1}] \setminus 0$ be given such that $f(a_1,\dots,a_d,a) = 0$. Then $f = \sum_0^m f_i x_{d+1}^i$, with $f_m \neq 0$. Since X is algebraically independent over K, this implies $f_m(a_1,\dots,a_d) \neq 0$. Then

$$g := f_m(a_1,\dots,a_d)^{-1} f(a_1,\dots,a_d,x)$$

is non-zero, monic and has a as a root, implying a is algebraic over $K(a_1,\dots,a_d)$.
"$\Leftarrow$": There is some $f = y^m + \sum_0^{m-1} b_i y^i \in K(a_1,\dots,a_d)[y] \setminus 0$ such that $f(a) = 0$.

Then $g := cf \in K[a_1,\dots,a_d][y]\setminus 0$, where c is a common denominator of the b_i's, hence

$$g = \sum_0^m g_i(a_1,\dots,a_d)y^i,$$

for suitable $g_i \in K[x_1,\dots,x_d]$, with $g_m = c \neq 0$. It follows that

$$h := \sum_0^m g_i y^i \in K[x_1,\dots,x_d,y],$$

such that $h(a_1,\dots,a_d,a) = g(a) = 0$, hence $X \cup \{a\}$ is algebraically dependent over K. □

Lemma 2.10.43. *Let $L \supset K$ and $L' \supset K$ be field extensions. Suppose there is a surjective K-algebra homomorphism $\sigma : L \to L'$. Then $\operatorname{trdeg}_K L \geq \operatorname{trdeg}_K L'$.*

Proof. Let $\alpha \in L'$ and set $n := \operatorname{trdeg} L$. Pick $\alpha_1,\dots,\alpha_n \in L'$ and pick $\beta,\beta_1,\dots,\beta_n \in L$ such that $\sigma(\beta),\sigma(\beta_i) = \alpha$. Pick $f \in K(\beta_1,\dots,\beta_n)[x]\setminus 0$ be monic such that $f(\beta) = 0$. Let $\overline{\sigma}$ denote the restriction of $L[x] \to L'[x]$, the induced K-algebra map induced by σ to a K-algebra map $K(\beta_1,\dots,\beta_n) \to K(\alpha_1,\dots,\alpha_n)$. Then

$$\overline{\sigma}(f)(\alpha) = \overline{\sigma}(f)(\sigma(\beta)) = \sigma(f(\beta)) = \sigma(0) = 0,$$

and since $\overline{\sigma}(f)$ is monic, this shows that α is algebraic over $K(\alpha_1,\dots,\alpha_n)$. So every sequence of n elements in L' are algebraically dependent over K by the prior lemma. It follows that $\operatorname{trdeg}_K L' \leq n = \operatorname{trdeg}_K L$. □

Remark 2.10.44. Note that the assumption that X is algebraically independent over K is not necessary to prove "$\Leftarrow$".

Lemma 2.10.45. *Let $L \supset K$ be a finite field extension generated by $a_1,\dots,a_d \in L$.*

1. *There is a subset of $X := \{a_1,\dots,a_r\}$ that is a transcendence basis for L over K.*

2. *Let $Y = \{a \in X : a$ is transcendental over $K(b_1,\dots,b_s)\}$. If $b_1,\dots,b_s \in L$ are algebraically independent over K, then for some $Z \subset Y$, $\{b_1,\dots,b_s\} \cup Y$ is a transcendence basis of L over K*

Proof. 1. If no subset of X is algebraically independent over K, then each a_i is algebraic over K, hence L is algebraic over K. This means that $\emptyset$ is a transcendence basis for L over K.

Suppose now that there is a subset of $\{a_1,\dots,a_r\}$ whose elements are algebraically independent over K. Let $Y \subset \{a_1, \ldots, a_r\}$ be a maximal subset of elements that are algebraically independent over K. After a permutation, we can write $Y = \{a_1, \ldots, a_k\}$ for some $k < r$. Then $a_1, \ldots, a_k, a_i$ are algebraically dependent over K for each $i \in \{k+1, \ldots, r\}$. Then by Lemma 2.10.42, a_i is algebraic over $K(a_1, \ldots a_k)$ for each $i \in \{k+1, \ldots, r\}$. Then L is algebraic over $K(a_1, \ldots, a_k)$, implying Y is a transcendence basis of L over K.

2. We proceed by induction in k: if a_i is algebraic over $K(b_1, \ldots, b_s)$ for every $i \in \{1, \ldots, r\}$, then L is algebraic over $K(b_1, \ldots, b_s)$, hence $\{b_1, \ldots, b_s\}$ is a transcendence basis for L over K by Lemma 2.10.42.

Suppose the statement is true for some $k < n$. Consider WLOG $Y = \{a_1, \ldots, a_{k+1}\}$. By Lemma 2.10.42 $b_1, \ldots, b_s, a_1$ are algebraically independent over K. Every element in $X \backslash Y$ is algebraic over $K(b_1, \ldots, b_s, a_1)$, hence $Y' := \{a \in X : a$ is transcendental over $K\} \subset Y$. It thus follows by induction that for some $Z' \subset Y'$,

$$\{b_1, \ldots, b_s\} \cup \underbrace{\{a_1\} \cup Z'}_{=:Z}$$

is a transcendence basis of L over K. □

Theorem 2.10.46. *Let $L \supset K$ be a field extension. L has a transcendence basis $X = \{a_1, \ldots, a_d\}$ if and only if* $\mathrm{trdeg}_K\, L = d$

Proof. "$\Rightarrow$": **To prove this we make the following claim**: L is algebraic over $K(b_1, \ldots, b_k, a'_{k+1}, \ldots, a'_d)$ for any $k \in \{0, \ldots, d\}$ for some subset $\{a'_{k+1}, \ldots, a'_d\} \subset \{a_1, \ldots, a_d\}$ with $d - k$ elements. We prove this using induction in r. For $k = 0$, we have that $L \supset K(a_1, \ldots, a_d)$ is algebraic by the assumption that $\{a_1, \ldots, a_d\}$ is a transcendence basis for L over K. Suppose that we the statement holds for some $k \in \{0, \ldots, d-1\}$. Then by induction hypothesis L is algebraic over $M := K(b_1, \ldots, b_k, a'_{r+1}, \ldots, a'_d)$ for a suitable subset $\{a'_{k+1}, \ldots, a'_d\} \subset \{a_1, \ldots, a_d\}$. This means b_{k+1} is algebraic over M, hence by Lemma 2.10.42 $b_1, \ldots, b_{k+1}, a'_{r+1}, \ldots, a_d$ are algebraically dependent over K. By Lemma 2.10.45 there is an integer $r \leq s < d$ and a subset $\{a''_{k+1}, \ldots, a''_{s+1}\} \subset \{a'_{k+1}, \ldots, a'_d\}$ such that $b_1, \ldots, b_{k+1}, a''_{k+2}, \ldots, a''_{s+1}$ are algebraically independent over K and $b_1, \ldots, b_{r+1}, a''_{k+1}, \ldots, a''_{s+1}$ are algebraically dependent over K. Again by Lemma 2.10.42, we have that a''_{k+1} is algebraic over $M' := K(b_1, \ldots, b_{k+1}, a''_{k+2}, \ldots, a''_d)$. Now $L \supset M(b_{r+1})$ is algebraic and $M'(a''_{r+1}) \supset M'$ is algebraic. Since $M(b_{r+1}) = M'(a''_{r+1})$ we find that L is algebraic over M', finishing the proof of the claim. **Application of the claim:** Suppose for a contradiction that there are $b_1, \ldots, b_{d+1} \in L$ that are algebraically independent over K. Then $b_1, \ldots, b_d$ are algebraically independent over K. But then by the claim $L \supset K(b_1, \ldots, b_d)$ is algebraic, hence b_{d+1} is

algebraic over $K(b_1,\ldots,b_d)$, but then $b_1,\ldots,b_{d+1}$ are algebraically dependent over K.
"$\Leftarrow$": There are algebraically independent $a_1,\ldots,a_d \in L$ and for every $a \in L$ $a,a_1,\ldots,a_d$ are algebraically dependent, hence by Lemma 2.10.42 $\{a_1,\ldots,a_d\}$ is a transcendence basis of L over K. $\square$

Remark 2.10.47. Note the claim of "$\Leftarrow$", simply proves that every algebraically independent $b_1,\ldots,b_d$ forms a transcendence basis of L over K

Corollary 2.10.48. *Two transcendence bases have the same cardinality.*

Example 2.10.49. Consider $L = K(x_1,\ldots,x_n) = Q(K[x_1,\ldots,x_n])$. The elements $x_1,\ldots,x_n \in K[x_1,\ldots,x_n]$ are algebraically independent over K, hence $x_1,\ldots,x_n \in K(x_1,\ldots,x_n)$ are algebraically independent over K. It follows that $\{x_1,\ldots,x_n\}$ is a transcendence basis of $K(\mathbf{x}) \supset K$, hence $\mathbf{trdeg}_K\ K(\mathbf{x}) = n$.

Corollary 2.10.50. *Consider a tower of field extensions $M \supset L \supset K$. Let $X = \{a_1,\ldots,a_n\} \subset L$ be a transcendence basis for L over K. Suppose M is finitely generated L-module (in other words a finite dimensional vector space over L). Then X is transcendence basis for M over K and hence $trdeg_K\ M = trdeg_K\ L$.*

Proof. We need to check that M is algebraic over $K(a_1,\ldots,a_n)$. Note that $M \supset L$ being module-finite, implies $M \supset L$ is algebraic (cf. Lemma 2.10.21), hence $M \supset K(a_1,\ldots,a_n)$ is algebraic. $\square$

Lemma 2.10.51. *Let $L \supset K$ and $M \supset K$ be field extension with an injective K-algebra homomorphism $\sigma : L \hookrightarrow M$. If $a_1,\ldots,a_n \in L$ are algebraically independent over K, then so are $\sigma(a_1),\ldots,\sigma(a_n)$. It follows that* $\mathbf{trdeg}_K\ L \le \mathbf{trdeg}_K\ M$.

Proof. Let $a_1,\ldots,a_n \in L$ be algebraically independent over K. Let $f \in K[x_1,\ldots,x_n] \setminus 0$. Note that $\sigma(k) = k$ for every $k \in K$ and $f(a_1,\ldots,a_n) \neq 0$ by the assumption that $a_1,\ldots,a_n$ are algebraically independent over K. It follows that

$$f(\sigma(a_1),\ldots,\sigma(a_n)) = \sigma(f(a_1,\ldots,a_n)) \neq 0,$$

hence $\sigma(a_1),\ldots,\sigma(a_n)$ are algebraically independent. Hence for any $n \le \mathbf{trdeg}\, L$ there are algebraically independent $b_1,\ldots,b_n \in M$, hence $\mathbf{trdeg}\, L \le \mathbf{trdeg}\, M$. $\square$

Remark 2.10.52. Note that $\mathbf{trdeg}_K$ is an integer invariant (cf. Example 1.3.19) of the category with objects being field extensions over K and morphisms being injective K-algebra homomorphism.

Lemma 2.10.53. *Let $M \supset L \supset K$ be a tower of field extensions.*

$$\mathbf{trdeg}_K\ M = \mathbf{trdeg}_K\ L + \mathbf{trdeg}_L\ M.$$

If L has transcendence basis $X = \{a_1, \ldots, a_d\}$ over K and M has transcendence basis $Y = \{b_1, \ldots, b_\delta\}$ over L, then $X \cup Y$ is a transcendence basis for M over K. It follows that

Proof. If $\mathbf{trdeg}_L\ M = \infty$, then $\mathbf{trdeg}_K\ M = \infty$. Suppose this is not the case. Then we may assume that L has transcendence basis $X = \{a_1, \ldots, a_d\}$ over K and M has transcendence basis $Y = \{b_1, \ldots, b_\delta\}$ over L. Then $M \supset L(b_1, \ldots, b_\delta)$ is algebraic and $L \supset K(a_1, \ldots, a_d)$ is algebraic, hence

$$L(b_1, \ldots, b_\delta) \supset K(a_1, \ldots, a_d)(b_1, \ldots, b_\delta) = K(a_1, \ldots, a_d, b_1, \ldots, b_\delta)$$

is algebraic. It follows that $M \supset K(a_1, \ldots, a_d, b_1, \ldots, b_\delta)$ is algebraic. Let $f \in K[x_1, \ldots, x_d, y_1, \ldots, y_\delta] \setminus 0$. Write

$$f = \sum_{v \in \mathbb{N}^\delta} f_v \mathbf{y}^v.$$

For some v, $f_v \neq 0$, hence since $a_1, \ldots, a_d$ are algebraically independent over K, $f_v(a_1, \ldots, a_d) \neq 0$. Then

$$g := f(a_1, \ldots, a_d, \mathbf{y}) = \sum_{v \in \mathbb{N}^\delta} f_v(a_1, \ldots, a_d)\mathbf{y}^v \in L[\mathbf{y}] \setminus 0.$$

Since $b_1, \ldots, b_\delta$ are algebraically independent over L, it follows that

$$f(a_1, \ldots, a_d, b_1, \ldots, b_\delta) = g(b_1, \ldots, b_\delta) \neq 0,$$

hence $a_1, \ldots, a_d, b_1, \ldots, b_\delta$ are algebraically independent over K. Hence $X \sqcup Y$ is a transcendence basis for M over K, hence

$$\mathbf{trdeg}_K\ M = d + \delta = \mathbf{trdeg}_K\ L + \mathbf{trdeg}_L\ M.$$

□

Lemma 2.10.54. *Let $L \supset K$ be an algebraic function field in one variable with K algebraically closed. Let $a \in L \setminus K$. Then*

1. *$L \supset K(a)$ is algebraic*

2. *Suppose* $\mathbf{char}\ K = 0$*. Then there is a $b \in L$ such that $L = K(a, b)$.*

3. *Consider an integral domain R with $Q(R) = L$, $K \subset R$ algebraically closed. Suppose there is a non-trivial prime ideal $I \subset R$. Then $\sigma : K \to R/I, a \mapsto a + I$ is an isomorphism.*

Proof. 1. L is algebraic over $K(t)$ for some $t \in L$. Then a is algebraic over $K(t)$, hence we may find $f \in K[x,y] \setminus 0$ such that $f(a,t) = 0$. Note that since $a \notin K$, a cannot be algebraic over K (Lemma 2.10.22). Then $\deg_y f > 0$, hence $g = f(a,y) \neq 0$ is polynomial that vanishes on t, hence $K(a,t) \supset K(a)$ is algebraic and since $L \supset K(a,t)$ is algebraic, it follows that $L \supset K(a)$ is algebraic.
2. Since $L \supset K(a)$ is algebraic it is finite, hence by the Theorem of the Primitive Element $L = K(a,b)$ for some b.
3. We prove the contrapositive: Let $I \supset R$ be any prime ideal. Suppose $\sigma : K \hookrightarrow R/I$ is not surjective. Pick $a \in R$ such that $a + I \in R/I$ is not in K and pick $b \in I$. Note that $b = 0$ or $b \in R \setminus K$, since otherwise $1 \in I$. Then since $L \supset K(a)$ is algebraic, we can find a $f = \sum_0^d g_i(a) y^i \in K[a][y] \setminus 0$ of minimal degree such that $f(b) = 0$. By Lemma 2.10.22 $g_0 = 0$ or $g_0(a) \neq 0$. In the first case we clearly have that $b = 0$, since then $f = y(\sum_1^d g_i(a) y^{i-1})$, hence $f = y$ by minimality. In the second case $f_0(a) \in I$, hence $a + I$ is integral over K which would imply that $a + I \in K$ by Lemma 2.10.22, leading to a contradiction. Since b was chosen arbitrarily it follows that $I = 0$. □

2.10.6 Graph Ideals & Algebraic Dependence of Polynomials

Definition 2.10.55. Consider a ring R. The *graph ideal* for polynomials $f_1, \ldots, f_m \in R[x_1, \ldots, x_n]$ is defined to be ideal $\langle y_1 - f_m, \ldots, y_m - f_m \rangle \subset R[\mathbf{x}, y_1, \ldots, y_m]$

Remark 2.10.56. Note that the graph ideal of $f_1, \ldots, f_m$ is just the point ideal (cf. Proposition 2.9.38) of $(x_1, \ldots, x_n, f_1, \ldots, f_m) \in R[\mathbf{x}, \mathbf{y}]^{n+m}$. Hence a polynomial $f \in K[\mathbf{x}, \mathbf{y}]$ is in the graph ideal of $f_1, \ldots, f_m$ if and only if $f(\mathbf{x}, f_1, \ldots, f_m) = 0$.

Lemma 2.10.57. *Let K be a field. Consider $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$ and denote the graph ideal of $f_1, \ldots, f_m$ by I. Then I is a proper ideal by Lemma 2.4.38*

Proof. Since 1 doesn't vanish on $(\mathbf{x}, f_1, \ldots, f_m)$ it follows that $1 \notin I$. Hence I is proper by Lemma 2.4.38 □

Proposition 2.10.58. *Consider $f_1, \ldots, f_{n+1} \in K[x_1, \ldots, x_n]$ and let I denote the graph ideal in $K[\mathbf{x}, y_1, \ldots, y_{n+1}]$ for these polynomials. Let $G \subset K[\mathbf{x}, \mathbf{y}]$ be a Gröbner basis for I with respect to the lexicographic term order with $x_1 > \cdots > x_n > y_1 > \cdots > y_{n+1}$. Then there is a non-zero polynomial $g \in G \cap K[\mathbf{x}]$ such that $g(f_1, \ldots, f_{n+1}) = 0$.*

Proof. By Example 2.10.49 there is a polynomial $h \in K[\mathbf{y}]\backslash 0$ such that $h(f_1,\dots,f_{n+1}) = 0$. Then $h \in I$ and in particular $h \in I \cap K[\mathbf{y}]$. By Proposition 2.9.91, $G' := G \cap K[\mathbf{y}]$ is a Gröbner basis for $I \cap K[\mathbf{y}]$ with respect to the lexicographic term order with $y_1 < \cdots < y_{n+1}$. Then $h^{G'} = 0$ by Proposition 2.9.69, meaning G' must contain a non-zero polynomial g. Since $g \in I$ we again by Lemma 2.9.91 get that $g(f_1,\dots,f_{n+1}) = 0$. □

2.10.7 Finite Algebra Homomorphisms

Definition 2.10.59. Let S,T be R-algebras. An R-algebra homomorphism, $\sigma : S \to T$ is called *finite* if $T \supset \sigma(S)$ is module finite.

Lemma 2.10.60. *Let S,T,Q be R-algebras and $\sigma : S \to T$ and $\omega : T \to Q$ be finite R-algebra homomorphisms. Then $\omega\sigma : S \to Q$ is finite.*

Proof. For some $t_1,\dots,t_m \in T$ and $q_1,\dots,q_n \in Q$ we have $S = \sum_1^m \sigma(S)t_i$ and $Q = \sum_1^k \omega(T)q_i$. We Then get that

$$= \sum_1^k \omega(T)q_i = \sum_1^k \omega\left(\sum_1^m \sigma(S)t_j\right)q_i = \sum_1^k\sum_1^m (\omega\circ\sigma)(S)\omega(t_j)q_i,$$

hence Q is a finitely generated over $(\alpha\circ\beta)(R)$ with generators

$$\omega(t_j)q_i,\ (1 \le i \le k, 1 \le j \le m).$$

Therefor, we can conclude that $\omega\circ\sigma$ is finite. □

Lemma 2.10.61. *Let S,T be R-algebras and $\sigma : S \to T$ be a surjective R-algebra homomorphism. Then σ is finite.*

Proof. Trivial since $T = \sigma(S)$. □

2.10.8 Perron's Theorem of Effective Algebraic Dependence of Polynomials

Lemma 2.10.62. *Let K be any field and $d_1,\dots,d_n > 0$, $\mathcal{S} \subset \mathbb{N}^n$ containing $d_ie_i \in \mathcal{S}$ for each i. Set $L := K(y_v^{[i]} : v \in \mathcal{S}) = Q(K[y_v^{[i]} : v \in \mathcal{S})$. For each $i \in \{1,\dots,n\}$, set*

$$g_i := \sum_{v\in\mathcal{S}} y_v^{[i]}\mathbf{x}^v \in L[x_1,\dots,x_n]$$

and $d_i := \deg\ g_i$ Let $N \ge 0$ be given. Define $\Delta := \{v \in \mathbb{N}^n : |v| \le N\}$. Then

$$B := \{g_1^{q_1}\cdots g_n^{q_n} x_1^{r_1}\cdots x_n^{r_n} : 0 \le r_i < d_i, \sum_1^n q_id_i + r_i \le N\}$$

is a basis for $L[\mathbf{x}]_{\le N}$ over K.

Proof. For each $v=(v_1,\dots,v_n)\in\Delta$ there are unique pair of tuples

$$(q_1(v_1),\dots,q_n(v_n)),(r_1(v_1),\dots,r_n(v_n))\in\mathbb{N}^n$$

such that for each $i\in\{1,\dots,n\}$, $0\le r_i<d_i$ and $v=(q_1d_1+r_1,\dots,q_nd_n+r_n)$. We $\nabla=\{(q_1,\dots,q_n),(r_1,\dots,r_n)\in\mathbb{N}^n:0\le r_i<d_i,\sum_1^n(q_id_i+r_i)\le N\}$. We thus have that

$$(q,r):\Delta\to\nabla$$
$$v=(v_1,\dots,v_n)\mapsto((q_1(v_1),\dots,q_n(v_n)),(r_1(v_1),\dots,r_n(v_n)))$$

defines a bijection. We define for each $v\in\Delta$,

$$\Lambda_v:=\Lambda_{q(v),r(v)}:=\left(\prod_1^n g_i^{r_i(v_i)}\right)\left(\prod_1^n x_i^{r_i(v_i)}\right)\in K[\mathbf{x}][y_v^{[i]}:v\in\mathcal{S}].$$

We thus have that $B=\{\Lambda_v:v\in\Delta\}$. Note that $\deg\ \Lambda_v=|v|$ for each $v\in\Delta$, which for one means that $\Lambda_v\in L[\mathbf{x}]_{\le N}$. Let $\sigma:K[y_v^{[i]}:v\in\mathcal{S}]\to K$ be the unique K-algebra homomorphism defined by $y_v^{[i]}\mapsto 0$ when $(v,i)\ne(d_1e_1,1)\dots,(d_ne_n,n)$ and $y_{d_ie_i}^{[i]}\mapsto 1$. This map naturally extends to a $K[\mathbf{x}]$-algebra homomorphism which we also denote σ. Then $\sigma(g_i)=\mathbf{x_i}^{d_i}$ and $\sigma(\mathbf{x}^v)=\mathbf{x}^v$, hence

$$\sigma(\Lambda_v)=x_1^{q_1(v_1)d_1+r_1(v_1)}\cdots x_n^{q_n(v_n)d_n+r_n(v_n)}=\mathbf{x}^v. \tag{5}$$

Write for each $v\in\Delta$

$$\Lambda_v=\sum_{v\in\Delta}c_{vw}\mathbf{x}^w.$$

By (5)

$$\sigma(c_{vw})=\begin{cases}1 & \text{if } w=v\\ 0 & \text{if } w\ne v\end{cases}$$

Let D denote $\#\Delta=\frac{N(N+1)}{2}$. σ naturally induces a homomorphism

$$\sigma:M_D(K[y_v:v\in\mathcal{S}])\to M_D(K)\subset M_D(K[y_v:v\in\mathcal{S}])$$

defined by entry-wise application for which $\sigma((c_{vw}))=(\sigma(c_{vw}))=(e_{vw})=I_D$. Set $\mathcal{V}$ to be equal to $\{\mathbf{x}^v:v\in\Delta\}$; i.e. the standard basis for $L[\mathbf{x}]_{\le N}$ over L. Then ${}_{\mathcal{V}}T_B=(c_{vw})\in M_D([y_v:v\bigcup\mathcal{S}_i])$. Moreover,

$$\sigma(\det\ {}_{\mathcal{V}}T_B)=\det\ \sigma({}_{\mathcal{V}}T_B)=\det\ I_D=1\ne0\Rightarrow\det\ {}_{\mathcal{V}}T_B\ne0.$$

This means ${}_{\mathcal{V}}T_B$ is invertible in $M_D(L)$, hence B is a basis by Theorem 2.7.3. □

Remark 2.10.63. In the above setup we can therefor given any $f \in K[\mathbf{x}]$ find a family of polynomials

$$f_{r_1,\dots,r_n} \in L[z_1,\dots,z_n] \quad (0 \le r_i < d_i),$$

such that

1. $f = \sum_{r_1,\dots,r_n} f_{r_1,\dots,r_n}(g_1,\dots,g_n)x_1^{r_1}\cdots x_n^{r_n}$
2. $\deg\ f_{r_1,\dots,r_n}(z_1^{d_1},\dots,z_n^{d_n}) + \sum_1^n r_i \le \deg\ f.$

Indeed set $N := \deg f$ and write

$$f = \sum_{v\in\Delta} a_v\Lambda_v = \sum_{(r_1,\dots,r_n)\in\mathbb{N}^n : r_i<d_i} \left[\sum_{(q_1,\dots,q_n):\sum_1^n(q_id_i+r_i)\le N} a_{(q_1d_1+r_1,\dots,q_nd_n+r_n)} g_1^{q_1}\cdots g_n^{q_n} \right] x_1^{r_1}\cdots x_n^{r_n}.$$

Setting

$$f_{r_1,\dots,r_n} := \sum_{(q_1,\dots,q_n)\in\mathbb{N}^n:\sum_1^n(q_id_i+r_i)\le N} a_{(q_1d_1+r_1,\dots,q_nd_n+r_n)} z_1^{q_1}\cdots z_n^{q_n} \quad (0\le r_i<d_i)$$

These polynomials will satisfy property 1. Secondly,

$$\begin{aligned} \deg\ f_{r_1,\dots,r_n}(z_1^{d_1},\dots,z_n^{d_n}) + \sum_1^n r_i &\le \max_{(q_1,\dots,q_n):\sum_1^n(q_id_i+r_i)} \deg\ z_1^{q_1d_1}\cdots z_n^{q_nd_n} + \sum_1^n r_i \\ &= \max_{(q_1,\dots,q_n):\sum_1^n(q_id_i+r_i)} \sum_1^n (q_id_i+r_i) \le N = \deg\ f. \end{aligned}$$

Lemma 2.10.64. *Let K be any field and $d_1,\dots,d_n > 0$, $\mathcal{S} \subset \mathbb{N}^n$ containing $d_ie_i \in \mathcal{S}$ for each i. Set $L := K(y_v^{[i]} : v \in \mathcal{S}) = Q(K[y_v^{[i]} : v \in \mathcal{S})$. For each $i \in \{1,\dots,n\}$, set*

$$g_i := \sum_{v\in\mathcal{S}} y_v^{[i]}\mathbf{x}^v \in L[x_1,\dots,x_n]$$

and $d_i := \deg\ g_i$ Then for every $g_{n+1} \in L[\mathbf{x}]$ with $d_{n+1} := \deg\ g_{n+1}$ there is polynomial $P \in L[z_1,\dots,z_{n+1}]$ that is monic in $L[z_1,\dots,z_n][z_{n+1}]$ satisfying

1. $P(g_1,\dots,g_{n+1}) = 0$
2. $\deg\ P(z_1^{d_1},\dots,z_{n+1}^{d_{n+1}}) \le \prod_1^{n+1} d_i.$

Proof. There is $d := \prod_1^n d_i$ elements in $\Omega = \{v \in \mathbb{N}^n : v_i < d_i\}$. Denote the elements in $\{\mathbf{x}^v : v \in \Omega\} = \{M_1,\dots,M_d\}$. Then by Lemma 2.10.62 for each $i \in \{1,\dots,d\}$ there are polynomials

$$P_{ij} \in L[\mathbf{x}] \quad (j \in \{1,\dots d\})$$

such that

a. $M_i g_{n+1} = \sum_1^d P_{ij}(g_1,\dots,g_n) M_i$

b. $\deg\ P_{ij}(z_1^{d_1},\dots,z_n^{d_n}) + \deg M_i \le \deg\ g_{n+1} M_i = d_{n+1} + \deg\ M_i$.

Property a. shows that

$$g_{n+1}\begin{pmatrix} M_1 \\ \vdots \\ M_d \end{pmatrix} = (P_{ij}(g_1,\dots,g_n))\begin{pmatrix} M_1 \\ \vdots \\ M_d \end{pmatrix},$$

I.e. g_{n+1} is an eigenvalue of $(P_{ij}(g_1,\dots,g_n))$, hence by Cramer's rule,

$$\det(P_{ij}(g_1,\dots,g_n) - \delta_{ij} g_{n+1}) = 0$$

. Then $P := (-1)^d \det(P_{ij} - \delta_{ij} g_{n+1}) = \sum_{\pi \in S_d} \prod_1^d (P_{i\pi(i)} - \delta_{ij} g_{n+1}) \in L[\mathbf{x}] \setminus 0$ satisfies $P(g_1,\dots,g_{n+1}) = 0$. From b. we find that

$$\deg\ P_{ij}(z_1^{d_1},\dots,z_n^{d_n}) - \delta_{ij} z_{n+1}^{d_{n+1}} \le d_{n+1} + \deg\ M_i - \deg\ M_j.$$

For an arbitrary permutation $\pi \in S_d$,

$$\begin{aligned}\deg \prod_1^d (P_{i\pi(i)}(z_1^{d_1},\dots,z_d^{d_n}) - \delta_{i\pi(i)} z_{n+1}^{d_{n+1}}) &\le \sum_1^d (d_{n+1} + \deg\ M_{\pi(i)} - \deg\ M_i) \\ &= \sum_1^d d_{n+1} + \sum_1^d \deg\ M_i - \sum_1^d \deg\ M_i \\ &= d d_{n+1} = \prod_1^{n+1} d_i.\end{aligned}$$

□

Lemma 2.10.65. *Let K be some field. For $f_1,\dots,f_n \in K[x_1,\dots,x_n]$ with $d_i := \deg\ f_i > 0$, and*

$$f_i = \sum_{v \in \mathbb{N}^n} a_v^{[i]} \mathbf{x}^v.$$

Define $L := K\left(y_v^{[i]} : v \in \mathbb{N}^n, a_v^{[i]} \neq 0 \text{ or } v = d_i e_i\right)$ and set $Y := \left\{y_v^{[i]} : v \in \mathbb{N}^n, a_v^{[i]} \neq 0 \text{ or } v = d_i e_i\right\}$. Lastly define

$$g_i := \sum_{v \in \mathbb{N}^n} y_v^{[i]} \mathbf{x}^v$$

for every i. Then there is a $K[\mathbf{x}]$-algebra homomorphism $\sigma : K[Y][\mathbf{x}] \to K[\mathbf{x}]$ such that $\sigma(g_i) = f_i$

Proof. Indeed, take σ to be the K-algebra homomorphism such that $y_v^{[i]} \mapsto a_v^{[i]}$. This trivially extends to a $K[\mathbf{x}]$-algebra homomorphism. $\square$

Theorem 2.10.66. *(Perron's Theorem) Let K be any field and let $f_1, \ldots, f_{n+1} \in K[x_1, \ldots, x_n]$ and put $d_i := \deg f_i$ for each i. Then there is a $P \in K[z_1, \ldots, z_{n+1}] \setminus 0$ satisfying*

1. $P(f_1, \ldots, f_{n+1}) = 0$,
2. $\deg P(z_1^{d_1}, \ldots, z_{n+1}^{d_{n+1}}) \leq \prod_1^{n+1} d_i$.

Proof. First a slight reformulation. Let M be some field and consider $H = \{h_1, \ldots, h_{n+1}\} \subset M[x_1, \ldots, x_n]$, $\delta_i := \deg h_i$. Set

$$\Delta_{\delta_1, \ldots, \delta_{n+1}} := \left\{ v \in \mathbb{N}^{n+1} : \sum_1^{n+1} v_i \delta_i \leq \prod_1^{n+1} \delta_i \right\}.$$

Then define

$$B(H) := \left\{ h_1^{q_1} \cdots h_{n+1}^{q_{n+1}} : (q_1, \ldots, q_{n+1}) \in \Delta_{\delta_1, \ldots, \delta_{n+1}} \right\}.$$

And let $\mathcal{V}$ be the standard basis of $\{\mathbf{z}^v : v \in \Delta_{\delta_1, \ldots, \delta_n}$. If ${}_{\mathcal{V}}T_{B(H)}$ is not invertible if and only if for suitable $a_v \in M$

$$\sum_{v \in \Delta} a_v h_1^{v_1} \cdots h_{n+1}^{v_{n+1}} = \mathrm{ev}_{h_1, \ldots, h_{n+1}} \left(\underbrace{\sum_{v \in \Delta} a_v \mathbf{z}^v}_{P_H} \right),$$

where

$$\deg P_H(z_1^{\delta_1}, \ldots, z_{n+1}^{\delta_{n+1}}) \leq \max_{v \in \Delta} \deg \mathbf{z}^{(\delta_1 v_1, \ldots, \delta_{n+1} v_{n+1})} = \max_{v \in \Delta} \sum_1^{n+1} \delta_i v_i \leq \prod_1^{n+1} \delta_i.$$

Set $F = \{f_1, \ldots, f_{n+1}\}$ and $B := \{\mathbf{z}^v : v \in \Delta_{d_1, \ldots, d_{n+1}}\}$. To prove the theorem we can equivalently prove that $\det {}_B T_{B(F)} = 0$. With this in mind, we proceed with the proof of the theorem. Write $f_i = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v$. Define L, g_i and σ as in the prior lemma. By Lemma 2.10.64, there is a $Q \in L[\mathbf{x}] \setminus 0$ such that

1. $Q(g_1, \ldots, g_n, f_{n+1}) = 0$,
2. $\deg Q(z_1^{d_1}, \ldots, z_{n+1}^{d_{n+1}}) \leq \prod_1^{n+1} d_i$.

Set $G := \{g_1, \ldots, g_n, f_{n+1}\}$. By small easy lemma $\sigma({}_B T_{B(G)}) = {}_B T_{B(F)}$. It thus follows that

$$\det {}_B T_{B(F)} = \det \sigma({}_B T_{B(G)}) = \sigma(\det {}_B T_{B(G)}) = \sigma(0) = 0.$$

$\square$

2.10.9 Noether Normalizations

Lemma 2.10.67. *Let K be a field and let $f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v \in K[x_1,\dots,x_n]$ be given with $d := \deg f > 0$. Then we have the following*

1. *There are elements $y_1,\dots,y_{n-1} \in K[x_1,\dots,x_n]$ such that $x_i = y_i + x_n^{r_i}$ for $i \in \{1,\dots,n-1\}$ for suitable $r_i > 0$ and*
$$f = ax_n^m + \sum_i^{m-1} G_i x_n^i,$$
for some $a \in K\setminus\{0\}$, $m > 0$ and $G_i \in K[y_1,\dots,y_{n-1}]$.

2. *If $\#K = \infty$ we get the same result as in (a) with $x_i = y_i + a_i x_n$ for*

Proof. 1. Set $k = d+1$ and put $r_i = k^i$ for $i \in \{1,\dots,n-1\}$. Then for $v,w \in \mathbb{N}^n$ with $|v|,|w| \le d$ and $v \neq w$ we get
$$v_n + \sum_1^{n-1} v_i r_i = v_n + \sum_1^{n-1} v_i k^i \neq w_n + \sum_1^{k} w_i k^i = w_n + \sum_1^{n} w_i r_i,$$
by the uniqueness of k-adic expansions. This mean that we can define
$m := \max_{v\in\mathbb{N}^n:|v|\le d, a_v\neq 0}\left\{v_n + \sum_1^{n-1} v_i r_i\right\}$, $v_m = \operatorname{argmax}_{v\in\mathbb{N}^n:|v|\le d, a_v\neq 0}\left\{v_n + \sum_1^{n-1} v_i r_i\right\}$.
We thus pick $y_i = x_i - x_n^{r_i}$ and see that
$$f = f\left(y_1 + x_n^{r_1},\dots,y_{n-1}+x_n^{r_{n-1}},x_n\right) = \sum_{v\in\mathbb{N}^n:|v|\le d} a_v \left(\prod_1^{n-1}\left(y_i + x_n^{r_i}\right)^{v_i}\right)x_n^{v_n}$$
$$= \sum_{v\in\mathbb{N}^n:|v|\le d} a_v \left(x_n^{v_n + \sum_1^{n-1} v_i r_i} + \underbrace{\cdots}_{\text{lower degree terms}}\right) = \sum_{v\in\mathbb{N}^n:|v|\le d} a_v x_n^{v_n+\sum_1^{n-1} v_i r_i} + \underbrace{\cdots}_{\text{lower degree terms}}$$
$$= a_{v_m} x_n^m + \underbrace{\cdots}_{\text{lower degree terms}}.$$
We can write the lower order terms on the form $\sum_1^{m-1} G_i x_n^i$ for suitable $G_i \in K[y_1,\dots,y_{n-1}]$.
2. Write $f = \sum_0^d f_i$ with $f_d \neq 0$ for homogeneous $f_i \in K[x_1,\dots,x_n]$ of degree i. By Lemma 2.9.136, $f_d(x_1,\dots,x_{n-1},1) \neq 0$. Since $\#K = \infty$ we get that there are $a_1,\dots,a_{n-1} \in K$ such that $f_d(a_1,\dots,a_{n-1},1) \neq 0$. We now pick $y_i = x_i - a_i x_i$ for $i \in \{1,\dots,n-1\}$. Note that
$$\sum_{v\in\mathbb{N}^n:|v|=d} a_v (a_1x_n)^{v_1}\cdots(a_{n-1}x_n)^{v_{n-1}} x_n^{v_n} = \left[\sum_{v\in\mathbb{N}^n:|v|=d} a_v a_1^{v_1}\cdots a_{n-1}^{v_{n-1}}\cdot 1^{v_n}\right] x_n^d = f_d(a_1,\dots,a_{n-1},1)x_n^d.$$

From this it follows that

$$
\begin{aligned}
f &= f(y_1 + a_1x_n, \dots, y_{n-1} + a_{n-1}x_n, x_n) \\
&= \sum_{v\in\mathbb{N}^n:|v|=d} a_v(y_1 + a_1x_n)^{v_1} \cdots (y_{n-1} + a_{n-1}x_n)^{v_{n-1}} x_n^{v_n} + \sum_1^{m-1} \sigma(F_i) \\
&\overset{(*)}{=} \left[\sum_{v\in\mathbb{N}^n:|v|=d} a_v a_1^{v_1} \cdots a_{n-1}^{v_{n-1}} \cdot 1^{v_n} \right] x_n^d + \cdots = f_d(a_1, \dots, a_{n-1}, 1)x_n^d + \dots.
\end{aligned}
$$

We then set $a = f_d(a_1,\dots,a_{n-1},1)$. The $\dots$ in the expressions following $(*)$ in the above signify remaining terms. One readily verifies that these have x_n-degree strictly smaller than d. Again these terms can clearly be written on the form $\sum_1^{d-1} G_i x_n^i$ for suitable $G_i \in K[y_1,\dots,y_{n-1}]$. ☐

Theorem 2.10.68. *(Noether Normalization Theorem) Let $A = K[x_1,\dots,x_n]/J$ for some field K and some ideal $J \subsetneq K[x_1,\dots,x_n]$. Let also $I \subsetneq A$ be an ideal.*

(a) Then there are elements $y_1,\dots,y_d \in A$, which are algebraically independent such that A is a finitely generated $K[y_1,\dots,y_d]$-module. Furthermore, for some $\delta \le d$, $I \cap K[y_1,\dots,y_d] = \langle y_{\delta+1},\dots,y_d\rangle$.

(b) In addition if $\#K = \infty$, we have that $y_i = \sum_{j=i}^n a_{ij}x_j$ for suitable $a_{ij} \in K$.

Proof. 1. **case 1:** We first consider the case where $A = K[\mathbf{x}]$ and $I = \langle f\rangle$ for some $f \in A$ with $\deg f > 0$. We put $y_n = f$ and apply Lemma 2.10.67 1. to obtain $y_i = x_i - x_n^{r_i} \in A$ for suitable $r_i > 0$ for $i \in \{1,\dots,n-1\}$, such that

$$
y_n = f = ax_n^m + \sum_1^{m-1} G_i(y_1,\dots,y_{n-1})x_n^i \iff y_n - ax_n^m + \sum_1^{m-1} G_i(y_1,\dots,y_{n-1})x_n^i = 0,
$$

for some $a \in K\setminus\{0\}$, $m > 0$, $G_i \in K[y_1,\dots,y_{n-1}]$. Then x_n is integral over $K[y_1,\dots,y_n]$. Since $x_i = y_i + x_n^{r_i} \in A$ we get that $A = K[y_1,\dots,y_n][x_n]$, hence A is a finitely generated $K[y_1,\dots,y_n]$-module.

We now claim that $y_1,\dots,y_n$ are algebraically independent over K. Suppose for contradiction that this is not the case. Then $Y := \{y_1,\dots,y_n\}$ is not a transcendence basis of $K(y_1,\dots,y_n)$. However, by Lemma 2.10.45 (a) there is a subset of Y, say $y_{l_1},\dots y_{l_k}$ for $k < n$, which constitutes a transcendence basis for $K(y_1,\dots,y_n)$. Then by Corollary 2.10.50

$$
k = \mathbf{trdeg}K(y_1,\dots,y_n) = \mathbf{trdeg}K(x_1,\dots,x_n) = n > k,
$$

leading to a contradiction.

Let $\lambda \in I \cap K[y_1,\dots,y_n]$. Then $\lambda = gf = gy_n$ for some $g \in A$. g is integral over

$K[y_1,\dots,y_n]$, hence for suitable $h_1,\dots,h_{k-1} \in K[y_1,\dots,y_n]$,

$$g^k + \sum_{i=1}^{k-1} h_i g^i = 0 \Rightarrow \lambda^k = f^k g^k = -\sum_{i=1}^{k-1} h_i f^k g^i = -\sum_{i=1}^{k-1} h_i \lambda^i y_n^{k-i}.$$

This means $y_n \mid \lambda^k$, implying $y_n \mid \lambda$. From this we conclude $I \cap K[y_1,\dots,y_n] = \langle y_n \rangle$.

Case 2: We now prove the statement for $A = K[x_1,\dots,x_n]$ and an arbitrary ideal $I \subsetneq A$. For $I = 0$, we are done after choosing $y_i = x_i$ and $\delta = n$. We prove the statement for $I \neq 0$ by induction in $n \geq 1$. For $n = 1$, A is a PID, so I is generated by some non-zero polynomial. Then the statement follows from case 1.

Suppose now that $n > 1$ and let $f \in I \setminus \{0\}$. Again using Lemma 2.10.67 we find $y_1,\dots,y_n \in A$ that are algebraically independent over K with $y_n = f$. Then $y_1,\dots,y_{n-1}$ are also algebraically independent over K. By the induction hypothesis, we can find elements $t_1,\dots,t_{d-1} \in K[y_1,\dots,y_{n-1}]$ algebraically independent over K such that $K[y_1,\dots,y_{n-1}]$ is a finitely generated $K[t_1,\dots,t_{d-1}]$-module and $I \cap K[t_1,\dots,t_{d-1}] = \langle t_{\delta+1},\dots,t_{d-1} \rangle$ for some $\delta < d$. We then get that $K[y_1,\dots,y_n]$ is a finitely generated $K[t_1,\dots,t_{d-1},y_n]$-module, and hence A is a finitely generated $K[t_1,\dots,t_{d-1},y_n]$-module. Thus by a similar contradiction argument to that of case 1 I feel there is an argument that captures the fact better, we conclude that $d = n$ and $t_1,\dots,t_{n-1},y_n$ are algebraically independent over K.

Let $\lambda \in I \cap K[t_1,\dots,t_{n-1},y_n]$. Then $\lambda = g + hy_n$ for some $g \in I \cap K[t_1,\dots,t_{n-1}] = \langle t_{\delta+1},\dots,t_{n-1} \rangle$ and $h \in K[t_1,\dots,t_{n-1},y_n]$, then $I \cap K[t_1,\dots,t_{n-1},y_n] = \langle t_{\delta+1},\dots,t_{n-1},y_n \rangle$.

case 3: We now generalize to the case where $A = K[x_1,\dots,x_n]/J$ and $I \subsetneq A$ for an ideal $J \subsetneq K[x_1,\dots,x_n]$. We apply case 2 to J and find $y_1,\dots,y_n \in A$ algebraically independent over K such that $K[x_1,\dots,x_n]$ is a finitely generated $K[y_1,\dots,y_n]$-module and $J \cap K[x_1,\dots,x_n] = \langle y_{d+1},\dots,y_n \rangle$ for some $d \leq n$. Consider the embedding $\iota : K[y_1,\dots,y_n] \hookrightarrow A$. By construction we have that A is a finitely generated $\iota(K[y_1,\dots,y_n])$-module. It is easy to check that

$$\iota(K[y_1,\dots,y_n]) \simeq \frac{K[y_1,\dots,y_n]}{(J \cap K[y_1,\dots,y_n])} = \frac{K[y_1,\dots,y_n]}{\langle y_{d+1},\dots,y_n \rangle} \simeq K[y_1,\dots,y_d].$$

From which it follows that A is a finitely generated $K[y_1,\dots,y_d]$-module.

Let $I' = I \cap K[y_1,\dots,y_d]$. Then using case 2 we find $t_1,\dots,t_d \in K[y_1,\dots,y_d]$ algebraically independent over K such that $K[y_1,\dots,y_d]$ is a finitely generated $K[t_1,\dots,t_d]$-module and $I' \cap K[t_1,\dots,t_d] = \langle t_{\delta+1},\dots,t_d \rangle$ for some $\delta \leq d$. It then also follows that A is a finitely generated $K[t_1,\dots,t_d]$-module.

2.Suppose now that $\#K = \infty$. In case 1 the construction is also valid with $y_i = x_i - a_i x_n$ for suitable $a_i \in K$ by Lemma 2.10.67 2.
In case 2 we can choose $t_1,\ldots,t_{n-1}$ and $y_1,\ldots,y_{n-1}$ in the same way. In case 3 we can again choose $y_i = x_i - a_i x_n$ for suitable $a_i \in K$. It follows from case 2 that we can choose

$$t_j = y_i - b_j y_d = x_i - x_d - (a_i - b_j a_d)x_n + J,$$

which is of the desired form. $\square$

Definition 2.10.69. Let A be a finitely generated K-algebra. A sequence of elements $y_1,\ldots,y_d \in A$ with the properties specified in the above theorem is called a *Noether normalization* of A.

Corollary 2.10.70. *Consider* $A = K[x_1,\ldots,x_n]/I$*, with* $I \subset K[\mathbf{x}]$ *a prime ideal.*

1. *If* $y_1,\ldots,y_d$ *is a Noether normalization of* A*, then* $X = \{y_1,\ldots,y_d\}$ *defines a transcendence basis of* $L := Q(A) \supset K$*.*

2. *If* $\mathbf{trdeg}_K\ Q(A) = d$*, then* A *has a Noether normalization* $y_1,\ldots,y_d$

Proof. 1. By assumption $y_1,\ldots,y_d$ are algebraically independent over K. Secondly A is a finitely generated $K[\mathbf{y}]$-module, hence A is integral over $K[\mathbf{y}]$. Then L is integral over $K[\mathbf{y}]$. It follows that since $L \supset K(\mathbf{y}) \supset K[\mathbf{y}]$ $K(\mathbf{y})$ must be algebraic Lemma not yet written.
2. NNT there is a Noether normalization $y_1,\ldots,y_\delta \in A$. By 1. $\delta = \mathbf{trdeg}_K Q(A) = d$. $\square$

Corollary 2.10.71. *Let* A *be finitely generated* K*-algebra and* $y_1,\ldots,y_d \in A$ *be its Noether normalization. Then* $\iota : K[y_1,\ldots,y_d] \hookrightarrow A$ *is finite* K*-algebra homomorphism.*

3 The Real and Complex Numbers

3.1 A Topological Aside: Completion - a construction of $\mathbb{R}$

Definition 3.1.1. Let X be a non-empty set, K an ordered field. A map $d : X \times X \to K_{\geq 0} := \{a \in K : a \geq 0\}$ is called a *metric over* K if for every $x,y,z \in X$

1. $d(x,y) = 0 \iff x = y$.

2. $d(x,y) = d(y,x)$.

3. $d(x,y)+d(y,z) \geq d(x,z)$.

A set X with such a map d is called a *metric space over* K

Remark 3.1.2. Consider a normed vector space V over an ordered field K. Then

$$V \times V \to K_{\geq 0}, (v,w) \mapsto \|v-w\|$$

defines a metric on V.
1. $\|v-w\| = 0 \iff v-w = 0 \iff v = w$.
2. $\|v-w\| = |-1|\|v-w\| = \|-(v-w)\| = \|w-v\|$.
3. $\|v-w\| = \|(v-u)+(u-w)\| \leq \|v-u\| + \|w-u\|$.

Definition 3.1.3. Let (X,K,d) be a metric space and $\xi \in K_{>0} := \{a \in K : a > 0\}$, $x \in X$. We define *the ball of radius* ξ *with center* x to be

$$B_\xi(x) := \{y \in X : d(x,y) < \xi\}.$$

Lemma 3.1.4. *Let* (X,K,d) *be a metric space. Then*

$$\tau := \left\{U \subset X : \forall x(x \in U \Rightarrow \exists \xi > 0, B_\xi\right\},$$

defines a topology on X

Proof. Trivially $\emptyset \in \tau$, since $x \in \emptyset$ is false. It is also obvious that $X \in \tau$, since a ball of any radius is a subset of X. Consider a family of sets in τ, $\{U_\alpha\}_{\alpha \in A}$. Then any point x in the union of these subsets is in at least one of these subsets, U_β say. Then there is an $\xi > 0$ such that $B_\xi(x) \subset U_\beta \subset \bigcup_{\alpha \in A} U_\alpha$. Consider $U_1, \ldots, U_n \in \tau$. Then for each point x in the intersection of these sets there are $\xi_1, \ldots, \xi_n > 0$, such $B_{\xi_i}(x) \subset U_i$. picking $\xi = \min_{i \in \{1,\ldots,n\}} \xi_i$, it follows that $B_\xi(x) \subset \bigcap_1^n U_i$. □

Definition 3.1.5. Let (X,K,d) be a metric space. A sequence $(q_n) \in \prod_{\mathbb{N}} X$ is said *to converge to* q if for all $\xi > 0$ there exists a $N \geq 1$ such that for every $n \geq N$, $q_n \in B_\xi(q)$. We call q *the limit of* q_n *as* n *approaches infinity* and denote it $\lim_{n\to\infty} q_n$. We shall see that such a limit is unique.
A sequence is called *Cauchy* if for every $\xi > 0$ there exists an $N \geq 0$ such that for every $n,m \geq N$, $q_n - q_m \in B_\xi(0)$. We say that two Cauchy sequences $(a_n),(b_n) \in \prod_{\mathbb{N}} X$ are equivalent if $\lim_{n\to\infty} d(x_n,y_n) = 0$. We write $(a_n) \equiv_d (b_n)$. We define *the completion of* X *with respect to* d to be the set

$$\overline{X} := \left\{(q_n) \in \prod_{\mathbb{N}} X : (q_n) \text{ is Cauchy}\right\} / \equiv_d.$$

Remark 3.1.6. Let (x_n) be a convergent sequence with limit x. Let $y \in X \setminus \{x\}$. Set $\epsilon := d(x,y)/2$. Then for every $n \geq N$ for some $N \geq 1$, $x_n \notin B_\epsilon(y)$, hence the limit of a convergent sequence is unique. If V is a normed vector space, $\lim_{n\to\infty}$ defines a linear operator from the space of convergent sequences to K which readily follows from the homogeneity of the norm and from the triangle inequality. One also easily checks that $\lim_{n\to\infty} \circ \|\cdot\| = \|\lim_{n\to\infty}\|$. On K consider convergent sequences $(a_n),(b_n) \in \prod_{\mathbb{N}} K$. Then for any $\xi > 0$ for sufficiently large n

$$|a_n b_n - ab| \leq |b_n||a_n - a| + |a||b_n - b| \overset{n\to\infty}{\to} 0.$$

On K consider a sequence $(a_n) \in \prod_{\mathbb{N}} K \setminus 0$ that is convergent with limit $a \in K \setminus 0$. Let $\xi > 0$. For large enough N for $n \geq N$, $|a_n||a_n - a| < \xi$. Pick $0 < \delta < |a|$. For some M, for $n \geq M$, $|a_n| \in B_{|a|-\delta}(|a|)$, hence $|a| - |a_n| < |a| - \delta$. Then setting $\epsilon := \min\{|a_1|,\dots,|a_{\max(N,M)}|,\delta\}$. Then for $n \geq \max(N,M)$,

$$\left|\frac{1}{a_n} - \frac{1}{a}\right| = \left|\frac{a_n - a}{a_n a}\right| < \frac{\xi}{|a|\epsilon}.$$

Lemma 3.1.7. *Let (X,K,d) be a metric space. Let (x_n) be a Cauchy sequence in X. Then (x_n) is bounded.*

Proof. Let $\xi > 0$ be given, then for a sufficiently large N, $d(x_n,x_N) < \xi$ for every $n \geq N$. Pick $\delta = \max\{d(x_1,x_N),\dots,d(x_{N-1},x_N),\xi\}$. Then $x_n \in B_\delta(x_N)$ for every $n \geq 1$. □

Definition 3.1.8. A function between metric spaces over a fixed field $f : X \to Y$ is said to be *sequentially continuous* if for every convergent sequence $(x_n) \in \prod_{\mathbb{N}} X$, $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$.

Remark 3.1.9. Being sequentially continuous is equivalent to being continuous with respect to the topology induced by the metric on X, respectively on Y.

In general, it may be difficult to endow the completion of a metric space with a topology. If we are working over complete (i.e. every Cauchy sequence is convergent) ordered field, $([a_n],[b_n]) \mapsto \lim_{n\to\infty} d(a_n,b_n)$ defines a metric on the completion of X. Any two complete ordered fields are isomorphic as topological fields. Moreover, the completion of an ordered field can be endowed with structure of a complete ordered field.

Lemma 3.1.10. *Let K be an ordered field. We define addition on $\overline{K}$ to be*

$$[(a_n)] + [(b_n)] := [(a_n + b_n)] \quad ([(a_n)],[(b_n)] \in \overline{K}).$$

We also define multiplication by

$$[(a_n)][(b_n)] := [(a_n b_n)].$$

For $[(a_n)],[(b_n)] \in \overline{K}$ *we say that* $[(a_n)] \geq [(b_n)]$ *if there is an* $N \geq 0$ *such that for* $n \geq N$, $a_n \geq b_n$. *With these definitions* $\overline{K}$ *becomes an ordered field. The subfield*

$$K' := \{[(a)] \in \overline{K} : a \in K\}$$

is isomorphic to K *as topological fields.*

Proof. Suppose $([(a_n)],[(b_n)]) = ([(a'_n)],[(b'_n)])$. Then addition is well-defined by a single application of the triangle inequality. Multiplication is also well-defined since

$$\begin{aligned}|a_n b_n - a'_n b'_n| &= |a_n b_n - a_n b'_n + a_n b'_n - a'_n b'_n| = |a_n(b_n - b'_n) + b'_n(a_n - a'_n)| \\ &\leq |a_n||b_n - b'_n| + |b'_n||a_n - a'_n| \leq \epsilon|b_n - b'_n| + \delta|a_n - a'_n| \overset{n\to\infty}{\to} 0\end{aligned}$$

where $\epsilon,\delta > 0$ are obtained from Lemma 3.1.7. It is fairly easy to check that upon defining $0_{\overline{K}} := 0 := [(0)]$, $1_{\overline{K}} := 1 := [(1)]$, $-[(a_n)] := [(-a_n)]$, for $[(a_n)] \in \overline{K}$, $\overline{K}$ becomes a commutative ring. Let $[(a_n)] \neq 0$. Then for n greater than some N, $a_n \neq 0$. Indeed, suppose (c_n) is Cauchy such that $c_n = 0$ for infinitely many n. Then for any $\xi > 0$ there is an $N \geq 0$ such that $|c_n - c_m| < \xi$ for $n,m \geq N$. In particular we may choose m such that $c_m = 0$, hence $|c_n - 0| < \xi$, implying $\lim_{n\to\infty} c_n = 0$, hence $[(c_n)] = 0$. We then define $b_n = 0$ for $n < N$ and $b_n = a_n^{-1}$ for $n \geq N$ and see that $[(a_n)][(b_n)] = [(1)]$, implying $\overline{K}$ is a field. It is fairly easy to check that $\leq$ defines a partial order on $\overline{K}$. We now check that $\leq$, defines a total order. Let $[(a_n)],[(b_n)] \in \overline{K}$. We prove first that the statement is true for $[(b_n)] = 0 = [(0)]$. The statement is obvious when $[(a_n)] = 0$. Note that $a_n > 0$ or $a_n < 0$ for every sufficiently large n. Suppose now (c_n) is Cauchy and $c_n > 0$ for infinitely many n and $c_n < 0$ for infinitely many n. Let (c'_n) and (c''_n) the sequences entries of (c_n) being > 0 resp. < 0. Then for sufficiently large n,

$$|c'_n| + |c''_n| = c'_n - c''_n = |c'_n - c''_n| < \xi$$

for every $\xi > 0$, using the fact that (c_n) is Cauchy, meaning $\lim_{n\to\infty} c_n = 0$. It thus follows that $a_n > 0$ for every sufficiently large n or $a_n < 0$ for every sufficiently large n, hence $[(a_n)] > 0$ or $[(a_n)] < 0$. In the general setting, we thus have that $[(a_n - b_n)] \geq 0$ or $[(a_n - b_n)] \leq 0$. This is equivalent to $a_n - b_n \geq 0$ for every large n or $a_n - b_n \leq 0$ for every large n. It follows that $[(a_n)] \geq [(b_n)]$ or $[(a_n)] \leq [(b_n)]$. It is easy to check that $\overline{K}$ becomes an ordered field with this total ordering. For instance if $[(a_n)] \leq [(b_n)]$

then $a_n + c_n \leq b_n + c_n$ for sufficiently large n, hence $[(a_n)] + [(c_n)] \leq [(b_n)] + [(c_n)]$. The map

$$K \to K'$$
$$a \mapsto [(a)]$$

is readily seen to be a ring isomorphism with mutual inverse $[(a)] \mapsto a$. Both of these maps are sequentially continuous, hence K and K' are isomorphic as topological fields. □

Remark 3.1.11. From this point on we identify K' with K.

Lemma 3.1.12. *Let $(K, |\cdot|)$ be an ordered field considered with structure of metric space in the natural way. Then K is a topological field.*

Proof. Using sequential continuity, the result follows from Remark 3.1.6. □

Lemma 3.1.13. *An ordered field is characteristic 0.*

Proof. Indeed, $0 < 1$. Suppose $0 < n$ for some $n \geq 1$. Then $1 < n+1$, hence $0 < 1 < n+1$. □

Remark 3.1.14. The utility of this lemma is that we may regard $\mathbb{Z}$ and $\mathbb{Q}$ as subrings in a canonical way.

Proposition 3.1.15. *For an ordered field K, we have that $\overline{K}$ is Dedekind complete, i.e. every non-empty bounded set has a least upper bound.*

Proof. Let $\emptyset \neq B \subset \overline{K}$ be bounded from above. Let $[(b_n)]$ be an upper bound. (b_n) is bounded, hence $[(b_n)] < [(u)]$ for some $u \in K$. Let $[(c_n)] \in B$ be arbitrary. Again, since $([c_n])$ is bounded, $([l]) < ([c_n])$ for some $l \in K$. Set $u_1 := u$ and $l_1 := l$. For each $n \geq 1$, if $(u_n + l_n)/2$ is an upper bound, set $u_{n+1} := (u_n + l_n)/2$ and $l_{n+1} := l_n$. Note that $|u_n - l_n| = \frac{1}{2^n}(u - l)$, which is easily verified by induction in n, hence $\lim_{n\to\infty} u_n - l_n = 0$. Then

$$|u_n - u_m| = |u_n - l_n| + |l_n - u_m| < \epsilon$$

for every $\epsilon > 0$ and sufficiently large n, m, hence (u_n) is Cauchy. A similar argument shows that (l_n) is Cauchy. We then have that $[(l_n)] = [(u_n)]$. By construction each u_n is an upper bound and each l_n is not an upper bound. We thus have that $[(u_n)]$

is an upper bound. Let $u' := [(a'_n)] \leq [(u_n)]$ be another upper bound of B. Then for every $\epsilon > 0$, for sufficiently large n

$$|a'_n - l_n| = a'_n - l_n \leq u_n - l_n = |u_n - l_n| < \epsilon,$$

hence $[(a'_n)] = [(l_n)] = [(u_n)]$, proving that $[(u_n)]$ is the least upper bound. $\square$

Lemma 3.1.16. *Let K be a Dedekind complete ordered field. Then a monotone sequence bounded sequence is convergent.*

Proof. Let (a_n) be an increasing sequence in K. Since the image of (a_n) is bounded, it has a least upper bound a. We claim that (a_n) converges to a. Let $\epsilon > 0$. For some $N \geq 0$, $a - \epsilon < a_N \leq a < a + \epsilon$, hence for every $n \geq N$, $-\epsilon < a_n - a < \epsilon$, hence $a_n \in B_\epsilon(a)$. $\square$

Definition 3.1.17. We define *the field of real numbers* as $\mathbb{R} := \overline{Q}$.

Definition 3.1.18. An ordered field K is *Archimedean* if $\mathbb{Q} \subset K$ is not bounded from above

Lemma 3.1.19. *Let K be an Archimedean ordered field. Then $\mathbb{Q}$ is dense in K.*

Proof. Let $0 < c \leq d$. There is a least upper bound u to the set $\{n \in \mathbb{N} : nc \leq d\}$. Since K is Archimedean we may pick a natural number $N > u$, hence $Nc > d$. The cases $a \leq 0 < b$ and $a < 0 \leq b$ are obvious. Suppose $0 < a < b$. Claim: if $c < d$ are such that $d - c > 1$, then $\{m \in \mathbb{Z} : c < m < d\} \neq \emptyset$. Let m_0 be the greatest integer smaller than c. Then $c < m_0 + 1$ and $d - m_0 > d - c > 1$ hence $c < m_0 + 1 < d$. This proves that if $0 < a < b$ such that $a - b > 1$ then there is a rational number in between a and b. Suppose $a - b < 1$. Then there is a positive integer n such that $n(a - b) > 1$ and by the claim there is then an integer m such that $na < m < nb$ hence $a < m/n < b$. $\square$

Definition 3.1.20. A metric space (X,K,d) is called *Cauchy complete* if every Cauchy sequence is convergent

Lemma 3.1.21. *An ordered field K is Dedekind complete if and only if it is Archimedean and Cauchy complete.*

Proof. "$\Rightarrow$": Suppose for a contradiction that $\mathbb{Q}$ is bounded from above. Set $s := \sup\ \mathbb{Q}$. We must have that s is not rational, since if it were $s+1$ would be a rational greater than s. Then $s - 1 \leq q < s$ for some $q \in \mathbb{Q}$, but then $q + 1$ is a rational number greater than s, leading to a contradiction. Let (a_n) be a Cauchy sequence in K. Set

$(u_n) := (\sup_{k\geq n}\ a_n)$. This is a decreasing bounded sequence, hence u_n converges to some a. Let $\epsilon > 0$. Then there is some $N \geq 1$ such that $|a_n - a_m| < \epsilon/3$ and $|u_n - l| < \epsilon/3$ for every $n, m \geq N$. We may also find an $M \geq N$ such $u_N - \epsilon/3 < a_M$. It thus follows that

$$|a_n - a| = |a_n - a_M + a_M - u_N + u_N - a| \leq |a_n - a_M| + |a_M - u_N| + |u_N - a| < \epsilon,$$

meaning (a_n) converges to a.
"$\Leftarrow$": Let S be a non-empty, bounded subset of K. Let u be a rational upper bound of S (we may pick this as K is Archimedean) and l be a rational number smaller than an $s \in S$ (we may pick such a rational number since $\mathbb{Q}$ is dense in K which follows from K being Archimedean). Define (u_n) and (l_n) as in the prior proposition. Then (u_n) and (l_n) are Cauchy sequences, which are convergent by the assumption that K is Cauchy complete. It follows that since $|l_n - u_n| \to 0$, $\lim_{n\to\infty} u_n = \lim_{n\to\infty} l_n = u$. Since every u_n is an upper bound of S so is u. Let $v < u$. Then for large n, $u - l_n < u - v$, hence $l_n > v$. Since l_n is never an upper bound of S, we may pick a $t \in S$ such that $t \geq l_n > v$, hence u is the smallest upper bound of S. This shows that K is Dedekind complete. $\square$

Definition 3.1.22. We say that a function $f : X \to Y$ of metric space is *uniformly continuous* if for every $\epsilon > 0$ there is a $\delta > 0$ such that for every $x, y \in X$, $d(x, y) < \delta \Rightarrow d(f(x), f(y)) < \epsilon$

Lemma 3.1.23. *If $f : X \to Y$ is uniformly continuous and (x_n) is Cauchy, then $(f(x_n))$ is Cauchy.*

Proof. Let $\epsilon > 0$. We can pick $\delta > 0$ such that for $x, y \in X$, if $d(x, y) < \delta$, $d(f(x), f(y)) < \epsilon$. Pick $N \geq 1$ such that $d(x_n, x_m) < \delta$ for every $n, m \geq N$. Then $d(f(x_n), f(x_m)) < \epsilon$ for every $n, m \geq N$. $\square$

Remark 3.1.24. Every uniformly continuous function is clearly continuous.

Lemma 3.1.25. *Let X, Y be metric spaces, where Y is complete and $f : A \to B$ be uniformly continuous where $A \subset X$ and $B \subset Y$. Then f can be uniquely extended to a uniformly continuous function $f : \mathrm{cl}(A) \to \mathrm{cl}(B)$.*

Proof. Let (x_n) be a convergent sequence in A. Then $(f(x_n))$ in Y is Cauchy and hence also convergent by completeness. We then define

$$f : \mathrm{cl}(A) \to \mathrm{cl}(B)$$
$$x \to \lim_{n\to\infty} f(x_n),$$

where (x_n) is a sequence in A converging to x. We need to check that this is well-defined. Note first that $f(A) \subset B \Rightarrow \mathrm{cl}(f(A)) \subset \mathrm{cl}(B)$, hence $\lim_{n\to\infty} f(x_n) \in \mathrm{cl}(B)$. Let $(x_n),(y_n)$ be two sequences in A converging to $x \in \mathrm{cl}(A)$. By uniform continuity (skipping a step) for large n, we get that $d(f(x_n), f(y_n)) < \epsilon$ for every $\epsilon > 0$, hence $\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} f(y_n)$. Let $\epsilon > 0$ be given. Pick $\delta > 0$ small enough such that for every $p,q \in A$ with $d(p,q) < 2\delta \Rightarrow d(f(p), f(q)) < \epsilon/3$. Pick $x,y \in \mathrm{cl}(A)$ such that $d(x,y) < \delta$. Pick sequences $(x_n),(y_n)$ in A converging to x respectively y. Then for large enough n, $d(x_n, y_n) \le d(x_n, x) + d(y_n, y) < 2\delta$, hence, again for large n,

$$d(f(x), f(y)) \le d(f(x_n), f(x)) + d(f(x_n), f(y_n)) + d(f(y_n), f(y)) < \epsilon.$$

It follows that the constructed extension of f is uniformly continuous. Let g be another such extension. Let $x \in \mathrm{cl}(A)$ and (x_n) a sequence in A converging to x. Then since g is continuous,

$$g(x) = g(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} f(x_n) = f(x).$$

$\square$

Proposition 3.1.26. *Any two Dedekind complete ordered fields are isomorphic as topological fields.*

Proof. Let K,L be two such fields. There is a canonical isomorphism between the rational numbers in K and the rational numbers in L. each copy of $\mathbb{Q}$ is a dense subset of the respective fields. It is easy to show that any point in a metric space over a complete metric space can be approximated by a Cauchy sequence in a dense subset. The isomorphism $\sigma : \mathbb{Q} \subset K \to \mathbb{Q} \subset L, q \mapsto q$ is clearly uniformly continuous and can therefor by the prior lemma be extended to an isomorphism $\sigma : \mathrm{cl}(\mathbb{Q}) = K \to \mathrm{cl}(\mathbb{Q}) = L$. $\square$

3.2 From $\mathbb{R}^2$ to $\mathbb{C}$

Note that $(\mathbb{R}^2, +, |\cdot|)$ is a normed topological vector space. We seek to endow this vector space with a multiplication such that it becomes a topological field, containing $\mathbb{R}$ as a topological subfield. We will call this field *the complex numbers* and denote it $\mathbb{C}$.

Definition 3.2.1. We define

$$\cdot : \mathbb{R}^2 \to \mathbb{R}^2$$

$$((a,b),(c,d)) \mapsto (ac - bd, ad + cb)$$

Proposition 3.2.2. *With this operation $\mathbb{C} := (\mathbb{R}^2, +, \cdot)$ is a field. The multiplication is obviously continuous w.r.t. $|\cdot|$ making $\mathbb{C}$ a topological field. The set*

$$R := \{(a,0) \in \mathbb{C} : a \in \mathbb{R}\}$$

is a subfield of $\mathbb{C}$ isomorphic to $\mathbb{R}$.

Proof. Let $(a,b),(c,d),(e,f) \in \mathbb{C}$. We then have that

$$\begin{aligned}((a,b)(c,d))(e,f) &= (ac-bd, ad+cb)(e,f) = (ace-bde-adf-cbf, acf-bdf+ade+cbe)\\ &= (a(ce-df)-b(de+cf), a(cf+de)+b(ce-df)) = (a,b)(ce-df, cf+de)\\ &= (a,b)((c,d)(e,f)).\end{aligned}$$

We define $1_{\mathbb{C}} = (1,0)$. One easily checks that $(a,b)(c,d) = (c,d)(a,b)$. Then

$$1_{\mathbb{C}}(a,b) = (1,0)(a,b) = (1a-0b, 1b+0a) = (a,b),$$

shows that $1_{\mathbb{C}}$ is the neutral element with respect to the multiplication. We moreover have that

$$\begin{aligned}(a,b)(c+e,d+f) &= (ac+ae-bd-bf, ad+bf+bc+be)\\ &= (ac-bd, ad+bc)+(ae-bf, bf+be) = (a,b)(c,d)+(a,b)(e,f).\end{aligned}$$

Suppose $(a,b) \neq 0$. Then

$$(a,b)\left(\frac{1}{a^2+b^2}(a,-b)\right) = \left(\frac{a^2+b^2}{a^2+b^2}, \frac{-ab+ab}{a^2+b^2}\right) = (1,0).$$

We have thus shown that $\mathbb{C}$ is a field. Note that R is subspace of $\mathbb{C}$ and that for $(a,0),(b,0) \in R$,

$$(a,0)(b,0) = (ab-0, a0+0b) = (ab,0) \in R.$$

Moreover, we clearly have that $(1,0) \in R$ and that for $a \neq 0$ $(a,0)(1/a,0) = (1,0)$. We thus see that R is a subfield of $\mathbb{C}$. The map

$$\mathbb{R} \to R, a \mapsto (a,0),$$

clearly defines a bijective $\mathbb{R}$-algebra homomorphism, hence $R \simeq \mathbb{R}$. □

Remark 3.2.3. One further notices that $\mathbb{C}$ is an $\mathbb{R}$ vector space with basis $\{1, i\}$, where we define $i = (0,1)$. Note that $\pm i$ are the roots of the polynomial $x^2+1 \in \mathbb{C}[x]$. We therefor also write $i = \sqrt{-1}$.

4 Classical Affine Algebraic Geometry

4.1 Introducing Algebraic Sets

4.1.1 Introducing Affine Algebraic Sets & the Affine Zariski Topology

Definition 4.1.1. Let K be any field. By *the affine n-space over K*, for some positive integer n, we mean the n-fold cartesian product of K with itself. We denote affine n-space over K by $\mathbb{A}^n(K)$ or just $\mathbb{A}^n$ when this will not lead to ambiguity. Thus we have

$$\mathbb{A}^n(K) := \mathbb{A}^n := K^n.$$

We also refer to Affine 1-space as the affine line, and Affine 2-space as the affine plane.

Remark 4.1.2. Whenever we write $\mathbb{A}^n$ and no field is given prior, implicitly a field K is given and we mean $\mathbb{A}^n = \mathbb{A}^n(K)$.

Definition 4.1.3. Consider a field K. we define *the vanishing set of S over K* denoted $V_K^{\mathbb{A}}(S) := V_K(S) := V(S)$ to be the set

$$\left\{v \in \mathbb{A}^n(K) : f(v) = 0 \text{ for every } f \in S\right\}.$$

similarly we define *the vanishing set of S over L* denoted $V_L(S)$ to be the set

$$\left\{v \in \mathbb{A}^n(L) : f(v) = 0 \text{ for every } f \in S\right\}.$$

When M is finite, say $M = \{f_1, \dots, f_m\}$, we define

$$V_F(f_1, \dots, f_m) := V_F(\{f_1, \dots, f_m\}) \quad V_L(f_1, \dots, f_m) := V_L(\{f_1, \dots, f_m\}).$$

Our convention will be that $V(M) := V_K(M)$

Definition 4.1.4. Let K be any field. A subset $X \subset \mathbb{A}^n$ is called an *affine algebraic subset*, or an *algebraic subset* when no ambiguity arises, if there is a subset $M \subset K[x_1, \dots, x_n]$ such that $X = V(M)$

Lemma 4.1.5. *Let $M, M' \subset K[x_1, \dots, x_n]$ with $M \supset M'$. Then $V(M) \subset V(M')$.*

Proof. Let $v \in V(M)$ and $f \in M'$. Then by assumption $f \in M$ and hence $f(v) = 0$, implying $v \in V(M')$. □

Proposition 4.1.6. *For any subset $M \subset K[x_1, \dots, x_n]$. Set $I := \langle M \rangle$. Then $V(M) = V(I)$*

Proof. Let $v \in V(M)$. Let $f \in I$. Then for some $f_1, \dots, f_m \in S$

$$f = \sum_1^m c_i f_i,$$

for suitable $c_1, \dots, c_m \in K[x_1, \dots, x_n]$. Then $f_i(v) = 0$ for every $i \in \{1, \dots, m\}$ and hence

$$f(v) = \sum_1^m c_i(v) f_i(v) = \sum_1^m c_i(v) \cdot 0 = 0,$$

meaning $v \in V(I)$ and hence $V(M) \subset V(I)$.
The converse inclusion follows from Lemma 4.1.5 since $I \supset M$. □

It follows as a corollary that every affine algebraic set arises as a vanishing of some polynomial ideal.

Corollary 4.1.7. *Let $X \subset \mathbb{A}^n$ be algebraic. Then $X = V(I)$ for some ideal $I \subset K[x_1, \dots, x_n]$.*

As another corollary to the above proposition we have that every algebraic set arises as .

Corollary 4.1.8. *Let $X \subset \mathbb{A}^n$ be an algebraic subset. Then $X = V(f_1, \dots, f_m)$ for suitable $f_1, \dots, f_m \in K[x_1, \dots, x_n]$.*

Proof. By Corollary 4.1.7 $X = V(I)$ for some ideal $I \subset K[x_1, \dots, x_n]$. By Corollary 2.9.49 $I = \langle f_1, \dots, f_m \rangle$ for suitable $f_1, \dots, f_m \in K[\mathbf{x}]$. Thus it follows from the above proposition that

$$X = V(I) = V(\langle f_1, \dots, f_m \rangle) = V(f_1, \dots, f_m).$$

□

Remark 4.1.9. Let an algebraic subset $X \subset \mathbb{A}^n$ be given. We can find $f_1, \dots, f_m \in K[x_1, \dots, x_n]$ such that $X = V(f_1, \dots, f_m)$. Consider the map

$$\varphi : \mathbb{A}^n \to \mathbb{A}^m$$
$$v \mapsto (\mathrm{ev}_v(f_1), \mathrm{ev}_v(f_2), \dots, \mathrm{ev}_v(f_m)) = (f_1(v), f_2(v), \dots, f_m(v)).$$

It is then clear that $V(f_1, \dots, f_m) = \varphi^{-1}(\mathbf{0})$. Thus one concludes that every algebraic subset of $\mathbb{A}^n$ arises the zero set of a map of the same form as φ. We shall later see that φ is an element of a special class of maps, called *polynomial maps*, which are central to study of affine algebraic geometry.

Lemma 4.1.10. *We collect the following results*

(i) Let A be some indexing set. Consider a family of algebraic sets $\{X_\alpha\}_{\alpha \in A}$ in $\mathbb{A}^n$. Then $\bigcap_\alpha X_\alpha$ is an algebraic set.

(ii) Consider algebraic sets $X, Y \subset \mathbb{A}^n$. Then $X \cup Y$ is algebraic. It follows by induction that $\bigcup_1^k X_i$ is algebraic for any finite sequence of algebraic sets $X_1, \ldots, X_k$ in $\mathbb{A}^n$.

Proof. (i) By Corollary 4.1.7 for every $\alpha \in A$, we can find an ideal $I_\alpha \subset K[x_1, \ldots, x_n]$ such that $X_\alpha = V(I_\alpha)$. We claim that $\bigcap_\alpha V(I_\alpha) = V(\bigcup_\alpha I_\alpha)$.
Let $v \in \bigcap_\alpha V(I_\alpha)$. Let $f \in \bigcup_\alpha I_\alpha$. Then $f \in I_\beta$ for some $\beta \in A$. Then since $v \in \bigcap_\alpha V(I_\alpha)$, in particular $v \in V(I_\beta)$, implying $f(v) = 0$, hence $v \in V(\bigcup_\alpha I_\alpha)$.
Let $\beta \in A$, then $I_\beta \subset \bigcup_\alpha I_\alpha$. By Lemma 4.1.5 $V(\bigcup_\alpha I_\alpha) \supset V(I_\beta)$, hence $V(\bigcup_\alpha I_\alpha) \supset \bigcap_\alpha V(I_\alpha)$.
It follows that

$$\bigcap_\alpha X_\alpha = \bigcap_\alpha V(I_\alpha) = V\left(\bigcup_\alpha I_\alpha\right).$$

(ii) We have that $X = V(I)$ and $Y = V(J)$ for some ideals $I, J \subset K[x_1, \ldots, x_n]$. We claim that $V(I) \cup V(J) = V(IJ)$.
$IJ \subset I$ and $IJ \subset J$, hence by Corollary 4.1.5 $V(I), V(J) \subset V(IJ)$, meaning $V(I) \cup V(J) \subset V(IJ)$.
Let $v \in V(IJ)$. If $v \in V(I)$ we find that $v \in V(I) \cup V(J)$. Suppose $v \notin V(I)$. Then for some $f \in I$ $f(v) \neq 0$. Let $g \in J$ be given. then $fg \in IJ$, hence $0 = (fg)(v) = f(v)g(v)$, since $f(v) \neq 0$ it follows that $g(v) = 0$, hence $v \in V(J) \subset V(I) \cup V(J)$.
It follows that

$$X \cup Y = V(I) \cup V(J) = V(IJ).$$

□

We collect the most fundamental examples of (affine) algebraic sets

Example 4.1.11. Fix a field K

1. Trivially the vanishing set of some subset $M \subset K[x_1, \ldots, x_n]$ is a an algebraic set.

2. Affine n-space is an algebraic set. Indeed, see that

$$\mathbb{A}^n = V(0)$$

3. The empty set $\emptyset \subset \mathbb{A}^n$ is an algebraic set. Indeed one readily verifies that

$$\emptyset = V(1).$$

4. A singleton or *point*, $\{(a_1,\ldots,a_n)\}$ for $a_1,\ldots,a_n \in K$, is an algebraic set. Indeed, one checks that

$$\{(a_1,\ldots,a_n)\} = V(x_1 - a_1,\ldots,x_n - a_n).$$

   It thus follows that from Lemma 4.1.10 (ii) that any finite subset of $\mathbb{A}^n$ is algebraic

We can represent some of the data collected thus far as a topology on affine n-space

Theorem 4.1.12. *Consider the family of sets*

$$\tau_{\mathcal{Z}} := \left\{\mathbb{A}^n \setminus X : X \subset \mathbb{A}^n \text{ is an algebraic set}\right\}.$$

This constitutes a topology on $\mathbb{A}^n$ called the Zariski topology.

Proof. One sees from example 4.1.11 2. & 3. that $\mathbb{A}^n, \emptyset \in \tau_{\mathcal{Z}}$.
Consider some indexing set A. Let $\{B_\alpha\}_{\alpha\in A} \subset \tau_{\mathcal{Z}}$ be given. For each $\alpha \in A$ we can find $X_\alpha \subset \mathbb{A}^n$ such that $B_\alpha = \mathbb{A}^n \setminus X_\alpha$. Then it follows from Proposition 4.1.10 (i) that

$$\bigcup_\alpha B_\alpha = \bigcup_\alpha \mathbb{A}^n \setminus X_\alpha = \mathbb{A}^n \setminus \left(\bigcap_\alpha X_\alpha\right) \in \tau_{\mathcal{Z}}$$

Let $B_1,\ldots,B_k \in \tau_{\mathcal{Z}}$. For suitable $X_1,\ldots,X_k \subset \mathbb{A}^n$, $B_i = \mathbb{A}^n \setminus X_i$. Then it follows from Proposition 4.1.10 (ii) that

$$\bigcap_1^k B_i = \bigcap_1^k \mathbb{A}^n \setminus X_i = \mathbb{A}^n \setminus \left(\bigcup_1^k X_i\right) \in \tau_{\mathcal{Z}}.$$

□

Remark 4.1.13. The closed subsets of $\mathbb{A}^n$ under the Zariski topology are exactly the algebraic sets in $\mathbb{A}^n$. Thus we may opt to say that an algebraic subset $X \subset \mathbb{A}^n$ is a *Zariski closed subset.*

4.1.2 Miscellaneous Result about Algebraic sets, Examples & Non-examples

Proposition 4.1.14. *The Zariski closed subsets of $\mathbb{A}^1 = K$ are the finite subsets of $\mathbb{A}^1$ and $\mathbb{A}^1$ itself.*

Proof. Let $X \subset \mathbb{A}^1$ be Zariski closed. Suppose $X \neq V(0) = \mathbb{A}^1$, i.e. that $X = V(f_1,\ldots,f_m)$ for $f_1,\ldots,f_m \in K[x_1,\ldots,x_n]$, where $f_i \neq 0$ for at least one $i \in \{1,\ldots,m\}$. Then $X \subset V(f_i)$ by Lemma 4.1.5. $V(f_i)$ is the set of roots of f_i. Since $f_i \neq 0$, $\#V(f_i) \leq \deg\ f_i < \infty$ GIVE REFERENCE, hence

$$\#X \leq \#V(f_i) \leq \deg\ f_i < \infty.$$

$\square$

Proposition 4.1.15. *Let K be a field with $\#K < \infty$. Let $X \subset \mathbb{A}^n$. Then X is Zariski closed.*

Proof. By assumption $\#X < \infty$, hence there are points $v_1,\ldots,v_k \in \mathbb{A}^n$ such that

$$X = \bigcup_1^k \{v_i\}.$$

By Example 4.1.11 X is Zariski closed. $\square$

We now give an example of a countable family of algebraic subsets whose union is not algebraic.

Example 4.1.16. Let $K = \mathbb{Q}$ and let and consider the family of algebraic sets $\{\{i\}\}_{i \in \mathbb{N}}$. The only algebraic subset of $\mathbb{A}^1$ are $\mathbb{A}^1$ it self and the finite subsets of $\mathbb{A}^1$ (cf. Proposition 4.1.14), but $\bigcup_{i \in \mathbb{Z}} \{i\} = \mathbb{Z}$, which is neither finite or equal to $\mathbb{A}^1 = \mathbb{Q}$.

Here are some examples of algebraic subsets

Example 4.1.17.

1. Let $X = \left\{(t,t^2,t^3) \in \mathbb{A}^3 : t \in K\right\}$. We prove that $X = V(f_1,f_2)$, where $f_1 = x^2 - y, f_2 = x^3 - z \in K[x,y,z]$. Let $v \in X$. Then $v = (t,t^2,t^3)$ for some $t \in K$. Note that

   $$f_1(v) = t^2 - t^2 = 0 \text{ and } f_2(v) = t^3 - t^3 = 0,$$

   implying that $v \in V(f_1,f_2)$.
   Let $v = (v_1,v_2,v_3) \in V(f_1,f_2)$. Put $t = v_1$. Then $t^2 = v_1^2 = v_2$ and $t^3 = v_1^3 = v_3$, hence $v = (v_1,v_2,v_3) = (t,t^2,t^3) \in X$.

2. Let $X = \left\{(\cos(t),\sin(t)) \in \mathbb{A}^2(\mathbb{R}) : t \in \mathbb{R}\right\}$. One sees that $X = V(x^2+y^2-1)$. Indeed, for $v = (\cos(t),\sin(t)) \in X$, $\sin^2(t) + \cos^2(t) = 1$, implying that $v \in V(x^2+y^2-1)$. For $v = (v_1,v_2) \in V(x^2+y^2-1)$, we have that $|v|^2 = v_1^2 + v_2^2 = 1$, implying that $v \in S^1 = \mathrm{im}(\mathbb{R} \ni t \mapsto (\cos(t),\sin(t) \in \mathbb{A}^2(\mathbb{R}))$.

3. Let $X = \{(\sin(t)\cos(t), \sin^2(t)) \in \mathbb{A}^2(\mathbb{R}) : t \in [0,2\pi)\}$, i.e. the points in $\mathbb{A}^2(\mathbb{R})$ with polar coordinates (r,t) such that $r = \sin(t)$. Recall the trigonometric identities

$$\sin(\alpha+\beta) = \sin(\alpha)\cos(\alpha) + \sin(\beta)\cos(\alpha),$$
$$\cos(\alpha+\beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$$

for $\alpha, \beta \in [0,2\pi)$. For $t \in [0,2\pi)$ this is equivalent to

$$\sin(2t) = 2\sin(t)\cos(t) \iff \sin(t)\cos(t) = \frac{1}{2}\sin(2t) \text{ and}$$
$$\cos(2t) = \cos^2(t) - \sin^2(t) = 1 - 2\sin^2(t) \iff \sin^2(t) = \frac{1}{2}(1-\cos(2t)),$$

hence $X = \{(\frac{1}{2}\sin(2t), \frac{1}{2}(1-\cos(2t))) : t \in [0,2\pi)\}$. Let $f = x^2 + (\frac{1}{2} - y)^2 - \frac{1}{4}$. One sees that $X = V(f)$. Indeed, if $v = ((\frac{1}{2}\sin(2t), \frac{1}{2}(1-\cos(2t))) \in X$ we have that

$$f(v) = \frac{1}{4}\cos^2(2t) + \left(\frac{1}{2} - \frac{1}{2} + \frac{1}{2}\sin(2t)\right)^2 - \frac{1}{4} = \frac{1}{4}\left(\cos^2(2t) + \sin^2(2t)\right) - \frac{1}{4} = \frac{1}{4} - \frac{1}{4} = 0,$$

thus we conclude $v \in V(f)$. Let $v = (v_1, v_2) \in V(f)$. We set $t = \frac{1}{2}\arcsin(2v_1)$. Then one easily checks that $\cos(2t) = v_1$. We also note that

$$v_1^2 = -\left(\frac{1}{2} - v_2\right)^2 + \frac{1}{4}.$$

Therefor we get that

$$\begin{aligned}
\frac{1}{2}(1-\cos(2t)) &= \frac{1}{2}\left(1 - \cos\left(2\frac{1}{2}\arcsin(2v_1)\right)\right) = \frac{1}{2}\left(1 - \sqrt{1-4v_1^2}\right) \\
&= \frac{1}{2}\left(1 - \sqrt{1 + 4\left(\left(\frac{1}{2} - v_2\right)^2 - \frac{1}{4}\right)}\right) = \frac{1}{2}\left(1 - \sqrt{1 - 1 + 4\left(\frac{1}{2} - v_2\right)^2}\right) \\
&= \frac{1}{2}\left(1 - 2\left(\frac{1}{2} - v_2\right)\right) = \frac{1}{2}\cdot 2v_2 = v_2.
\end{aligned}$$

Hence $v = (v_1, v_2) = (\frac{1}{2}\sin(2t), \frac{1}{2}(1-\cos(2t))) \in X$

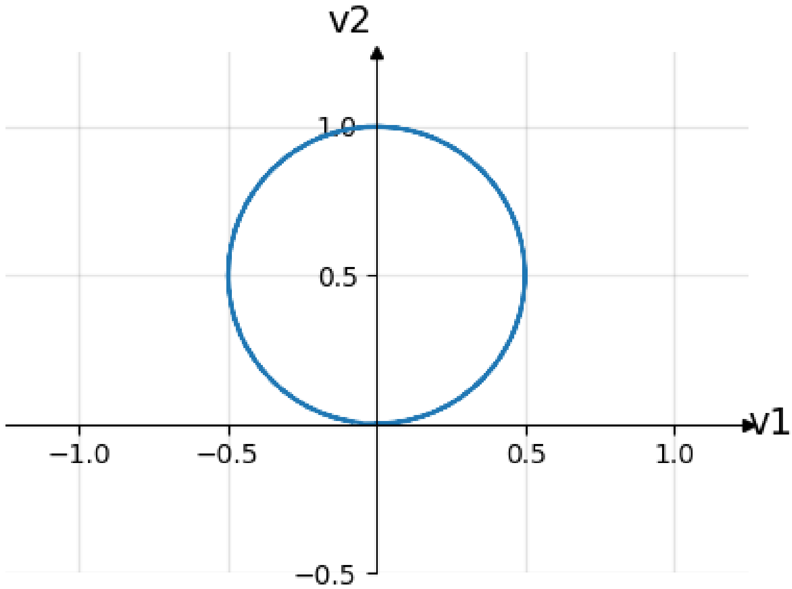


Figure 1: The algebraic curve given by $V(x^2-(1/2-y)^2-1/4)$ is a circle of radius $1/2$ and center $(0,1/2)$

Definition 4.1.18. An algebraic subset of an $\mathbb{A}^2$ is called an *algebraic curve*.

Definition 4.1.19. The vanishing set of a non-zero polynomial $f \in K[x_1,\dots,x_n]$ is called a *hypersurface*. If $\deg\ f = 1$, then its vanishing set is called a *hyperplane*.

Definition 4.1.20. A hypersurface in $\mathbb{A}^2$ is called an *affine plane curve*. A hyperplane in $\mathbb{A}^2$ is called a *line*.

Remark 4.1.21. Consider a line $V(f) \subset \mathbb{A}^2$, where $f = ax+by+c$, for $a,b,c \in K$ where a or b are non-zero. Without loss of generality we may assume that $b \neq 0$. Thus

$$V(f) = V(f) \cup \emptyset = V(f) \cup V\left(b^{-1}\right) = V\left(b^{-1}(ax+by+c)\right) = V\left(y+ab^{-1}x+cb^{-1}\right).$$

This means that any can be expressed as the vanishing set of some polynomial of the form

$$y - Ax + B,$$

for some $A, B \in K$.

Proposition 4.1.22. *Consider* $l = y - (ax+b) \in K[x,y]$ *and a line* $L = V(l)$ *and an affine plane curve* $C = V(f)$*, where* $f \in K[x,y]$ *with* $n = \deg f$ *such that* $L \not\subset C$*. Then* $C \cap L$ *is a finite set containing at most* n *points.*

Proof. The case where $\#K < \infty$ is trivial. Suppose then that $\#K = \infty$. Set

$$X := \{(t, at+b) \in \mathbb{A}^2 : f(t, at+b) = 0\}.$$

Note that $t \in V(f(x, ax+b))$ if and only if $(t, at+b) \in X$ and thus that there is a $t \in \mathbb{A}^1 \setminus V(f(x, ax+b)$, since $L \not\subset C$. This then implies that $\#V(f(x, ax+b)) \leq n$ (cf. Example 4.1.14), and hence $\#X \leq n$. It is then sufficient to prove that $C \cap L = X$. Let $v = (v_1, v_2) \in C \cap L$. Since $v \in L$ we have that

$$v_2 = av_1 + b,$$

meaning $v = (v_1, v_2) = (v_1, av_1 + b) \in X$. Conversely, let $v = (t, at+b) \in X$. Then

$$l(v) = at + b - (at+b) = 0,$$

hence $v \in L$. Clearly we also have that $v \in C$, meaning $v \in C \cap L$. □

Corollary 4.1.23. *An algebraic curve* $X \subset \mathbb{A}^2$ *intersects a line* L *not contained in* X *in only finitely many points.*

Proof. $X = V(f_1, \dots, f_m)$, for some $f_1, \dots, f_m \in K[x,y]$, where without loss of generality $f_1 \neq 0$. Then $X \subset V(f_1)$ and $L \not\subset V(f_1)$, which by Proposition 4.1.22 implies that $\#X \leq \#V(f_1) \leq \deg f_1 < \infty$. □

We use this result to give some non-examples of algebraic subsets

Example 4.1.24.

1. Consider the graph of the sine function,

$$G := \{(t, \sin(t)) \in \mathbb{A}^2(\mathbb{R}) : t \in \mathbb{R}\}$$

   Consider also the line $L := V(y)$. One easily verifies that $L = \{(t,0) \in \mathbb{A}^2(\mathbb{R}) : t \in \mathbb{R}\}$. Thus $G \cap L = \{(n\pi, 0) \in \mathbb{A}^2(\mathbb{R}) : n \in \mathbb{Z}\} \neq L$, hence $\#(G \cap L) = \infty$, and G is not algebraic by Corollary 4.1.23.

2. Consider the sphere

$$S := \left\{(z,w) \in \mathbb{A}^2(\mathbb{C}) : |z|^2 + |w|^2 = 1\right\}.$$

   Consider again the line $L := V(y)$. Note that $S \cap L = \left\{(z,0) \in \mathbb{A}^2(\mathbb{C}) : |z| = 1\right\} \neq L$ and that this set is in bijection with S^1. But clearly $\#S^1 = \infty$, hence S cannot be algebraic by Corollary 4.1.23.

Definition 4.1.25. Given an algebraic set $X \subset \mathbb{A}^n$ and $(b_1, \ldots, b_n) \in \mathbb{A}^n$, a *translation* is a map

$$\begin{aligned} \varphi : X &\to \mathbb{A}^n \\ (v_1, \ldots, v_n) &\mapsto (v_1 + b_1, \ldots, v_n + b_n) \end{aligned}$$

Lemma 4.1.26. *Algebraic sets are closed under translation. In other words, let $X := V(f_1, \ldots, f_m) \subset \mathbb{A}^n$ be an algebraic set. Let $b_1, \ldots, b_n \in K$. Then the image of the map*

$$\begin{aligned} \varphi : X &\to \mathbb{A}^n \\ (v_1, \ldots, v_n) &\mapsto (v_1 + b_1, \ldots, v_n + b_n) \end{aligned}$$

is an algebraic set

Proof. One checks that im $\varphi = V(f_1(x_1 - b_1, \ldots, x_n - b_n), \ldots, f_m(x_1 - b_1, \ldots, x_n - b_n))$. Indeed, let $(v_1 + b_1, \ldots, v_n + b_n) \in$ im φ. Then

$$f_i(v_1 + b_1 - b_1, \ldots, v_n + b_n - b_n) = f_i(v_1, \ldots, v_n) = 0.$$

On the other hand, let

$$(v_1, \ldots, v_n) \in V(f_1(x_1 - b_1, \ldots, x_n - b_n), \ldots, f_m(x_1 - b_1, \ldots, x_n - b_n)),$$

Then $(v_1 - b_1, \ldots, v_n - b_n) \in V(f_1, \ldots, f_n)$ and

$$(v_1, \ldots, v_n) = \varphi(v_1 - b_1, \ldots, v_n - b_n) \Rightarrow (v_1, \ldots, v_n) \in \text{im } \varphi.$$

□

Lemma 4.1.27. *Let $X \subset \mathbb{A}^n$ be algebraic and $\omega \in \mathcal{S}_n$. Then the image of*

$$\begin{aligned} \varphi : X &\to \mathbb{A}^n \\ (v_1, \ldots, v_n) &\mapsto (v_{\omega(1)}, \ldots, v_{\omega(n)}) \end{aligned}$$

is algebraic.

Proof. For suitable $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$, $X = V(f_1, \ldots, f_m)$. Clearly,

$$\textbf{im } \varphi = V\left(f_1\left(x_{\omega^{-1}(1)}, \ldots, x_{\omega^{-1}(n)}\right), \ldots, f_m\left(x_{\omega^{-1}(1)}, \ldots, x_{\omega^{-1}(n)}\right)\right).$$

Indeed, if $w = (v_{\omega(1)}, \ldots, v_{\omega(n)}) \in \textbf{im } \varphi$. Then

$$f_i\left(w_{\omega^{-1}(1)}, \ldots, w_{\omega^{-1}(n)}\right) = f_i\left(v_{\omega^{-1}(\omega(1))}, \ldots, v_{\omega^{-1}(\omega(n))}\right) = f_i(v_1, \ldots, v_n) = 0.$$

Conversely, if $v = (v_1, \ldots, v_n)$ is an element of the right-hand side, then picking $w = \left((v_{\omega^{-1}(1)}, \ldots, v_{\omega^{-1}(n)}\right)$, then $f_i(w) = 0$ and $\varphi(w) = v$. □

Definition 4.1.28. We define a *line* in $\mathbb{A}^n$ given by $a_1,\dots,a_n,b_1,\dots,b_n \in K$, where $a_j \neq 0$ for some j to be the set

$$\left\{\begin{pmatrix} a_1t+b_1 \\ \vdots \\ a_nt+b_n \end{pmatrix} \in \mathbb{A}^n : t \in K\right\}$$

Lemma 4.1.29. *A line in $\mathbb{A}^n$ is Zariski closed.*

Proof. We consider first the case where $a_1 \neq 0$. It is easy to check that

$$V(\{a_1x_i - a_ix_1 - a_1b_i + a_ib_1 : i \neq 1\}) = \left\{\begin{pmatrix} a_1t+b_1 \\ a_2t+b_2 \\ \vdots \\ a_nt+b_n \end{pmatrix} \in \mathbb{A}^n : t \in K\right\}.$$

Any other line can be obtained as the image translation composed with a map induced by a permutation of a line of the above form, hence by Lemmas 4.1.26 and 4.1.27 any line is Zariski closed. □

The intersection of a proper configuration of $n-1$ hyperplanes is a line for $n \geq 2$.

Proposition 4.1.30. *Consider a hyperplanes $V(p_1+b_i),\dots,V(p_{n-1}+c_{n-1})$, with $p_i = \sum_1^n a_{ij}x_j \in K[x_1,\dots,x_n]$, $b_i \in K$ such that the vectors*

$$a_i := \begin{pmatrix} a_{i1} \\ \vdots \\ a_{in} \end{pmatrix}, \quad (i \in \{1,\dots,n-1\})$$

are linearly independent. The intersection of these hyperplanes, L say, is a line.

Proof. Our assumptions lets us show that WLOG

$$L = \left\{ v \in \mathbb{A}^n : \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{(n-1)1} & \cdots & a_{(n-1)n} \\ 0 & \cdots & 0 \end{pmatrix} v = \begin{pmatrix} b_1 \\ \vdots \\ b_{n-1} \\ 0 \end{pmatrix} \right\} = \left\{ v \in \mathbb{A}^n : \begin{pmatrix} c_{ij} \end{pmatrix} v = \begin{pmatrix} b'_1 \\ \vdots \\ b'_{n-1} \\ 0 \end{pmatrix} \right\}$$

where $c_{ii} = 1$ for $i \in \{1,\dots,n-1\}$, $c_{nn} = 0$ and $c_{ij} = 0$ for $i,j \in \{1,\dots,n\}$ where $i > j$ and $n > j > i$. In other words

$$L = \bigcap_1^{n-1} V(x_i - c_ix_n - b'_i),$$

for suitable $c_i', b_i' \in K$. This means that

$$L = \{(c_1 t + b_1', \dots, c_{n-1} t + b_{n-1}', t) \in \mathbb{A}^n : t \in K\}.$$

$\square$

We can extend the result in Proposition 4.1.22 in the following way

Proposition 4.1.31. *Consider also a hypersurface $H := V(f) \subset \mathbb{A}^n$ with $d := \deg f > 0$ and a line $L \subset \mathbb{A}^n$ with $n \geq 2$ such that $L \not\subset H$. Then $\#(L \cap H) \leq d$*

Proof. For suitable $a_1, \dots, a_n, b_1, \dots, b_n \in K$ with $a_j \neq 0$ for some j we have that

$$L = \left\{ \begin{pmatrix} a_1 t + b_1 \\ \vdots \\ a_n t + b_n \end{pmatrix} \in \mathbb{A}^n : t \in K \right\}$$

We may then easily prove that

$$L \cap H = \left\{(a_1 t + b_1, \dots, a_n t + b_n) \in \mathbb{A}^n : f(a_1 t + b_1, \dots, a_n t + b_n) = 0\right\}$$

and that $(a_1 t + b_1, \dots, a_n t + b_n) \in L \cap H$ if and only if

$$t \in V(f(a_1 x + b_1, \dots, a_n x + b_n)).$$

Since $L \cap H \neq L$, there is a $t \in \mathbb{A}^1$ such that $t \notin V(f(a_1 x + b_1, \dots, a_n x + b_n))$, hence

$$\#(L \cap H) = \#V(f(a_1 x + b_1, \dots, a_n x + b_n)) \leq d$$

$\square$

Corollary 4.1.32. *Consider a line $L \in \mathbb{A}^n$. Then L intersects any algebraic set not containing L only finitely many times.*

Example 4.1.33. Consider the helix

$$H := \left\{(\cos(t), \sin(t), t) \in \mathbb{A}^3(\mathbb{R}) : t \in \mathbb{R}\right\}.$$

Consider also the line $L := \{(0, 1, t) \in \mathbb{A}^3(\mathbb{R}) : t \in \mathbb{R}\}$ and note that $L \cap H = \{(0, 1, 2n\pi + \pi/2) : n \in \mathbb{Z}\} \neq \mathbb{A}^1$, but we also have that $\#L \cap H = \infty$, hence H cannot be algebraic by 4.1.32.

Proposition 4.1.34. *Suppose K is an infinite field. Let $f \in K[x_1, \dots, x_n]$ with $\deg f > 0$*

1. *Suppose $n \geq 1$. Then $\#(\mathbb{A}^n \setminus V(f)) = \infty$.*

2. *Suppose that K is algebraically closed and that $n \geq 2$. Then* $\#V(f) = \infty$

Proof. 1. In the one variable case every algebraic subset that is not affine n-space is finite, hence $\#\left(\mathbb{A}^1 \setminus V(f)\right) = \infty$. Suppose then that $n \geq 2$. Since f is non-constant, we can write (cf. ref)

$$f = \sum_0^d f_i x_n^i,$$

for a $d \geq 1$ and for suitable $f_i \in K[x_1, \ldots, x_{n-1}]$ with $f_j \neq 0$ for some $j \in \{1, \ldots, d\}$. By Proposition **??** there is a point $v \in \mathbb{A}^{n-1}$ such that $f_j(v) \neq 0$. Hence

$$f' := \sum_0^d f_i(v) x_n^i \in K[x_n]$$

is a non-zero polynomial in one variable, implying that that that for every infinitely many $v_n \in K$ such that $f'(v_n) = f(v_1, \ldots, v_n) \neq 0$, hence $\mathbb{A}^n \setminus V(f)$ is infinite.
2. We again write $f = \sum_0^d f_i x_n^i$, for suitable $f_i \in K[x_1, \ldots, x_{n-1}]$, with $f_j \neq 0$ for some $j \in \{1, \ldots, d\}$. If every non-zero f_i is constant, then for any choice of $v \in \mathbb{A}^n$, there exists $a_v \in K$ such that for $f_v = \sum_0^d f_i(v) x_n^i$,

$$0 = f_v(a_v) = f(v_1, \ldots, v_{n-1}, a_v).$$

If $\deg f_j > 0$ for some j. Then there are infinitely many $v \in \mathbb{A}^{n-1}$ such $f_j(v) \neq 0$. For such a v there is at least one $a_v \in K$ that is a root in $f(v_1, \ldots, v_{n-1}, x_n)$. Thus in any case f has infinitely many zeroes, i.e. $V(f)$ is infinite. □

We present a way of constructing algebraic sets via Cartesian product

Proposition 4.1.35. *Let $X \subset \mathbb{A}^n$ and $Y \subset \mathbb{A}^m$ be algebraic then. $X \times Y \subset \mathbb{A}^{n+m}$ is algebraic.*

Proof. For suitable $f_1, \ldots, f_k \in K[x_1, \ldots, x_n]$ and $g_1, \ldots, g_l \in K[y_1, \ldots, y_m]$, we have $X = V(f_1, \ldots, f_m)$ and $Y = V(g_1, \ldots, g_l)$. We prove that $X \times Y = V(f_1, \ldots, f_k, g_1, \ldots, g_l)$, where we consider $f_1, \ldots, f_k, g_1, \ldots, g_l$ as elements of $K[x_1, \ldots, x_n, y_1, \ldots, y_m]$. Let $(v, w) \in X \times Y$. then clearly $f_i(v, w) = 0$ and $g_j(v, w) = 0$ for every $i \in \{1, \ldots, k\}$ and every $j \in \{1, \ldots, l\}$. Let $(v, w) \in V(f_1, \ldots, f_k, g_1, \ldots, g_l)$. Considering $f_1, \ldots, f_k$ and $g_1, \ldots, g_l$ as elements of the subsrings $K[x_1, \ldots, x_n]$ and $K[y_1, \ldots, y_m]$ respectively. We easily see that $f_i(v) = 0$ for every i and $g_j(w) = 0$ for every j, hence $v \in X$ and $w \in Y$, meaning $(v, w) \in X \times Y$. □

4.1.3 A Correspondence between Algebraic sets and Polynomial Ideals

Definition 4.1.36. Let $X \subset \mathbb{A}^n$ be any subset. We define the *ideal of* X to be the set

$$I^{\mathbb{A}}(X) := I(X) := \left\{ f \in K[x_1, \ldots, x_n] : f(v) = 0 \text{ for every } v \in \mathbb{A}^n \right\}$$

Lemma 4.1.37. *The ideal of any subset $X \subset \mathbb{A}^n$ is an ideal in $K[x_1, \ldots, x_n]$.*

Proof. Let $f, g \in I(X)$ and $h \in K[x_1, \ldots, x_n]$. Let $v \in X$. Note first that $0 \in K[x_1, \ldots, x_n]$ trivially vanishes on v, hence $0 \in I(X)$. Furthermore we have that

$$(f+g)(v) = f(v) + g(v) = 0 \Rightarrow f + g \in I(X)$$

and that

$$(hf)(v) = h(v)f(v) = h(v) \cdot 0 = 0 \Rightarrow hf \in I(X).$$

□

Example 4.1.38. We consider some initial examples of ideals of subsets of $\mathbb{A}^n$.

1. $I(\emptyset) = K[x_1, \ldots, x_n]$. For $f \in K[x_1, \ldots, x_n]$, the statement, f vanishes on every $v \in \emptyset$ is vacuously true, hence $1 \in I(\emptyset)$.

2. Suppose $\#K = \infty$. Then $I(\mathbb{A}^n) = 0$. Since K is infinite, if $f \in I(\mathbb{A}^n)$, then $f(v) = 0$ for every $v \in \mathbb{A}^n$, hence $f = 0$ by Proposition **??**

3. Consider a point $v \in \mathbb{A}^n$. Then $I(\{v\}) = \langle x_1 - v_1, \ldots, x_n - v_n \rangle$. This follows from Proposition 2.9.38.

Lemma 4.1.39. *Let $X, Y \subset \mathbb{A}^n$ with $X \subset Y$. Then $I(X) \supset I(Y)$.*

Proof. Let $f \in I(Y)$, and let $v \in X$. Then $v \in Y$, hence $f(v) = 0$, which implies $f \in I(X)$. □

Lemma 4.1.40. *Let $M \subset K[x_1, \ldots, x_n]$ and $X \subset \mathbb{A}^n$. Then we have the following*

1. $I(V(M)) \supset M$.

2. $V(I(X)) \supset X$.

3. $V(I(V(M))) = V(M)$. *Hence if X is algebraic $X = V(I(X))$.*

4. $I(V(I(X))) = I(X)$. *Hence if M is an ideal of some algebraic set, $M = I(V(M))$*

Proof. 1. Let $f \in M$. Let $v \in V(M)$. Then $f(v) = 0$, hence $f \in I(V(M))$.
2. Let $v \in X$. Let $f \in I(X)$. Then $f(v) = 0$, hence $v \in V(I(X))$.
3. By 1. $I(V(M)) \supset M$, hence $V(M) \supset V(I(V(M)))$ by Lemma 4.1.5. By 2. $V(I(V(M))) \supset V(M)$.
4. Since $V(I(X)) \supset X$ by 2, it follows that $I(V(I(X))) \subset I(X)$ by Lemma 4.1.39. By 1. $I(V(I(X))) \supset I(X)$. □

Lemma 4.1.41. *Let $X \subset \mathbb{A}^n$. Then $I(X)$ is radical.*

Proof. Let $F \in \mathbf{rad}(I(X))$, then $F^n \in I(X)$ for some $n > 0$. Let $v \in X$. Then

$$F(v)^n = \left(F^n\right)(v) = 0.$$

Since K is an integral domain, this implies that $F(v) = 0$, hence
$F \in I(X)$. □

Lemma 4.1.42. *Let $X, Y \subset \mathbb{A}^n$ be algebraic subsets. Then*

$$X = Y \iff I(X) = I(Y).$$

Proof. "$\Rightarrow$": Follows from Lemma 4.1.39.
"$\Leftarrow$": By Lemma 4.1.40 3.

$$X = V(I(X)) = V(I(Y)) = Y.$$

□

Corollary 4.1.43. *Let $X \subsetneq \mathbb{A}^n$ be an algebraic subset.*

1. *Consider $p \in \mathbb{A}^n \setminus X$. Then there is some polynomial $f \in K[x_1, \ldots, x_n]$ such that $f(q) = 0$ for every $q \in X$ and $f(p) = 1$.*

2. *Consider distinct points $p_1, \ldots, p_k \in \mathbb{A}^n \setminus X$. Then there are polynomials $f_1, \ldots, f_k \in I(X)$ such that $f_i(p_j) = 0$ whenever $i \neq j$ and $f_i(p_i) = 1$.*

3. *Let $p_1, \ldots, p_k \in \mathbb{A}^n \setminus X$ and $a_{ij} \in K$ for $i, j \in \{1, \ldots, k\}$. Then there are $G_i \in I(X)$ with $G_i(p_j) = a_{ij}$ for every i and j.*

Proof. 1.Note that $X \subsetneq X \cup \{p\}$. By Lemma 4.1.39 and Lemma 4.1.42, this implies $I(X) \supsetneq I(X \cup \{p\})$. Hence there is some $g \in I(X) \setminus I(X \cup \{p\})$. Clearly $g(q) = 0$ for every $q \in X$, while $g(q) \neq 0$, for otherwise $g \in I(X \cup \{p\})$. Upon taking

$$f = g(p)^{-1} g \in K[x_1, \ldots, x_n],$$

we are done.
2. Let $i \in \{1,\dots,k\}$ and put

$$Y = X \cup \bigcup_{j\in\{1,\dots,k\}:j\neq i} \{p_j\}.$$

Using 1. we can then find $f_i \in K[x_1,\dots,x_n]$ such that $f_i(v) = 0$ for every $v \in Y$ and $f_i(p_i) = 1$. In particular, $f_i(p_j) = 0$ whenever $i \neq j$.
3. We construct $f_1,\dots,f_k \in I(X)$ as in 2. Set

$$G_i = \sum_{h=1}^{k} a_{ij} f_j \in I(X).$$

Then

$$G_i(p_j) = \sum_{h=1}^{k} a_{ih} f_h(p_j) = a_{ij} f_j(p_j) = a_{ij}.$$

□

From Lemma 4.1.42 we get that

$$I(\bullet) : \tau_{\mathcal{Z}} \ni X \mapsto I(X) \in \{I \subset K[x_1,\dots,x_n] : I \text{ is a radical ideal}\},$$

is an injective well defined function. A simple counter example shows that it is not surjective

Example 4.1.44. Consider $I := \langle x^2+1\rangle \subset \mathbb{R}[x]$. We prove that I is prime and therefor radical by Lemma 2.8.30. Put $f = x^2+1$ and let $f_1, f_2 \in \mathbb{R}[x]$ such that $f_1 f_2 \in I$, i.e. such that $f \mid f_1 f_2$. Note that f has no roots in $\mathbb{R}$ and therefor is irreducible in $\mathbb{R}[x]$. Since $\mathbb{R}[x]$ is a UFD it follows that $f \mid f_1$ or $f \mid f_2$, hence $f_1 \in I$ or $f_2 \in I$. From the fact that x^2+1 vanishes on no points in $\mathbb{R}$, it follows that $x^2+1 \notin I(X)$ for any non-empty $X \subset \mathbb{R}$, hence $I(X) \neq I$. Since $\deg\ 1 = 0$ it follows that $x \notin I$, hence $I \subsetneq I(\emptyset) = \langle 1\rangle = \mathbb{R}[x]$. In conclusion, $I \neq I(X)$ for any set $X \subset \mathbb{R}$.

Lemma 4.1.45. *Let $I \subset K[x_1,\dots,x_n]$ be an ideal. Then $V(I) = V(\mathbf{rad}(I))$.*

Proof. Since $I \subset \mathbf{rad}(I)$, $V(I) \supset V(\mathbf{rad}(I))$ by Lemma 4.1.5. Let $v \in V(I)$ and let $f \in \mathbf{rad}(I)$. Then for some $n > 0$, $f^n \in I$. Then $0 = f(v)^n$, and since K is an integral domain, this implies $f(v) = 0$, hence $v \in V(\mathbf{rad}(I))$. In conclusion $V(I) = V(\mathbf{rad}(I))$. □

Lemma 4.1.46. *Since $I \subset K[x_1,\dots,x_n]$ be an ideal. Then $\mathbf{rad}(I) \subset I(V(I))$.*

Proof. Let $f \in \mathbf{rad}(I)$. Then for some $n > 0$, $f^n \in I$. Let $v \in V(I)$. Then $f^n(v) = 0$ hence $f(v) = 0$, implying that $f \in I(V(I))$. □

Lemma 4.1.47. *Let $a_1,\dots,a_n \in K$ and $I := \langle x_1 - a_1,\dots,x_n - a_n\rangle$. Then I is maximal ideal and $K \simeq K[x_1,\dots,x_n]/I$ via the canonical embedding of K in $K[x_1,\dots,x_n]/I$*

Proof. Proving that I is maximal is equivalent to proving that $K[x_1,\dots,x_n]/I$ is a field. Hence if we prove that $K \simeq K[x_1,\dots,x_n]/I$ the first claim follows. Consider the canonical embedding of K in $K[x_1,\dots,x_n]/I$

$$\iota : K \to K[x_1,\dots,x_n]/I$$
$$a \to a + I$$

which is a injective ring homomorphism by Lemma Make lemma. It remains to check that ι is surjective. Let $f + I \in K[x_1,\dots,x_n]/I$. Then $f + I = f(a_1,\dots,a_n) + I$ by Make Lemma?, hence

$$\iota(f(a_1,\dots,a_n)) = f(a_1,\dots,a_n) + I = f + I.$$

□

4.2 Affine Varieties

Definition 4.2.1. An algebraic subset $X \subset \mathbb{A}^n$ is said to be *reducible* if there are algebraic subsets $Y, Z \subsetneq X$ such that $X = Y \cup Z$. An algebraic subset $V \subset \mathbb{A}^n$ that is not reducible is said to be *irreducible*. An irreducible algebraic subset is also called an *(affine) variety.*

Remark 4.2.2. If $f \in K[x_1,\dots,x_n]$ is a reducible polynomial. Then $V(f)$ is reducible. Indeed, $f = gh$ for some non-constant $g,h \in K[\mathbf{x}]$, hence $V(f) = V(gh) = V(g) \cup V(h)$. We also have that $\langle f\rangle \subsetneq \langle g\rangle$ and $\langle f\rangle \subsetneq \langle h\rangle$, since $g,h \mid f$ and $\deg\ g, \deg\ h < \deg\ f$. We shall later see an example of a variety $V(f)$ where f is not irreducible.

Proposition 4.2.3. *Let $V \subset \mathbb{A}^n$ be a finite non-empty algebraic set. Then V is a variety if and only if V is a point*

Proof. "$\Rightarrow$ ":Suppose V is not a point. Then $V = \{p_1,\dots,p_m$ for $m > 1$, hence $V = \{p_1\} \cup \bigcup_2^m \{p_i\}$. Noting that $\{p_1\}$ and $\bigcup_2^m \{p_i\}$ are disjoint algebraic sets contained in V, it follows that V is reducible.
"$\Leftarrow$": Suppose V is a point. Then if $V = X \cup Y$ for algebraic sets $X, Y \subset \mathbb{A}^n$, then either $X = V$ or $Y = V$, hence V is a variety. □

Definition 4.2.4. Let $V \subset \mathbb{A}^n$ be a variety. We define *the coordinate ring of V* to be the ring $\Gamma(V) := K[x_1,\dots,x_n]/I(V)$.

Proposition 4.2.5. *Let $V \subset \mathbb{A}^n$ be an algebraic set. The following are equivalent:*

1. *V is a variety.*

2. *$I(V)$ is a prime ideal.*

3. *$\Gamma(V)$ is an integral domain.*

Proof. "1.$\Rightarrow$ 2.": Suppose $I(V)$ is not prime. Then there are $f_1, f_2 \in K[x_1, \dots, x_n] \setminus I(V)$ such that $f_1 f_2 \in I(V)$ then $V \subset V(f_1 f_2)$, hence

$$V = V \cap V(f_1 f_2) = V \cap (V(f_1) \cup V(f_2)) = (V \cap V(f_1)) \cup (V \cap V(f_2)).$$

Since $f_1, f_2 \notin I(V)$, $I(V) \subsetneq I(V) + \langle f_i \rangle$ for $i = 1, 2$, hence

$$V \subsetneq V(I(V) + \langle f_i \rangle) = V(I(V)) \cap V(f_i) = V \cap V(f_i),$$

meaning V is reducible.
"2.$\Rightarrow$ 1.": Suppose V is reducible. Then $V = V_1 \cup V_2$ for algebraic subsets $V_1, V_2 \subsetneq V$. Then $I(V_1), I(V_2) \supsetneq I(V)$, meaning that there exists $f_1 \in I(V_1) \setminus I(V)$ and $f_2 \in I(V_2) \setminus I(V)$. Furthermore, $f_1 f_2$ vanish on ever point in $V_1 \cup V_2 = V$, hence $f_1 f_2 \in I(V)$, meaning $I(V)$ is not prime.
"2. $\iff$ 3.": This follows from Lemma 2.8.22. $\square$

We demonstrate the strength of the above theorem in the following example

Example 4.2.6. 1. Let $f = y - x^2 \in \mathbb{C}[x, y]$. Since $\deg_y f = 1$ and f is monic with respect to x, f is irreducible in $(\mathbb{C}[y])[x] \simeq \mathbb{C}[x, y]$ Result!!!, hence since $\mathbb{C}[x, y]$ is a UFD Result!!!, f is prime, and thus $\langle f \rangle$ is prime. Suppose $a \in I(V(f))$. We aim to prove that $a(x, x^2) = 0$, which would imply

$$a(x, y) + \langle f \rangle = a(x, x^2) + \langle f \rangle = 0 + \langle f \rangle \Rightarrow a \in \langle f \rangle.$$

Indeed every point in $V(f)$ is of the form $(\alpha, \alpha^2) \in \mathbb{A}^2(\mathbb{C})$, hence upon putting $b = a(x, x^2) \in \mathbb{C}[x]$, we see that $b(\alpha) = 0$ for every $\alpha \in \mathbb{C}$, implying $a(x, x^2) = b = 0$. We thus conclude $I(V(f)) = \langle f \rangle$, which means $V(f) \subset \mathbb{A}^2(\mathbb{C})$ is irreducible by the above theorem.

2. Let $g = y^4 - x^2 \in \mathbb{C}[x, y]$ and $h = y^4 - x^2 y^2 + x y^2 - x^3 \in \mathbb{C}[x, y]$. Setting $g_1 = y^2 - x$ and $g_2 = y^2 + x$, then $g = g_1 g_2$. We can thus write

$$V(g, h) = V(g) \cap V(h) = V(g_1 g_2) \cap V(h) = (V(g_1) \cup V(g_2)) \cap V(h) = V(g_1, h) \cup V(g_2, h).$$

Note also that

$$h = y^4 - x^2y^2 + xy^2 - x^3 = y^2(y^2+x) - x^2(y^2-x) = -x^2g_1 + y^2g_2 \in \langle g_1, g_2\rangle.$$

We aim to prove that $V(g_i,h) = \{0\}$. Let $v \in V(g_1,h)$. Then $v_1 = v_2^2$. Then $h(v) = v_2^2 g_2(v)$, hence $0 = v_2^2 = v_1$, in which case $v = 0$, or $g_2(v) = 0$. Since $V(g_1) \cap V(g_2) = \{0\}$, it follows that in the second case $v = 0$. One proves $V(g_2,h) = \{0\}$ similarly. We may write $\{0\} = V(x,y)$. Trivially a singleton is irreducible, since if $\{0\} = V \cup W$, then $V = \{0\}$ or $W = \{0\}$.

3. One should note that a polynomial being irreducible, does not imply that its vanishing set is irreducible. Take for instance, $f = y^2 + x^2(x-1)^2 \in \mathbb{R}[x,y]$. Consider f as a polynomial in $\mathbb{C}[x,y]$ we have that

$$f = (y - ix(x-1))(y + ix(x-1)),$$

hence since the prime factorization of f is unique (since $\mathbb{C}[x,y]$ is a UFD) and $y - ix(x-1), y + ix(x-1) \in \mathbb{C}[x,y] \setminus \mathbb{R}[x,y]$, there is no non-trivial factors of f in $\mathbb{R}[x,y]$, which means f is irreducible in $\mathbb{R}[x,y]$. Note that

$$V(f) \supset \{(0,1)\} \cup \{(0,0)\} = V(x-1,y) \cup V(x,y).$$

Let $(v_1,v_2) \in V(f)$. Then $v_2^2 = -v_1^2(v_1-1)^2$. Note that $v_2^2 \geq 0$, and $v_1^2(v_1-1)^2 = (v_1(v_1-1))^2 \geq 0$, hence $v_2^2 = -v_1^2(v_1-1)^2 \leq 0$. This implies $v_2 = 0$, hence $v_1^2 = 0$ or $(v_1-1)^2 = 0$, implying $v_1 = 0$ or $v_1 - 1 = 0$. In conclusion $(v_1,v_2) = (0,0)$ or $(v_1,v_2) = (1,0)$, meaning $(v_1,v_2) \in \{(0,1)\} \cup \{(0,0)\}$. We thus see that

$$V(f) = V(x-1,y) \cup V(x,y).$$

Lemma 4.2.7. *Let X be a set and $\mathscr{X} \subset 2^X$ be a non-empty family of subsets of X. Set*

$$\mathscr{Y} := \left\{X \in \mathscr{X} : X \subset Y \text{ for some } Y \in \mathscr{X}\right\}$$

and $\mathscr{Z} := \mathscr{X} \setminus \mathscr{Y}$. Then

$$\bigcup_{X \in \mathscr{X}} X = \bigcup_{X \in \mathscr{Z}} X.$$

Proof. The inclusion $\bigcup_{X\in\mathscr{X}} X \supset \bigcup_{X\in\mathscr{Z}} X$ is trivial. Let $x \in \bigcup_{X\in\mathscr{X}} X$. Then $x \in X$, for some $X \in \mathscr{X}$. To prove the inclusion, we need to show that $x \in Y$ for some $Y \in \mathscr{Z}$ □

Theorem 4.2.8. *Let $X \subset \mathbb{A}^n$ be an algebraic set. Then there is a finite sequence of varieties $V_1, \ldots, V_n \subset \mathbb{A}^n$ unique (up to re-ordering) such that*

$$X = \bigcup_1^n V_i$$

and $V_i \not\subset V_j$ for $i \neq j$.

Proof. **Existence:** Consider the set

$$\mathscr{A} := \{X \subset \mathbb{A}^n : X \text{ is algebraic and not a finite union of varieties}\}.$$

Suppose for a contradiction that $\mathscr{A} = \emptyset$. Then by Lemma **??**, there is a minimal element $X_0 \in \mathscr{A}$. On the one hand X_0 is not irreducible. Hence There are $Y, Z \subsetneq X_0$ such that $X_0 = Y \cup Z$. However, by minimality $Y, Z \notin \mathscr{A}$, hence $Y = \bigcup_1^k V_i$ and $Z = \bigcup_1^l W_i$ for varieties $V_1, \ldots, V_k, W_1, \ldots, W_l \subset \mathbb{A}^n$. But then

$$X_0 = Y \cup Z = \bigcup_1^k V_i \cup \bigcup_1^l W_i,$$

leading to a contradiction. This means $\mathscr{A} = \emptyset$. Hence every algebraic set can be written as a union of varieties.

Let $X \subset \mathbb{A}^n$ be an algebraic set, and pick varieties $V_1, \ldots, V_m \subset \mathbb{A}^n$ whose union is equal to X. Let $V_{k_1}, \ldots, V_{k_l}$ be the maximal elements of $\{V_1, \ldots, V_m\}$. Let $j \in \{1, \ldots, m\}$. The set $S := \{V_i : V_j \subset V_i\}$ has a maximal element V_k. Let $h \in \{1, \ldots, m\}$ and suppose $V_k \subset V_h$. Then in particular $V_j \subset V_h$, hence $V_h \in S$, implying $V_h = V_k$. This means V_k is a maximal element of $\{V_1, \ldots, V_m\}$. In other words $k = k_i$ for some i, hence $V_j \subset V_{k_i}$, implying

$$X = \bigcup_1^m V_i = \bigcup_1^l V_{k_i}$$

Uniqueness: Suppose $X \subset \mathbb{A}^n$ be algebraic sets. Suppose there are varieties $V_1, \ldots, V_l, W_1, \ldots, W_m \subset \mathbb{A}^n$ such that

$$\bigcup_1^l V_i = X = \bigcup_1^m W_i,$$

and $V_i \not\subset V_j$ for $j \neq i$ and $W_i \not\subset W_j$ for $i \neq j$. For each $i \in \{1, \ldots, l\}$ we have that

$$V_i = V_i \cap V = V_i \cap \bigcup_1^m W_j = \bigcup_1^m (V_i \cap W_j).$$

Then there is a $j(i)$ such that $V_i \subset W_{j(i)}$. By similar argument, $W_{i(j)} \subset V_k$ for some k hence $i = k$, implying $V_i = W_{j(i)}$. similarly $W_j = V_{i(j)}$ for every j some $i(j)$. Hence there is a one-to-one correspondence of $\{V_1, \ldots, V_l\}$ and $\{W_1, \ldots, W_m\}$. □

Definition 4.2.9. The varieties $V_1, \ldots, V_m \subset \mathbb{A}^n$ involved in the *decomposition*, described in the above theorem, of an algebraic set $X \subset \mathbb{A}^n$ are called the *(irreducible) components* of X.

Proposition 4.2.10. *Let $X, Y \subset \mathbb{A}^n$ be algebraic subsets such that $X \subset Y$ with irreducible components $V_1, \ldots, V_m$ respectively $W_1, \ldots, W_l$. Then for each $i \in \{1, \ldots, m\}$, V_i is contained in W_j for some $j \in \{1, \ldots, l\}$.*

Proof. Let $v \in V_i$. Then since

$$\bigcup_1^m V_i = X \subset Y = \bigcup_1^l W_j,$$

$v \in W_j$ for some $j \in \{1,\ldots,l\}$, hence $V_i \subset W_j$. □

Proposition 4.2.11. *Let $X \subset \mathbb{A}^n$ be an algebraic subset with irreducible components $V_1,\ldots,V_m$. Then for each $i \in \{1,\ldots,m\}$,*

$$V_i \not\subset \bigcup_{j\in\{1,\ldots,m\}\setminus\{i\}} V_j.$$

Proof. If $m = 1$, then $\bigcup_{j\in\{1,\ldots,m\}\setminus\{i\}} V_j = \emptyset$, hence clearly $V_i \not\subset \bigcup_{j\in\{1,\ldots,m\}\setminus\{i\}} V_j$. Suppose $m \geq 2$, i.e. that X is reducible. Note that $X \setminus V_i \subset \bigcup_{j\in\{1,\ldots,m\}\setminus\{i\}} V_j$. Then since $V_i \subsetneq X$, there is a point $v \in X\setminus$... □

Proposition 4.2.12. *For an infinite field K, $\mathbb{A}^n(K)$ is irreducible.*

Proof. When K is infinite, $I(\mathbb{A}^n(K)) = 0$ ref result!!!, which is trivially prime, hence $\mathbb{A}^n(K)$ is prime by Theorem 4.2.5 □

4.2.1 Classifying Algebraic Subsets of the Plane

Theorem 4.2.13. *Let $f,g \in K[x,y]$ such that $\gcd(f,g) = 1$. Then $\#V(f,g) < \infty$.*

Proof. By Add result 1 is also greatest common divisor of f and g in $K(x)[y]$. since $K(x)[y]$ is a PID it follows that $\langle f,g\rangle = \langle 1\rangle \subset K(x)[y]$ Result. Then there are $\mu,\lambda \in K(x)[y]$ such that $\mu f + \lambda g = 1$. For some $\delta \in K[x]$, we have that $\alpha f + \beta g = \delta$. Hence if $(v_1,v_2) \in V(f,g)$, then $v_1 \in V(\delta)$. Since $\#V(\delta) < \infty$, we have that

$$V_1 := \{v_1 \in K : (v_1,v_2) \in V(f,g) \text{ for some } v_2 \in K\}$$

is finite. We can similarly show by a symmetric argument that

$$V_2 := \{v_2 \in K : (v_1,v_2) \in V(f,g) \text{ for some } v_1 \in K\}$$

is finite. One easily sees that $V(f,g) \subset V_1 \times V_2$, hence $\#V(f,g) < \infty$. □

Corollary 4.2.14. *Let $f \in K[x,y]$ be irreducible such that $\#V(f) = \infty$. Then $I(V(f)) = \langle f\rangle$ and $V(f)$ is irreducible.*

Proof. If $g \in I(V(f))$ then $V(f,g) \supset V(f)$, hence $\#V(f,g) = \infty$. Then f and g has a non-trivial common factor by the above theorem. Then since f is assumed irreducible this means $f \mid g$ or equivalently that $g \in \langle f\rangle$, hence $I(V(f)) \subset \langle f\rangle$, implying

$I(V(f)) = \langle f \rangle$ by Lemma 4.1.40, this implies that $I(V(f))$ is prime by Lemma 2.8.26 and the fact that $K[x,y]$ is a UFD. This implies that $V(f)$ is irreducible by Proposition 4.2.5. □

Corollary 4.2.15. *Suppose K is infinite. The varieties in $\mathbb{A}^2$ are $\mathbb{A}^2$, $\emptyset$, points and plane curves $V(f)$ where $f \in K[x,y]$ is irreducible and $\#V(f) = \infty$.*

Proof. Suppose $V \subset \mathbb{A}^2$ is a non-empty variety. By Proposition 4.2.3, V is finite if and only if V is a point. Suppose V is infinite. If $I(V) = 0$ then $V = \mathbb{A}^2$. So suppose $I(V)$ contains a non-constant f. Since $K[x,y]$ is a UFD, we can write f as a product irreducible factors $f_1, \cdots, f_m \in K[x,y]$, since $I(V)$ is prime by Proposition 4.2.5 it follows that $f_i \in I(V)$ for some i. Suppose for a contradiction Do without contradiction proof! that there is a $g \in I(V) \setminus \langle f_i \rangle$, then $\langle g, f_i \rangle \subset I(V)$ and $f_i \nmid g$, hence $\gcd(g, f_i) = 1$ and $V(g, f_i) \supset V(I(V)) = V$, implying that V would be finite by the above theorem. It thus follows that $I(V) = \langle f_i \rangle$, hence $V = V(I(V)) = V(f_i)$. □

Corollary 4.2.16. *Let K be an algebraically closed field, $f \in K[x,y]$, $\deg f > 0$. Let distinct irreducible polynomials $f_1, \ldots, f_l \in K[x,y]$ and positive integers $r_1, \ldots, r_l$ be given such that $f = \prod_1^r f_i^{r_i}$. Then $V_i := V(f_i)$ for $i = 1, \ldots, l$ are the irreducible components of $V := V(f)$ and $V(f) = \langle \prod_1^l f_i \rangle$.*

Proof. Proposition 4.1.34 2. shows that $V(f_i)$ is an infinite set for each i. Hence by Corollary 4.2.14 V_i is irreducible and $I(V_i) = \langle f_i \rangle$ for each i. Furthermore, for each i, $V\left(f_i^{r_i}\right) = V\left(\mathbf{rad}\left(f_i^{r_i}\right)\right) = V(f_i)$, hence $V = \bigcup_1^l V_i$. Since $f_i \nmid f_j$ for $i \neq j$, $\langle f_j \rangle \not\subset \langle f_i \rangle$, hence $V_i \not\subset V_j$. Note that

$$I\left(\bigcup_1^l V_i\right) = \bigcap_1^l I(V_i) = \bigcap_1^l \langle f_i \rangle.$$

It follows from Reference small lemma from algebra section that $\bigcap_1^l \langle f_i \rangle = \langle \prod_1^l f_i \rangle$. □

Example 4.2.17. We give some examples of why it's not too much fun to work over non-algebraically closed field

1. Let $f = x^2 + y^2 + 1 \in \mathbb{R}[x,y]$, $I = \langle f \rangle$. Note that for any $(v_1, v_2) \in \mathbb{A}^2(\mathbb{R})$, $v_1^2 + v_2^2 \geq 0$, hence $v_1^2 + v_2^2 + 1 \geq 1 > 0$, hence $V(I) = \emptyset$. Hence $I(V(I)) = I(\emptyset) = \langle 1 \rangle$ Let $f_1 = a_1 x + b_1 y + c_1, f_2 = a_2 x + b_2 y + c_2 \in \mathbb{R}[x,y]$. Then

$$f_1 f_2 = a_1 a_2 x^2 + b_1 b_2 y^2 + (a_1 b_2 + a_2 b_1) xy + (a_1 c_2 + a_2 c_1) x + (b_1 c_2 + b_2 c_1) y + c_1 c_2.$$

Consider the following system of equations

$$\begin{cases} a_1a_2 = \alpha & b_1b_2 = \beta \\ (a_1b_2 + a_2b_1) = 0 \end{cases}$$

where $\alpha,\beta > 0$ This system of equations clearly cannot have any real solutions since using $a_1a_2 = 1$ and $b_1b_2 = 1$, we get that

$$(a_1b_2 + a_2b_1) = 0 \iff a_1b_2 = -a_2b_1 \iff a_1^2b_2^2 = -a_2a_1b_1b_2 = -1.$$

It thus follows that any polynomial in $\mathbb{R}[x,y]$ of the form $\alpha x^2 + \beta y^2 + 0xy + \ldots \in \mathbb{R}[x,y]$ where α,β cannot be a product of two degree 1 polynomials in $\mathbb{R}[x,y]$ and therefor is irreducible in $\mathbb{R}[x]$. Alternatively one could also see this by applying the rational root theorem and result saying that second degree pol. over field is red. iff has roots to $x^2 + y^2 + 1 \in \mathbb{R}(x)[y]$. The point in any case is that $I(V(I)) \neq I$.

2. Every algebraic subset of $\mathbb{A}^2(\mathbb{R})$ can be written as $V(f)$ for some $f \in \mathbb{R}[x,y]$: If $P = \{(v_1,v_2)\}$, then $P = V(g)$, where $g = (x-v_2)^2 + (y-v_1)^2$. Indeed, $g(v_1,v_2) = 0$ and if $(w_1,w_2) \in V(g)$, then since $(w_1-v_1)^2, (w_2-v_2)^2 \geq 0$ and $(w_1-v_1)^2 + (w_2-v_2)^2 = 0$, we have $w_1 - v_1 = 0$ and $w_2 - v_2 = 0$. Now if V is an arbitrary algebraic subset of $\mathbb{A}^2(\mathbb{R})$, it has a decomposition into irreducible components $V_1,\ldots,V_l$. By Corollary 4.2.15 and the remark we made about points at the beginning, it follows that $V_i = V(f_i)$ for some $f_i \in \mathbb{R}[x,y]$, hence
$$V = \bigcup_1^l V(f_i) = V\left(\prod_1^l f_i\right).$$

Example 4.2.18. We have now classified the varieties of the plane over an infinite field K. Let's determine the irreducible decompositions of some algebraic subsets

1. Let $f = y^2 - xy - x^2y + x^3$. Set $f_1 = x - y$ and $f_2 = y - x^2$. Then one sees that $f = f_1f_2$ and hence $V(f) = V(f_1) \cup V(f_2)$. Eisenstein's criterion write!!! shows that $f_1 \in K(x)[y]$, $f_2 \in K(y)[x]$ are irreducible, since $x \nmid 1, x \mid x, x^2 \nmid x$ and $y \nmid 1, y \mid y, y^2 \nmid y$. We then have that (since f_1, f_2 are primitive Ref result) that $f_1, f_2 \in K[x,y]$ are irreducible (This is for any field K). Note that $V_{\mathbb{R}}(f_1) = \{(\lambda,\lambda) \in \mathbb{R}^2 : \lambda \in \mathbb{R}\}$, which is infinite. Note also that $V_{\mathbb{R}}(f_2) = \{(\lambda,\lambda^2) \in \mathbb{R}^2 : \lambda \in \mathbb{R}\}$, which is also infinite since it is the graph of the function $\lambda \mapsto \lambda^2$. We thus see that $V_{\mathbb{R}}(f_1), V_{\mathbb{R}}(f_2)$ and $V_{\mathbb{C}}(f_1), V_{\mathbb{C}}(f_2)$ are the irreducible components of $V_{\mathbb{R}}(f)$ resp. $V_{\mathbb{C}}(f)$ by Corollary 4.2.15, since $V_{\mathbb{C}}(f_1) \cap V_{\mathbb{C}}(f_2) = \{(0,0),(1,1)\}$ implying that neither set is contained in the other.

2. Consider $f = y^2 - x\left(x^2-1\right)$. The set $V(f)$ is an example of an elliptic curve (these are in general of the form $V\left(y^2 - \left(ax^3+bx^2+cx+d\right)\right)$. Using Eisenstein's criterion for f in $K[x][y]$ performing the testing the criterion against $x \in K[x]$, it follows that $f \in K[x,y]$ is irreducible. We aim to show that $V_{\mathbb{R}}(f)$ is infinite. Let $\lambda \in \mathbb{R}_{\geq 1}$. Then $\lambda^2 - 1 \geq 0$, hence $\lambda\left(\lambda^2-1\right) \geq 0$. It thus follows that
$$\left\{\left(\lambda, \pm\sqrt{\lambda\left(\lambda^2-1\right)}\right) \in \mathbb{R}^2 : \lambda \in \mathbb{R}\right\} \subset V_{\mathbb{R}}(f) \subset V_{\mathbb{C}}(f).$$
Since this subset is infinite, it follows that both $V_{\mathbb{R}}(f), V_{\mathbb{C}}(f)$ are irreducible by Corollary 4.2.15.

3. Consider $f = x^3 + x - x^2y - y$. Put $f_1 = x - y$, $f_2 = x^2+1$. Then $f = f_1 f_2$. We have already seen that f_1 is irreducible in $K[x,y]$ and that $V(f_1)$ is infinite. We also know that $V_{\mathbb{R}}(f_2) = \emptyset$. It thus follows that $V_{\mathbb{R}}(f) = V_{\mathbb{R}}(f_1) \cup V_{\mathbb{R}}(f_2) = V_{\mathbb{R}}(f_1)$, hence $V_{\mathbb{R}}(f_1)$ is the only irreducible component over $\mathbb{R}$. Note that $f_2 = (x+i)(x-i)$ in $\mathbb{C}[x,y]$, and $f_2' = x+i$, and f_2'' are irreducible in $\mathbb{C}[x,y]$ being of degree 1. Since $V_{\mathbb{C}}(f_2') = \left\{(-i,\lambda) \in \mathbb{C}^2 : \lambda \in \mathbb{C}\right\}$ and $V_{\mathbb{C}}(f_2'') = \left\{(i,\lambda) \in \mathbb{C}^2 : \lambda \in \mathbb{C}\right\}$ are infinite sets, it thus follows that $V_{\mathbb{C}}(f_2'), V_{\mathbb{C}}(f_2'')$ are also irreducible sets in $\mathbb{A}^2(\mathbb{C})$. Clearly $V_{\mathbb{C}}(f_2')$ and $V_{\mathbb{C}}(f_2'')$ are disjoint. Furthermore, $V_{\mathbb{C}}(f_1) \cap V_{\mathbb{C}}(f_2') = \{(-i,-i)\}$ and $V_{\mathbb{C}}(f_1) \cap V_{\mathbb{C}}(f_2'') = \{(i,i)\}$, hence there is no containment of these sets in each other. It thus follows that the irreducible components of $V_{\mathbb{C}}(f)$ are $V_{\mathbb{C}}(f_1), V_{\mathbb{C}}(f_2'), V_{\mathbb{C}}(f_2'')$.

4.2.2 Hilbert's Nullstellensatz

In the discussion involving Hilbert's Nullstellensatz, we shall assume that K is algebraically closed. Before presenting the full Hilbert Nullstellensatz we prove a weak version, which we shall see implies Hilbert's Nullstellensatz.

Theorem 4.2.19. *(The Weak Nullstellensatz/WNS)*
Let K be an algebraically closed field and consider $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$. Then $V(f_1, \ldots, f_m) = \emptyset$ if and only if there are polynomials $\lambda_1, \ldots, \lambda_m \in K[\mathbf{x}]$ such that
$$\sum_1^m \lambda_i f_i = 1.$$
In other words
$$V(f_1, \ldots, f_m) = \emptyset \iff \langle f_1, \ldots, f_m \rangle = \langle 1 \rangle = K[\mathbf{x}].$$

Proof. Set $I := \langle f_1, \dots, f_m \rangle$.
"$\Leftarrow$": If $I = \langle 1 \rangle$, then $V(I) = V(1) = \emptyset$.
"$\Rightarrow$ ": Suppose $I \neq \langle 1 \rangle$, i.e. I is a proper ideal. We aim to show $V(I) \neq \emptyset$. $I \subset I'$ for some maximal ideal by reference to proper theorem. Hence $V(I) \supset V(I')$. This means that if $V(I') \neq \emptyset$, then $V(I) \neq \emptyset$. It is then sufficient to prove the statement in the case where I is maximal. It follows from Corollary 2.10.28 that $I = \langle x_1 - a_1, \dots, x_n - a_n \rangle$ for suitable $a_i \in K$. Hence $V(I) = \{(a_1, \dots, a_n)\} \neq \emptyset$. □

Remark 4.2.20. The equation

$$\sum_1^m f_i y_i = 1 \tag{6}$$

is called *the Hilbert equation associated with* $f_1, \dots, f_m$. Thus WNS says that $f_1, \dots, f_m$ has no common zeros if and only if the Hilbert equation associated with $f_1, \dots, f_m$ is soluble over $K[\mathbf{x}]$

Theorem 4.2.21. *(Hilbert's Nullstellensatz/HNS)*
Let K be an algebraically closed field and consider $f_1, \dots, f_m \in K[x_1, \dots, x_n]$. Put $I := \langle f_1, \dots, f_m \rangle$. Then for every $f \in I(V(I))$ there are $\lambda_1, \dots, \lambda_n \in K[\mathbf{x}]$ and an integer $k \geq 0$ such that

$$\sum_1^m \lambda_i f_i = f^k.$$

In other words $I(V(I)) = \mathrm{rad}(I)$.

Proof. $\mathrm{rad}(I) \subset I(V(I))$ by Lemma 4.1.46. Let $g \in I(V(I))$. Put $J := \langle f_1, \dots, f_m, x_{n+1}g - 1 \rangle \subset K[x_1, \dots, x_{n+1}]$. Suppose $v \in \mathbb{A}^{n+1}$ vanishes on $f_1, \dots, f_m$. Then $v_{n+1}g(v) - 1 = -1 \neq 0$. If $v \in \mathbb{A}^{n+1}$ vanishes on $x_{n+1}g - 1$. Then $v_{n+1}g(v) = 1 \neq 0$, hence v doesn't vanish on $f_1, \dots, f_m$. This means $V(J) = \emptyset$. By the weak Nullstellensatz we can then find $\lambda_1, \dots, \lambda_{m+1} \in K[x_1, \dots, x_{n+1}]$ such that

$$1 = \left[\sum_1^m \lambda_i f_i\right] + \lambda_{m+1}(x_{n+1}G - 1),$$

Let $y = \frac{1}{g} \in K(x_1, \dots, x_{n+1})$. Then

$$\begin{aligned} 1 = \mathrm{ev}_{x_1,\dots,x_n,y}(1) &= \mathrm{ev}_{x_1,\dots,x_n,y}\left(\left[\sum_1^m \lambda_i f_i\right] + \lambda_{m+1}(x_{n+1}G - 1)\right) \\ &= \left[\sum_1^m \lambda_i(x_1, \dots, x_n, y) f_i\right] + \lambda_{m+1}(x_1, \dots, x_n, y)(yG - 1) = \sum_1^m \lambda_i(x_1, \dots, x_n, y) f_i. \end{aligned}$$

Then taking $N \geq \max_{i \in \{1,\dots,m\}}(\deg_{x_{n+1}} \ \lambda_i f_i)$,

$$g^N = g^N \sum_1^m \lambda_i(x_1, \dots, x_n, y) f_i = \sum_1^m \lambda_i' f_i,$$

for suitable $\lambda_i' \in K[x_1, \dots, x_n]$. This implies $g \in \mathrm{rad}(I)$ □

We give a few corollaries the first of which is apparent.

Corollary 4.2.22. *The mapping taking a radical ideal* $I \subset K[x_1,\dots,x_n]$ *to* $V(I) \subset \mathbb{A}^n$ *and the mapping taking an algebraic set* $V = V(I) \subset \mathbb{A}^n$ *to* $I(V) = \mathrm{rad}(I)$ *establishes a one-to-one correspondence between algebraic sets and polynomial radical ideals for each* $n \geq 1$.

Corollary 4.2.23. *The mapping taking a prime ideal* $I \subset K[x_1,\dots,x_n]$ *to* $V(I) \subset \mathbb{A}^n$ *and the mapping taking a variety* $V = V(I) \subset \mathbb{A}^n$ *to* $I(V) = \mathrm{rad}(I)$ *establishes a one-to-one correspondence between irreducible algebraic sets and polynomial prime ideals for each* $n \geq 1$.

Proof. Let $I \subset K[\mathbf{x}]$ be a prime ideal. By Lemma 2.8.30 and HNS,

$$I(V(I)) = \mathrm{rad}(I) = I, \tag{7}$$

hence by Proposition 4.2.5, $V(I)$ is indeed irreducible. Given a variety $V \subset \mathbb{A}^n$, also by Proposition 4.2.5, $I(V)$ is prime. Furthermore, $V(I(V)) = V$ by Lemma 4.1.40. This with (7) shows the mappings are mutual inverses. □

Corollary 4.2.24. *The mapping taking a maximal ideal* $I \subset K[x_1,\dots,x_n]$ *to* $V(I) \subset \mathbb{A}^n$ *and the mapping taking a point* $P = \{(a_1,\dots,a_n)\} = V(x_1 - a_1,\dots,x_n - a_n) \subset \mathbb{A}^n$ *to* $I(P) = \mathrm{rad}(x_1 - a_1,\dots,x_n - a_n)$ *establishes a one-to-one correspondence between and polynomial maximal ideals for each* $n \geq 1$.

Proof. Let $I \subset K[\mathbf{x}]$ be a maximal ideal. Then $I = \langle x_1 - a_1,\dots,x_n - a_n\rangle$, hence $V(I) = \{(a_1,\dots,a_n)\}$. Since I is maximal, it is prime, hence radical, and thus $I(V(I)) = I$ by HNS. Let $P = \{(a_1,\dots,a_n)\} \subset \mathbb{A}^n$. Then $I(P) = I(V(\langle x_1 - a_1,\dots,x_n - a_n\rangle)) = \langle x_1 - a_1,\dots,x_n - a_n\rangle$, which is maximal, and of course we know $V(I(P)) = P$. □

HNS establishes a hands-on way of decomposing ANY hypersurface into its irreducible components, which generalizes the result already obtained in the two variable case.

Corollary 4.2.25. *Consider a hypersurface* $X := V(f) \subset \mathbb{A}^n$. *We can write*

$$f = \prod_1^r f_i^{v_i},$$

for suitable distinct irreducible polynomials $f_1,\dots,f_r \in K[x_1,\dots,x_n]$ *and* $v_1,\dots,v_r \geq 1$. *The irreducible components of* X *are* $V_1,\dots,V_r$, *where* $V_i := V(f_i)$. *We thus establish a one-to-one correspondence between irreducible hypersurfaces and irreducible polynomials up to scalar multiplication from* K. *Lastly* $I(X) = \left\langle \prod_1^r f_i \right\rangle$

Proof. Since the ideal $\langle f_i \rangle$ are prime, it follows from HNS that $V_i = V(f_i)$ is irreducible. Since $f_i \nmid f_j$ for $i \neq j$, it follows that $V_i \not\subset V_j$, hence $V_1, \dots, V_r$ are the irreducible components of V. To prove the last statement, it is sufficient, by HNS, to prove that $\mathbf{rad}(\langle f \rangle) = \langle \prod_1^r f_i \rangle$. This follows from Lemma 2.8.51. □

Corollary 4.2.26. *Consider* $I := \langle f_1, \dots, f_m \rangle \subset K[x_1, \dots, x_n]$. $V(I)$ *is finite if and only if* $K[\mathbf{x}]/I$ *is a finite dimensional vector space over* K. *If this occurs* $\#V(f) \leq \mathbf{dim}_K K[\mathbf{x}/I]$.

Proof. "$\Rightarrow$": Suppose $V(I) = \{p_1, \dots, p_l\}$, where $p_i = (a_{i1}, \dots, a_{il}) \in \mathbb{A}^n$. Define for $j \in \{1, \dots, n\}$, $g_j := \prod_{i=1}^l (x_j - a_{ij})$. Then $g_j \in I(V(I)) = \mathbf{rad}(I)$, hence for some $N_j \geq 1$, $g_j^{N_j} \in I$. Take $N = \mathbf{max}\ N_j$. Then for suitable $\lambda_{ij} \in K$,

$$0 + I = F_j^N + I = \left[X_j^{lN} + \sum_0^{lN-1} \lambda_{ij} X_j^i \right] + I \Rightarrow X_j^{lN} + I = \left[- \sum_0^{lN-1} \lambda_{ij} X_j^i \right] + I.$$

For each $s \geq 0$, we thus find that $X_j^s + I$ is a linear combination of $1 + I, X_j + I, \dots, X_j^{rN-1} + I$. This means

$$K[\mathbf{x}]/I = \mathrm{Span}_K \left(\left\{ x_j^k + I : j \in \{1, \dots, n\}, k \in \{0, \dots, lN\} \right\} \right),$$

hence $K[\mathbf{x}]/I$ is finite dimensional.

"$\Leftarrow$": Set $d := \mathbf{dim}_K\ K[\mathbf{x}]/I$ and let distinct points $p_1, \dots, p_l \in V(I)$ be given. There are polynomials $g_1, \dots, g_l \in K[\mathbf{x}]$ such that $g_i(p_i) = 1$ and $g_i(p_j) = 0$ for $i \neq j$ (which exist by Some result. Insert ref!). Let $\lambda_1, \dots, \lambda_l \in K$ be given such that

$$\left[\sum_1^l \lambda_i g_i \right] + I = 0 + I.$$

Then $\left[\sum_1^l \lambda_i g_i \right] \in I$, so for each i,

$$\lambda_i = \lambda_i g_i(p_i) = \sum_1^m \lambda_j f_j(p_i) = 0,$$

hence $\{g_1 + I, \dots, g_l + I\} \subset K[\mathbf{x}]/I$ are linearly independent over K. Then $\sum_1^l K(g_i + I) \subset K[\mathbf{x}]/I$ has dimension l, and $l \leq d$. This means there can be at most d distinct points in $V(I)$. □

Example 4.2.27. 1. In general the Nullstellensatz provides a way to find irreducible components. Consider for example $f := x^2 + y^2 - 1, g := x^2 - z^2 - 1 \in \mathbb{C}[x, y, z]$ and set $X := V(f, g)$. The fact that f, g are both irreducible will make the task of providing the irreducible components for X difficult. It is therefor useful to find an alternative generating set for $I := \langle f, g\rangle$. Note that $h := y^2 + z^2 = f - g \in I$ and $f = (f - g) + g = h + g \in J := \langle h, g\rangle$. It thus follows that $I = J$. Setting $h_1 := y - iz$ and $h_2 := y + iz$, one sees that $h = h_1 h_2$. This means

$$X = V(I) = V(J) = (V(h_1) \cup V(h_2)) \cap V(g) = \underbrace{(V(h_1) \cap V(g))}_{V_1} \cup \underbrace{(V(h_2) \cap V(g))}_{V_2}.$$

We now claim that V_1, V_2 are the irreducible components of X. Consider the ring homomorphism

$$\sigma : \mathbb{C}[x, y, z] \to \mathbb{C}[x, z]/\langle g\rangle$$
$$f \mapsto f(x, iz, z)$$

This is clearly a surjective ring homomorphism. Note that $\sigma(h_1) = 0$ and $\sigma(g) = 0$, hence $\langle h_1, f\rangle \subset \textbf{ker}\ \sigma$. Let $f \in \textbf{ker}\ \sigma$. Then $f(x, iz, z) + \langle g\rangle = 0$. This means

$$f + \langle h_1, g\rangle = f(x, iz, z) + \langle h_1, g\rangle = 0 + \langle h_1, g\rangle,$$

hence $\textbf{ker}\ \sigma = \langle h_1, g\rangle$, which shows that

$$\mathbb{C}[x, y, z]/\langle h_1, g\rangle \simeq \mathbb{C}[x, z]/\langle g\rangle.$$

Since g is irreducible, it is prime and thus $\mathbb{C}[x, z]/\langle g\rangle$ is an integral domain hence $\mathbb{C}[x, y, z]/\langle h_1, g\rangle$ is an integral, meaning $\langle h_1, g\rangle$ is a prime ideal. By HNS, it follows that V_1 is irreducible. Similarly one can show that $\mathbb{C}[x, y, z]/\langle h_2, g\rangle$ is an integral domain, and hence $\langle h_2, g\rangle$ is prime, meaning V_2 is irreducible. Any point vanishing on h_1 clearly does not vanish on h_2, hence $V_i \not\subset V_j$ for $i \neq j$. We thus conclude that V_1, V_2 are the irreducible components of X.

2. Set $X := \{(v, v^2, v^3) : v \in \mathbb{C}^3\}$. Note that by Example 4.1.17, $X = V(f, g)$, where $f_1 = x^2 - y$ and $f_2 = x^3 - z$. How does one determine $I(X)$?. Consider the surjective ring homomorphism

$$\sigma : \mathbb{C}[x, y, z] \to \mathbb{C}[t]$$
$$f \mapsto f\left(t, t^2, t^3\right)$$

Note that $I(V) = \textbf{ker}\ \sigma$. Indeed, $f \in \mathbb{C}[x, y, z]$ vanishes on $\left(v, v^2, v^3\right)$, then $f\left(t, t^2, t^3\right) = 0$, since $\mathbb{C}$ is infinite. Conversely if $f\left(t, t^2, t^3\right) = 0$, then for any $v \in \mathbb{C}$, $f\left(v, v^2, v^3\right) = 0$. Since $V(I(V)) = V$, it would be nice to compute the generators of $I(V)$, as this perhaps leads to a nicer sets of generators, that allows us to say something about irreducible components. Note that $\mathbf{ker}\ \sigma = \mathbb{C}[x,y,z] \cap \langle x-t, y-t^2, z-t^3 \rangle$. This is indeed just a reformulation of Lemma 2.9.38. In general, how do we compute the generators of $\mathbb{C}[x,y,z] \cap \langle x-t, y-t^2, z-t^3 \rangle$? First let $\leq$ be the lexicographic term order where $x < y < z < t$ or $x < y < z$. The answer is that we turn Gröbner basis theory to compute a Gröbner basis for the ideal $I := \langle x-t, y-t^2, z-t^3 \rangle \subset \mathbb{C}[x,y,z,t]$ with respect to $\leq$ using Buchberger's algorithm and then determine $G \cap K[x,y,z]$ which will be a Gröbner basis for $K[x,y,z] \cap I$ with respect to $\leq$ by Theorem 2.9.91. Set $g_1 := -t+x, g_2 := -t^2+y, g_3 := -t^3+z$. We follow Buchberger's algorithm

$$S(g_1,g_2) = tx - y,$$

which has residue $f_1 = -y+x$ when divided by $\{g_1,g_2,g_3\}$. For step 2 we find that

$$S(g_1,g_3) = t^2x - z,$$

which has residue $f_2 = -z+x^3$ when divided by $\{g_1,g_2,g_3,f_1\}$. For step 3 we find that

$$S(g_2,g_3) = ty - z,$$

which one verifies to have residue 0 when divided by $\{g_1,g_2,g_3,f_1,f_2\}$. Note now that $S(g_i,g_j) \to_{\{g_1,g_2,g_3,f_1,f_2\}} 0$ and the initial terms of the remaining pairs have pairwise greatest common divisor 1, hence the S-polynomial of these pairs also reduce to 0 modulo $\{g_1,g_2,g_3,f_1,f_2\}$ by Proposition 2.9.85. Thus $G := \{g_1,g_2,g_3,f_1,f_2\}$ is a Gröbner basis for I, hence $G' := G \cap \mathbb{C}[x,y,z] = \{f_1,f_2\}$ is a Gröbner basis for $\mathbf{ker}\ \sigma = I(V)$. From this one concludes that $\langle x-y^2, x-z^3 \rangle$ is radical. However, we can say even more! We actually find that

$$K[x,y,z]/I(V) = K[x,y,z]/\langle f_1,f_2 \rangle \simeq K[t].$$

This means $K[x,y,z]/I(V)$ is an integral domain, implying V is irreducible by Proposition 4.2.5.

Example 4.2.28. Consider $f = y^2 - x(x-1)(x-a) \in K[x,y]$ where K is an algebraically closed field and $a \in K$. $V(f)$ is irreducible. Indeed, when $a \neq 1$, $x-1 \nmid y^2$, $x \mid f$ and $(x-1)^2 \nmid f$ thus by Eisenstein's criterion f is irreducible/prime. When $a = 1$, $a \neq 0$, hence $x \nmid y^2, x \mid f$ and $x^2 \nmid f$ implies again by Eisenstein's criterion that f is irreducible/prime. The ideal $\langle f \rangle$ is thus prime, hence by HNS $V(f)$ is irreducible.

Example 4.2.29. Consider $f_1 = x^2 - y^2 = (x+y)(x-y) \in \mathbb{C}[x,y]$ and $f_2 = x^2 + y^2 = (x+iy)(x-iy) \in \mathbb{C}[x,y]$. Set $V := V(f_1, f_2)$. Let's find the irreducible components. There is a naive approach: Note that

$$\begin{aligned} V &= (V(x+y) \cup V(x-y)) \cap (V(x+iy) \cup V(x-iy)) \\ &= V(x+y, x+iy) \cup V(x+y, x-iy) \cup V(x-y, x+iy) \cup V(x-y, x-iy). \end{aligned}$$

We can thus decompose V into 4 linear equations all of which has $\{(0,0)\}$ as there only solution, hence $V = \{(0,0)\}$.
One can be more smarter: Note that $1/2(f_1 + f_2) = x^2$ and $1/2(f_2 - f_1) = y^2$, hence $I = J := \langle x^2, y^2 \rangle$, hence clearly if $v \in V(I)$, $v = (0,0)$. Note now that

$$\mathbb{C}[x,y]/I = \mathbb{C}[x,y]/J = \{ax + by + cxy + d + J : a,b,c,d \in \mathbb{C}\},$$

and hence is isomorphic to the 4 dimensional vector space $\mathbf{Span}_{\mathbb{C}}(1, x, y, xy) \subset \mathbb{C}[x,y]$. So the inequality from Corollary 4.2.26 doesn't always hold with equality. One notes, however that when decomposing V into a system of linear equations, we get 4 "copies" of the point $(0,0)$, hence it seems that the dimension of the vector space $\mathbb{C}[x,y]/I$ matches the number of points in $V(I)$ with some sort of multiplicity.

Corollary 4.2.30. *If a variety $V \subset \mathbb{A}^n$ is a hypersurface. It is either $\mathbb{A}^n$, $\emptyset$ or $V(f)$ where f is irreducible. As a consequence, such a hypersurface is of the third type, it contains no variety W.*

Proof. This follows from Lemma 2.8.52 and HNS. □

Definition 4.2.31. Let $V \subset \mathbb{A}^n$ be a variety. A variety $W \subset \mathbb{A}^n$ is a *subvariety of* V if it is also a subset of V.

Proposition 4.2.32. *Let $V \subset \mathbb{A}^n$ be a variety. We have one-to-one correspondences between*

1. *Algebraic subsets contained in V and radical ideals of $\Gamma(V)$.*

2. *Subvarieties of V and prime ideals of $\Gamma(V)$.*

3. *Points of V and maximal ideals of $\Gamma(V)$.*

Proof. This follows from HNS (or more precisely the one-to-one correspondences obtained from it) and Proposition 2.8.32. □

Definition 4.2.33. Let $V \subset \mathbb{A}^n$ be a variety and $W \subset V$ a subvariety. We define $I_V(W)$ to be the ideal $I(W)/I(V) \subset \Gamma(V)$.

Lemma 4.2.34. *Let $V \subset \mathbb{A}^n$ be a variety and $W \subset V$ a subvariety. Then* $\Gamma(W) \simeq \Gamma(V)/I_V(W)$

Proof. This follows from Is written since

$$\Gamma(V)/I_V(W) = \frac{K[\mathbf{x}]/I(V)}{I(W)/I(V)} \simeq K[\mathbf{x}]/I(W) = \Gamma(W).$$

□

Proposition 4.2.35. *Let $V \subset \mathbb{A}^n$ be a non-empty variety. The following are equivalent:*

1. *V is a point.*

2. *$\Gamma(V) = K$.*

3. *$\dim_K \Gamma(V) < \infty$.*

Proof. "1.⇒ 2.": By HNS, $\Gamma(V)$ is a field, and $K[x] \to \Gamma(V), f \mapsto f + I(V)$ is surjective K-algebra homomorphism hence by Corollary 2.10.27, $\Gamma(V) = K$.
"2.⇒ 3.": In this case $\dim_K \Gamma(V) = \dim_K K = 1 < \infty$.
"3.⇒ 1.": Since $\dim_K \Gamma(V) < \infty$, V is finite by Corollary 4.2.26, and since V it cannot consist of multiple points. □

4.2.3 Introduction to Effective Nullstellensätze: Degree Bounds and a Gröbner Basis Method

We fix an algebraically closed field K.

Remark 4.2.36. An *effective Nullstellensatz (ENS)* is a theorem that proves the weak Nullstellensatz in a way that gives rise to an algorithm for computing a solution to Hilbert Equations.

The proof of the weak Nullstellensatz provided in these notes thus far does not give a way for constructing an explicit solution to the Hilbert Equation associated with polynomials $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$, which vanish nowhere simultaneously and is therefor not an ENS. We are thus presented with the question of how to construct a solution to this Hilbert equation explicitly.
One way to accomplish this, is to prove that there is an upper bound $B \geq 0$ (which will typically dependent on $\deg f_1, \ldots, \deg f_m$) on $\{\deg \lambda_i\}_1^m$ or equivalently on $\{\deg \lambda_i f_i\}_1^m$.
To see that this is indeed the case, write

$$f_i = \sum_{v \in \mathbb{N}^n} a_v^{(i)} \mathbf{x}^v \in K[\mathbf{x}],$$

for suitable $a_v^{(i)} \in K$, $i \in \{1,\dots,m\}$. Define for $u \in \mathbb{N}^n$ with $|u| \le B$,

$$c_u = \sum_{\substack{i \in \{1,\dots,m\},\\ v,w \in \mathbb{N}^n:\\ v+w=u}} a_v^{(i)} y_w^{(i)},$$

where $y_w^{(i)}$ (say for $|w| \le B$) are unknowns. Consider the equation:

$$1 = \sum_1^m \sum_{v \in \mathbb{N}^n} \sum_{w \in \mathbb{N}^n} a_v^{(i)} y_w^{(i)} \mathbf{x}^{v+w} = \sum_{u \in \mathbb{N}^n : |u| \le B} c_u \mathbf{x}^u. \tag{8}$$

This equation has a solution if and only if the following linear equation has a solution,

$$\begin{cases} c_u = 0 & \text{for } 0 < |u| \le B, \\ c_{(0,\dots,0)} = 1. \end{cases} \tag{9}$$

The upshot is that the existence of a solution $b_w^{(i)} \in K$ to (8) is seen to be equivalent the existence of a solution to the Hilbert equation bounded by B when putting

$$\lambda_i = \sum_{w \in \mathbb{N}^n} b_w^{(i)} \mathbf{x}^v.$$

Indeed we get that

$$\sum_1^m \lambda_i f_i = \sum_{u \in \mathbb{N}^n : |u| \le B} c_u \mathbf{x}^u = 1$$

and by construction $\mathbf{deg}\ \lambda_i f_i \le B$. Conversely if a solution to the Hilbert equation $\lambda_i \in K[x_1,\dots,x_n]$ with $\mathbf{deg}\ \lambda_i f_i \le B$ exists then the coefficients of λ_i constitutes a solution to (8). The upshot is that we can produce an explicit solution to the Hilbert equation by solving the system of linear equations (9) and we know that a solution to this system of equations exists exactly when a degree bounded solution to the Hilbert equation exists.

Another approach to providing an ENS relies exclusively on Gröbner basis theory. We present this approach here. We first need to reformulate WNS such that the statement can be "observed" by a Gröbner basis. To do this we give the following definition

Definition 4.2.37. Consider polynomials $f_1,\dots,f_m \in K[x_1,\dots,x_n]$. A polynomial $p(\mathbf{x}, y_1 \dots, y_m) \in K[\mathbf{x}, y_1,\dots,y_m]$ is called a *final polynomial for* $f_1,\dots,f_m$ if

1. $p(\mathbf{x}, f_1,\dots,f_m) = 0$
2. $p(\mathbf{x}, \mathbf{0}) \in K \setminus \{0\}$.

Furthermore, if $p \in K[\mathbf{y}] \subset K[\mathbf{x},\mathbf{y}]$ is a final polynomial for $f_1,\dots,f_m$, it is called a *final syzygy* (for $f_1,\dots,f_m$).

Lemma 4.2.38. *Let $f_1,\ldots,f_m \in K[x_1,\ldots,x_n]$. Then there is a final polynomial for $f_1,\ldots,f_m$ if and only if there is a solution to the Hilbert equation for $f_1,\ldots,f_m$.*

Proof. "$\Rightarrow$": Suppose $p \in K[\mathbf{x}, y_1,\ldots,y_m]$ is a final polynomial for $f_1,\ldots,f_m$. We can then write

$$p = -c + \sum_1^m G_i y_i,$$

for suitable $G_i \in K[\mathbf{x},\mathbf{y}]$ and $c \in K$. The second defining property of final polynomials shows that $c \neq 0$. Put $\lambda_i = G_i(\mathbf{x}, f_1,\ldots,f_m) \in K[\mathbf{x}]$ for each i. From the first defining property for final polynomials it follows that

$$c = p(\mathbf{x}, f_1,\ldots,f_m) = \sum_1^m G_i(\mathbf{x}, f_1,\ldots,f_m) f_i = \sum_1^m \lambda_i f_i,$$

hence (after scaling by c^{-1}), $\{\lambda_i\}_1^m$ is a solution to the Hilbert equation for $f_1,\ldots,f_m$. "$\Leftarrow$": Let $\lambda_1,\ldots,\lambda_m \in K[\mathbf{x}]$ be a solution to the Hilbert equation for $f_1,\ldots,f_m$. Put $p = 1 - \sum_1^m \lambda_i y_i \in K[\mathbf{x},\mathbf{y}]$. Then

$$p(\mathbf{x}, f_1,\ldots,f_m) = 1 - \sum_1^m \lambda_i f_i = 1 - 1 = 0,$$

hence p satisfies the first defining property for final polynomials. Secondly, we have that

$$p(\mathbf{x}, \mathbf{0}) = 1 - \sum_1^m \lambda_i \cdot 0 = 1 \in K \setminus 0,$$

hence p is a final polynomial for $f_1,\ldots,f_m$. □

This gives rise to the following reformulation of the weak Nullstellensatz

Theorem 4.2.39. *Let K be algebraically closed and consider $f_1,\ldots,f_m \in K[x_1,\ldots,x_n]$. Then $V(f_1,\ldots,f_m) = \emptyset$ if and only if there exists a final polynomial for $f_1,\ldots,f_m$.*

Proposition 4.2.40. *Let $f_1,\ldots,f_m \in K[x_1,\ldots,x_n$, denote the graph ideal of these polynomials by I and let G be a Gröbner basis for I with respect $\leq_{lex}$ with $x_1 > \cdots > x_n > y_1 > \cdots > y_m$. Then $f_1,\ldots,f_m$ has a final syzygy if and only if G contains a final syzygy for $f_1,\ldots,f_m$.*

Proof. Note by Theorem 2.9.91 that $G' = G \cap K[\mathbf{y}] \subset G$ is a Gröbner basis for the ideal $I \cap K[\mathbf{y}]$ with respect to the $\leq_{\text{lex}}$ with $y_1 < \ldots < y_m$. Suppose $p \in K[\mathbf{y}]$ is a final syzygy for $f_1,\ldots,f_m$. We see that $p \in I \cap K[y]$ by Proposition 2.9.38, which implies $\widehat{G} \cap K[\mathbf{y}]$ contains a non-zero polynomial. The final syzygy p is on the form $-\left[\sum_1^m h_i y_i\right] + c$ for suitable $h_i \in K[\mathbf{y}]$, $c \in K \setminus 0$ and $p^{G'} = 0$ by Proposition 2.9.69.

For this to be true there must be a polynomial $p' \in G'$ with non-zero constant term, implying $p'(0) \in K \setminus 0$. Since $p' \in I$, it satisfies $p'(f_1,\dots,f_m) = 0$. One thus concludes that p' is a final syzygy for $f_1,\dots,f_m$. $\square$

To extend this approach to the case where only a final polynomial to exist we need to modify the definition of final polynomial slightly.

Definition 4.2.41. Consider $f_1,\dots,f_m \in K[x_1,\dots,x_n]$. A polynomial $p(z,\mathbf{x},\mathbf{y}) \in K[z,\mathbf{x},y_1,\dots,y_n]$ is called an *extended final polynomial* if

$$\begin{aligned} &1. \quad p(z,\mathbf{x},zf_1(\mathbf{x}),\dots,zf_m(\mathbf{x})) = 0 \\ &2. \quad p(z,\mathbf{x},0) = cz \quad (c \in K \setminus \{0\}). \end{aligned}$$

Remark 4.2.42. It is easy to see that an extended final polynomial $p(z,\mathbf{x},\mathbf{y})$ for $f_1,\dots,f_m \in K[\mathbf{x}]$ gives rise to a final polynomial for $f_1,\dots,f_m$ given by $p(1,\mathbf{x},\mathbf{y})$. Conversely, if $f_1,\dots,f_m$ admits a final polynomial $p(\mathbf{x},\mathbf{y}) = 1-\sum_1^m \lambda_i y_i \in K[\mathbf{x},\mathbf{y}]$ where $\lambda_i \in K[\mathbf{x}]$, we can construct an extended final polynomial for $f_1,\dots,f_m$ given by $z - \sum_1^m \lambda_i y_i$.

Theorem 4.2.43. *Let $f_1,\dots,f_m \in K[\mathbf{x}]$ and let $I = \langle zf_1-y_1,\dots,zf_m-y_m\rangle \subset K[z,\mathbf{x},y_1,\dots,y_m]$. Let $\leq$ be a term order on $K[z,\mathbf{x},\mathbf{y}]$, where z^k is greater than any monomial in $K[\mathbf{x},\mathbf{y}]$ for $k \geq 1$ (e.g. $\leq_{lex}$ with $z > x_1 > \cdots > x_n > y_1 > \cdots > y_m$). Consider a Gröbner basis $G \subset K[z,\mathbf{x},\mathbf{y}]$ of I with respect to $\leq$. Then $V(\mathscr{F}) = \emptyset$ if and only if G contains an extended final polynomial for $f_1,\dots,f_m$.*

Proof. The first direction is trivial. For the converse statement, suppose $V(\mathscr{F}) = \emptyset$. Then by the Nullstellensatz there exist $\lambda_i \in K[\mathbf{x}]$ such that

$$1 = \sum_1^m \lambda_i f_i,$$

hence

$$p(z,\mathbf{x},\mathbf{y}) = z - \sum_1^m \lambda_i y_i,$$

defines an extended final polynomial for $f_1,\dots,f_m$. By Lemma 2.10.57, $p \in I$. Then, since G is a Gröbner basis, there is a $g \in G$ such that $\mathbf{in}_\leq\, g \mid \mathbf{in}_\leq\, p = z$. Again by Lemma 2.10.57, $1 \notin I$, hence we can assume that $\mathbf{in}_\leq\, g = cz$ for some $c \in K \setminus \{0\}$. Since $g \in I$, we have that

$$g = \sum_1^m \mu_i(z,\mathbf{x},\mathbf{y})(zf_i - y_i),$$

for suitable $\mu_i \in K[z,\mathbf{x},\mathbf{y}]$. This, for one, implies that $g(z,\mathbf{x},zf_1,\dots,zf_m) = 0$. Furthermore, we can rearrange terms in the above expression to write

$$g = \underbrace{z\sum_1^m \mu_i' f_i}_{s_1} - \underbrace{\sum_1^m \mu_i'' y_i}_{s_2},$$

for suitable $\mu_i' \in K[z,\mathbf{x}]$ and $\mu_i'' \in K[z,\mathbf{x},\mathbf{y}]$. Since $s_2 \in \langle y_1,\dots,y_m\rangle$ and $s_1 \in K[z,\mathbf{x},\mathbf{y}] \setminus \langle y_1,\dots,y_m\rangle$ there is no cancellation between terms of s_1 and s_2. This means $\sum_1^m \mu_i' f_i = c$ and hence $g(z,\mathbf{x},0) = cz$, which implies g is an extended final polynomial for $f_1,\dots,f_m$. □

4.3 A Theory of Affine Varieties

In this section we will assume that any field is algebraically closed ref to def

4.3.1 Morphisms of Affine Varieties: Polynomial maps

Definition 4.3.1. Let $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ be affine varieties. A map $\varphi : V \to W$ is called a *polynomial map* if there are polynomials $f_1,\dots,f_m \in K[x_1,\dots,x_n]$ such that for $v \in V$,

$$\varphi(v) = (f_1(v),\dots,f_m(v)).$$

When $W = K$, we call such a polynomial map a *polynomial function.* We denote the set of polynomial maps from V to W by $\mathrm{Pol}(V,W)$

Remark 4.3.2. 1. Note that $\mathrm{Pol}(V,K)$ admits a structure of ring where $0,1 \in \mathrm{Pol}(V,K)$ are the constant functions mapping every point to respectively 0 and 1 and where for $f,g \in \mathrm{Pol}(V,K)$, $f+g \in \mathrm{Pol}(V,K)$ is defined by

$$(f+g)(v) := f(v) + g(v) \quad (v \in V)$$

and $fg \in \mathrm{Pol}(V,K)$ is defined by

$$(fg)(v) = f(v)g(v) \quad (v \in V).$$

2. Suppose we have polynomials $f_1,\dots,f_m,g_1,\dots,g_m \in K[x_1,\dots,x_n]$ such that for $v \in V$

$$(f_1(v),\dots,f_m(v)) = \varphi(v) = (g_1(v),\dots,g_m(v)),$$

then $(f_i - g_i)(v) = f_i(v) - g_i(v) = 0$ for every $v \in V$ and $i \in \{1,\dots,m\}$, i.e. $f_i - g_i \in I(V)$. Hence a polynomial map is uniquely defined up to $I(V)$-residues.

Example 4.3.3. Here are a few examples of polynomial maps.

1. Consider for $1 \le i \le j \le n$ the map

$$\pi_{i,j} : \mathbb{A}^n \to \mathbb{A}^m$$
$$(v_1,\dots,v_n) \mapsto (v_i,\dots,v_j)$$

   Note that setting $f_i = x_i \in K[x_1,\dots,x_n]$ we get that

$$(f_i(v),\dots,f_j(v)) = (v_i,\dots,v_j) = \pi_{i,j}(v),$$

   for every $v \in \mathbb{A}^n$, hence $\pi_{i,j}$ is polynomial. In particular the projection map $\pi_i = \pi_{i,i}$ is polynomial.

2. Let $\omega \in \mathcal{S}_n$, i.e. a permutation of n elements. Then

$$\varphi : \mathbb{A}^n \to \mathbb{A}^n$$
$$(v_1,\dots,v_n) \mapsto (v_{\omega(1)},\dots,v_{\omega(n)})$$

   Setting $f_i = x_{\omega(i)}$ for each i, we get that

$$(f_1(v),\dots,f_n(v)) = (v_{\omega(1)},\dots,v_{\omega(n)}) = \varphi(v),$$

   for every $v \in \mathbb{A}^n$.

3. Given a variety $V \subset \mathbb{A}^n$. The identity map is a polynomial map, given by $x_1,\dots,x_n \in K[x_1,\dots,x_n]$.

4. A simple example is given varieties $V \subset \mathbb{A}^n$, $W \subset \mathbb{A}^m$ and $f_1,\dots,f_m \in K[x_1,\dots,x_n]$, where $f_1(v),\dots,f_m(v) \in W$. $V \ni v \mapsto (f_1(v),\dots,f_m(v)) \in W$ is a polynomial map.

Lemma 4.3.4. *Let $V \subset \mathbb{A}^n$, $W \subset \mathbb{A}^m$ be varieties and $\varphi : V \to W$ a polynomial map defined by $f_1 \dots, f_m$. Let $X = V(g_1,\dots,g_l) \subset \mathbb{A}^m$ be an algebraic set, where $g_1,\dots,g_l \in K[y_1,\dots,y_m]$. We then have that*

1. *$\varphi^{-1}(X) \subset V$ is an algebraic set.*

2. *If $\varphi^{-1}(X)$ is a variety and $X \subset Im\ \varphi$, then X is irreducible.*

Proof. 1. We prove that $\varphi^{-1}(X) = V(g_1(f_1,\dots,f_m),\dots,g_l(f_1,\dots,f_m))$. Indeed,

$$\begin{aligned} v \in \varphi^{-1}(X) &\iff (f_1(v),\dots,f_m(v)) = \varphi(v) \in X \in X = V(g_1,\dots,g_l) \\ &\iff g_i(f_1,\dots,f_m)(v) = g_i(f_1(v),\dots,f_m(v)) = 0 \ \forall i \\ &\iff v \in V(g_1(f_1,\dots,f_m),\dots,g_l(f_1,\dots,f_m)) \end{aligned}$$

2. Suppose $X = Y \cup Z$ for some algebraic sets $Y, Z \subset X$. Then $\varphi^{-1}(Y), \varphi^{-1}(Z)$ are algebraic sets contained in V such that

$$\varphi^{-1}(X) = \varphi^{-1}(Y \cup Z) = \varphi^{-1}(Y) \cup \varphi^{-1}(Z),$$

hence WLOG $\varphi^{-1}(Y) = \varphi^{-1}(X)$. Then since $X \subset \varphi(V)$,

$$Y = \varphi(\varphi^{-1}(Y)) = \varphi(\varphi^{-1}(X)) = X.$$

□

The above result can be used to determine wether an algebraic set is a variety.

Example 4.3.5. 1. Consider the algebraic set $V = \{(t, t^2, t^3) : t \in \mathbb{C}\} \subset \mathbb{A}^3(\mathbb{C})$ from Example 4.2.27. There is an easier way of showing that it is a variety. Indeed consider the surjective polynomial map

$$\begin{aligned} \varphi : \mathbb{A}^1(\mathbb{C}) &\to V \\ t &\mapsto (t, t^2, t^3) \end{aligned}$$

This has an inverse given by projection onto the first coordinate, i.e. the inverse is the map

$$\begin{aligned} \varphi^{-1} : V &\to \mathbb{A}^1(\mathbb{C}) \\ (t, t^2, t^3) &\mapsto t \end{aligned}$$

This mean $\varphi^{-1}(V) = \mathbb{A}^1(\mathbb{C})$ which is a variety, hence by part 2. of the above proposition V is a variety.

2. Consider $I = \langle f_1, f_2, f_3 \rangle \subset \mathbb{C}[x, y, z]$, where $f_1 = xz - y^2$, $f_2 = yz - x^3$, $f_3 = z^2 - x^2 y$ and set $V = V(I)$. Consider

$$\begin{aligned} \varphi : \mathbb{A}(\mathbb{C})^1 &\to V \\ t &\mapsto (t^3, t^4, t^5) \end{aligned}$$

One easily sees that $f_i(t^3, t^4, t^5) = 0$ for $i = 1, 2, 3$. Let $v \in V$. And let t be any solution to the equation $X^3 - v_1 = 0$. If $v_1 = 0$, then $0 = v_1 v_3 - v_2^2 = -v_2^2$, hence $v_2 = 0$. Furthermore $v_3 = v_1^2 v_2 = 0$. In this case, we then have $\varphi(0) = v$. Suppose $v_1 \neq 0$. Then $v_1 v_3 = v_2^2$, hence $v_3 = v_2^2 / v_1$, hence $v_2 v_2^2 / v_1 = v_2 v_3 = v_1^3$, implying $v_2^3 = v_1^4 = t^{12}$, hence t is a solution to $X^4 - v_2 = 0$. Lastly $v_3^2 = v_1^2 v_2 = t^6 t^4 = t^1 0$, hence t is a solution to $X^5 - v_3 = 0$. It thus follows that $\varphi(t) = v$. Note that $\varphi^{-1}(V) = V(f_1(t^3, t^4, t^5), f_2(t^3, t^4, t^5), f_3(t^3, t^4, t^5)) = V(0) = \mathbb{A}^1(\mathbb{C})$, which is a variety. Again by part 2. of the above theorem V is a variety.

Proposition 4.3.6. *Affine varieties with polynomial maps chosen as morphisms define a category.*

Proof. It is clear that composition of polynomial maps is associative and the identity function $\mathrm{id}_V : \mathbb{A}^n \supset V \ni v \mapsto v \in V$ is the identity morphism, since this map is polynomial (defined by $x_1,\dots,x_n \in K[x_1,\dots,x_n]$). □

Proposition 4.3.7. *A polynomial map is continuous in the Zariski topology.*

Proof. Let varieties $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ be given. Consider a polynomial map $\varphi : V \to W$ defined by $f_1,\dots,f_m \in K[x_1,\dots,x_n]$. Consider a Zariski closed subset $U \subset W$. Then by Proposition 4.3.4 1. $\varphi^{-1}(U)$ is Zariski closed, hence it is know from topology ADD SOME TOPOLOGY EVENTUALLY that φ is continuous. □

Definition 4.3.8. A polynomial map $\varphi : V \to W$, where $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ are affine varieties is called an *isomorphism (of affine varieties)* if it is bijective and $\varphi^{-1} : W \to V$ is a polynomial map. In other words an isomorphism of affine varieties is a *bi-polynomial map*.

Proposition 4.3.9. *Let $V \subset \mathbb{A}^n$ be an affine variety. Then*

$$\Gamma(V) \simeq Pol(V,K).$$

Proof. Consider the ring homomorphism

$$\sigma : K[x_1,\dots,x_n] \to \mathrm{Pol}(V,K)$$
$$f \mapsto (V \ni v \mapsto f(v) \in K),$$

This is obviously surjective. Indeed, a polynomial function $\mathrm{Pol}(V,K)$ is given by $v \mapsto f(v)$ for some $f \in K[x_1,\dots,x_n]$ by definition. If $f \in K[x_1,\dots,x_n]$, is given such that $f(v) = \sigma(f)(v) = 0$ for all $v \in V$, then $f \in I(V)$. Hence one sees that $I(V) = \ker\ \sigma$. Then by Theorem 2.6.19,

$$\Gamma(V) = K[\mathbf{x}]/I(V) = K[\mathbf{x}]/\ker\ \sigma \simeq \mathrm{Pol}(V,K).$$

□

Definition 4.3.10. Let $V \subset \mathbb{A}^n, W \subset \mathbb{A}^n$ be varieties and $\varphi : V \to W$ be a polynomial map defined by polynomials $f_1,\dots,f_m \in K[x_1,\dots,x_n]$. The *$K$-algebra homomorphism induced by φ* is the map

$$\widetilde{\varphi} : \Gamma(W) \to \Gamma(V)$$
$$f + I(W) \mapsto f \circ \varphi + I(V)$$

where $f \circ \varphi + I(V) = f(f_1,\dots,f_m) + I(V)$

Remark 4.3.11. This map is well-defined. Indeed the map described above is due to Proposition 4.3.9 given by the composition

$$\Gamma(W) \xrightarrow{\sim} \mathrm{Pol}(W,K) \to \mathrm{Pol}(V,K) \xrightarrow{\sim} \Gamma(V)$$
$$f + I(W) \mapsto \mathrm{ev}_\bullet(f) \mapsto \mathrm{ev}_\bullet(f) \circ \varphi \mapsto f(f_1,\dots,f_m) + I(V).$$

Proposition 4.3.12. *Let $V \subset \mathbb{A}^n, W \subset \mathbb{A}^n$ be varieties and $\varphi : V \to W$ be a polynomial map defined by $f_1,\dots,f_m \in K[x_1,\dots,x_n]$. $\widetilde{\varphi} \in \mathrm{Hom}^{K-Alg}(\Gamma(W),\Gamma(V))$.*

Proof. The map

$$\sigma : K[y_1,\dots,y_m] \to \Gamma(V)$$
$$f \mapsto f \circ \varphi + I(V)$$

is clearly a ring homomorphism. Suppose $f \in I(W)$. Then for every $v \in V$, $\varphi(v) \in W$. Hence $(f \circ \varphi)(v) = f(\varphi(v)) = 0$, hence $\sigma(f) = f \circ \varphi \in I(V)$, hence by Corollary 2.6.12, $\widetilde{\varphi}$ is a well-defined ring homomorphism. Let $k \in K$. Then

$$\widetilde{\varphi}(k + I(W)) = k \circ \varphi + I(V) = k + I(V) = k(1 + I(V)) = k\widetilde{\varphi}(1 + I(W))$$

□

Lemma 4.3.13. *The mapping taking a variety V to $\Gamma(V)$ and a polynomial map $\varphi \in \mathrm{Pol}(V,W)$ to $\widetilde{\varphi} \in \mathrm{Hom}^{K-Alg}(\Gamma(W),\Gamma(V))$ is a contravariant functor.*

Proof. For varieties $V \subset \mathbb{A}^n$, $W \subset \mathbb{A}^m$, $U \subset \mathbb{A}^l$, let $\varphi \in \mathrm{Pol}(V,W)$ and $\psi \in \mathrm{Pol}(W,U)$ be defined by polynomials $f_1,\dots,f_m \in K[x_1,\dots,x_n]$ and $g_1,\dots,g_l \in K[y_1,\dots,y_m]$. Then viewing an arbitrary $f + I(W) \in \Gamma(W)$ as an element $f \in \Gamma(W,K)$, we have that

$$\widetilde{\psi\varphi}(f) = f(\psi\varphi) = (f\psi)\varphi = \widetilde{\varphi}(f\psi) = (\widetilde{\varphi}\widetilde{\psi})(f) \Rightarrow \widetilde{\psi\varphi} = \widetilde{\varphi}\widetilde{\psi}.$$

Furthermore we have that

$$\widetilde{\mathrm{id}_V}(f) = f\mathrm{id}_V = f = \mathrm{id}_{\Gamma(V)}f \Rightarrow \widetilde{\mathrm{id}_V} = \mathrm{id}_{\Gamma(V)}.$$

□

Theorem 4.3.14. *For varieties $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$, the map*

$$\widetilde{\bullet} : Pol(V,W) \ni \varphi \mapsto \widetilde{\varphi} \in Hom^{K-Alg}(\Gamma(W),\Gamma(V)),$$

is bijective.

Proof. Let $\sigma \in \mathbf{Hom}(\Gamma(W),\Gamma(V))$. For each $i \in \{1,\ldots,m\}$ there is some $f_i \in K[x_1,\ldots,x_n]$ such that $\sigma(y_i + I(W)) = f_i + I(V)$ for each $i \in \{1,\ldots,m\}$. Define

$$\begin{aligned}\psi_\sigma : \mathbb{A}^n &\to \mathbb{A}^m\\ v &\mapsto (f_1(v),\ldots,f_m(v))\end{aligned}$$

which is a polynomial map. Note that for $f \in I(W)$ Some lemmas on polynomial commuting with homomorphism need to be added

$$\begin{aligned}\widetilde{\psi_\sigma}(f)+I(V) &= f(f_1,\ldots,f_m)+I(V) = f(f_1+I(V),\ldots,f_m+I(V))\\ &= f(\sigma(y_1+I(W)),\ldots,\sigma(y_m+I(W))) = \sigma(f(y_1+I(W),\ldots,y_m+I(W)))\\ &= \sigma(f(y_1,\ldots,y_m)+I(W)) = \sigma(f+I(W)) = \sigma(0+I(W)) = 0+I(V).\end{aligned}$$

From this we see that $\widetilde{\psi_\sigma}(I(W)) \subset I(V)$. Note that by HNS $V = V(I(V))$ and $W = V(I(W))$. So given $\psi_\sigma(v) \in \psi_\sigma(V)$, if $f \in I(W)$, then $f(\psi_\sigma(v)) = \psi_\sigma(f)(v) = 0$, hence $\psi_\sigma(V) \subset W$. This implies that

$$\begin{aligned}\varphi_\sigma : V &\to W\\ v &\mapsto \widetilde{\psi_\sigma}(v)\end{aligned}$$

is a well-defined polynomial map. It remains to check that $\sigma \mapsto \varphi_\sigma$ is the mutual inverse of $\varphi \mapsto \widetilde{\varphi}$. Indeed, for $\varphi \in \mathbf{Pol}(V,W)$ defined by $f_1,\ldots,f_m \in K[\mathbf{y}]$, since $\widetilde{\varphi}(y_i + I(W)) = f_i + I(V)$, we have for any $v \in V$ that

$$\varphi_{\widetilde{\varphi}}(v) = (f_1(v),\ldots,f_m(v)) = \varphi(v) \Rightarrow\ \varphi_{\widetilde{\varphi}} = \varphi.$$

Conversely for $\sigma \in \mathbf{Hom}(\Gamma(W),\Gamma(V))$ for any $f + I(W) \in \Gamma(W)$

$$\begin{aligned}\widetilde{\varphi_\sigma}(f+I(W)) &= f\circ\varphi_\sigma + I(W) = f(f_1,\ldots,f_m)+I(W) = f(f_1+I(W),\ldots,f_m+I(W))\\ &= f(\sigma(y_1+I(V)),\ldots,\sigma(y_m+I(V))) = \sigma(f(y_1+I(V),\ldots,y_m+I(V))\\ &= \sigma(f(y_1,\ldots,y_m)+I(V)) = \sigma(f+I(V)),\end{aligned}$$

implying $\widetilde{\varphi_\sigma} = \sigma$. ☐

Remark 4.3.15. We thus have a contravariant functor $(\Gamma(\bullet),\widetilde{\bullet})$ mapping a variety V to $\Gamma(V)$ and a polynomial map $\varphi \in \mathbf{Pol}(V,W)$ to $\widetilde{\varphi} \in \mathbf{Hom}(\Gamma(W),\Gamma(V))$. The above theorem shows that this functor is *fully faithful*.

Corollary 4.3.16. *Let $V \subset \mathbb{A}^n, W \subset \mathbb{A}^m$ be varieties. Then*

$$V \simeq W \iff \Gamma(V) \simeq \Gamma(W),$$

where $\simeq$ refers to an isomorphism of varieties, resp. K-algebras.

Proposition 4.3.17. *Let $V \subset \mathbb{A}^n$ be a variety and $W \subset V$ a subvariety. The ideal $I_V(W) \subset \Gamma(V)$ corresponds to the ideal of polynomial functions in* $\mathrm{Pol}(V,K)$ *vanishing on W, which we therefor also denote by $I_V(W)$. Thus the map* $\mathrm{Pol}(V,K) \to \mathrm{Pol}(W,K), f \mapsto f\,|_W$ *has kernel $I_V(W)$.*

Proof. This follows from the above proposition and Proposition 4.2.33. □

Proposition 4.3.18. *Let $\varphi : V \to W$ be a polynomial map and $V' \subset V$, $W' \subset W$ subvarieties. Suppose $\varphi(V') \subset W'$. Thus $\varphi\,|_{V'}: V' \to W'$ is a polynomial map. Furthermore $\widetilde{\varphi}(I_W(W')) \subset I_V(V')$.*

Proof. Let $f + I(W) \in I(W')/I(W)$ and $v \in V'$. Then $\varphi(v) \in W'$, hence $f \circ \varphi(v) = 0$, meaning $f \circ \varphi \in I(W')$, hence $\widetilde{\varphi}(f + I(W)) = f \circ \varphi + I(V) \in I_V(V')$. □

Lemma 4.3.19. *Let $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ be algebraic set and $\varphi : V \to W$ be a map such that $\varphi(v) = (f_1(v),\dots,f_m(v))$ for $v \in V$ where $f_1,\dots,f_m \in K[x_1,\dots,x_n]$. We then have*

1. *The graph of φ,*

   $$G_\varphi := \left\{(v_1,\dots,v_n,\varphi(v)) \in \mathbb{A}^{n+m} : v \in V\right\} \subset V \times W,$$

   is an algebraic set.

2. *If V, W are varieties, φ is a polynomial map. In this case G_φ is a variety isomorphic to V*

Proof. For suitable $I = \langle g_1,\dots,g_l\rangle \subset K[x_1,\dots,x_n]$, $V = V(I)$.
1. One easily verifies that $G_\varphi = V(g_1,\dots,g_l,y_1,\dots,y_m)$. Indeed, if $(v,w) \in G_\varphi$, $v \in V$, and $w_i = f_i(v)$, hence $g_i(v,w) = 0$ (here we think of g_i in the canonical way as an element of $K[\mathbf{x},\mathbf{y}]$) and $\mathrm{ev}_{(v,w)}(y_i - f_i) = f_i(v) - f_i(v) = 0$. Conversely, if (v,w) is in the right-hand side. Then $g_i(v) = 0$ for every i, hence $v \in V$. Furthermore $w_i - f_i(v) = 0$, hence $w_i = f_i(v)$, hence $w = \varphi(v)$.
2. By HNS we can choose I to be a prime ideal. Let $J = \langle g_1,\dots,g_l,y_1 - f_1,\dots,y_m - f_m\rangle \subset K[\mathbf{x},\mathbf{y}]$. By Corollary 2.9.41

$$K[\mathbf{x}]/I \simeq K[\mathbf{x},\mathbf{y}]/J,$$

hence J is prime. It follows that $G_\varphi = V(J) \subset \mathbb{A}^{n+m}$ is a variety. We can see more from this isomorphism. By HNS $I = I(V)$ and $J = I(G_\varphi)$, thus $\Gamma(V) \simeq \Gamma(G_\varphi)$, hence by Corollary 4.3.16 $V \simeq G_\varphi$. Another way to explicitly construct an isomorphism. Is to consider $\phi : V \to G_\varphi, v \mapsto (v, \varphi(v))$ and $\pi : G_\varphi \to V, (v, \varphi(v)) \mapsto v$. These are clearly mutually inverse polynomial maps. □

Example 4.3.20. Let us consider a couple of examples of a bijective polynomial map that is NOT isomorphisms of varieties and a central example of an elliptic curve "almost" being in bijection with the affine line.

1. Consider the map

$$\begin{aligned} \varphi : \mathbb{A}^1 &\to V := V(y^2 - x^3) \\ t &\mapsto \left(t^2, t^3\right) \end{aligned}$$

This is clearly a well-defined polynomial map. Let $(a,b) \in V$. Let t be a solution to the equation $X^2 - a = 0$. Then $t^2 = a$. Moreover, $b^2 = a^3 = t^6$, hence $\pm t$ is a solution to $X^3 - b = 0$, since then $(b + t^3)(b - t^3) = 0$. Suppose $\left(t^2, t^3\right) = \left(s^2, s^3\right)$. Then $0 = t^2 - s^2 = (t+s)(t-s)$. Then either $t = -s$ or $t = s$. In the first case, we get $s^3 = t^3 = -s^3$, hence $2s^3 = 0$, implying $s = 0$, and hence $t = 0$. It follows that φ is a bijection. Note that $y^2 - x^3$ is irreducible hence $\Gamma(V) = K[x,y]/I$. The induced K-algebra homomorphism is given by

$$\begin{aligned} \widetilde{\varphi} : \Gamma(V) &\to K[z] \\ f + I &\mapsto f(z^2, z^3) \end{aligned}$$

One notes that $\widetilde{\varphi}(\Gamma(V)) = K[z^2, z^3] \subsetneq K[z]$, meaning that $\widetilde{\varphi}$ is not an isomorphism. By Corollary 4.3.16 it follows that φ cannot be an isomorphism.

2. This is exercise 2.13 in Fulton which is a continuation of exercise 1.40 both which are easy and I should get around to doing them at some point.

3. Consider the polynomial map

$$\begin{aligned} \varphi : \mathbb{A}^1 &\to V := V(y^2 - x^2(x+1)) \\ t &\mapsto (t^2 - 1, t(t^2 - 1)) \end{aligned}$$

One readily verifies that this is well-defined. Let $(a,b) \in V$. Let t be a solution to $X^2 - a + 1$. Then

$$b^2 = (t^2 - 1)^2 t^2 \Rightarrow b = \pm t(t^2 - 1).$$

It thus follows that

$$\varphi(\pm t) = (t^2 - 1, \pm t(t^2 - 1)) = (a, b).$$

Suppose $(t^2 - 1, t(t^2 - 1)) = (s^2 - 1, s(s^2 - 1))$ for $t \neq \pm 1$. Then $t^2 = s^2$, hence $s = \pm t$. If $s = -t$, then

$$0 = t(t^2 - 1) - s(s^2 - 1) = -2t(t^2 - 1) \Rightarrow t = 0 \text{ or } t^2 - 1 = 0.$$

by the assumption $t \neq \pm 1$, one concludes $s = t = 0$. Note that $\varphi(\pm 1) = (0, 0)$. One thus concludes that φ is onto and injective but for a double point $(0, 0)$.

4.3.2 (Affine) Coordinate Changes

Definition 4.3.21. Let $I = \langle g_1, \ldots, g_l \rangle \subset K[y_1, \ldots, y_m]$ and $\varphi : \mathbb{A}^n \to \mathbb{A}^m$ a polynomial map given by polynomials $f_1, \ldots, f_m \in K[x_1, \ldots, x_m]$. We define

$$I(\varphi) = \langle \{ f(f_1, \ldots, f_m) \in K[x_1, \ldots, x_n] : f \in I \} \rangle$$

Remark 4.3.22. One easily checks that

$$\langle \{ f(f_1, \ldots, f_m) \in K[x_1, \ldots, x_n] : f \in I \} \rangle = \langle g_1(f_1, \ldots, f_m), \ldots, g_m(f_1, \ldots, f_m) \rangle.$$

Note that the set $I(\varphi)$ is independent of choice of representative of φ, since φ is uniquely determined by it's defining polynomials. Let $W \subset \mathbb{A}^m$ be a variety. Note that $V^{\varphi} := \varphi^{-1}(W) = V(I(\varphi))$, setting $I := I(W)$.

Definition 4.3.23. Let $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ given by polynomials $f_1, \ldots, f_m \in K[x_1, \ldots, x_n]$. φ is called *an affine change of coordinates*, if $\deg f_i = 1$ and φ is bijective for each i.

Remark 4.3.24. For a moment let us remove the condition that φ is bijective. We may write $f_i = b_i + \sum_{j=1}^{n} a_{ij} x_i$, hence for each $v \in \mathbb{A}^n$ we get

$$\varphi(v) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} v + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

In other words φ is a given by a composition $t \circ l$, where $t : \mathbb{A}^n \to \mathbb{A}^n$ is a translation and $l : \mathbb{A}^n \to \mathbb{A}^n$ is a linear map. It follows that φ is an affine change of coordinates if and only if l is bijective, as t is always bijective. Note that the inverse of a linear map and a translation is also respectively a linear map and a translation, hence in particular these are polynomial maps. An affine change of coordinates is there automatically an isomorphism of affine varieties.

Definition 4.3.25. A variety $V \subset \mathbb{A}^n$ is called a *linear subvariety* if $V = V(f_1, \dots, f_m)$ for degree 1 polynomials $f_1, \dots, f_m \in K[x_1, \dots, x_n]$

Lemma 4.3.26. *Let R be a commutative ring. Let $f_1, \dots, f_n, g \in R[x_1, \dots, x_n]$ with $f_i = a_i + \sum_{j=1}^n a_{ij}x_j$ such that $(a_{ij}) \in M_n(R)$ is invertible and $g = c + \sum_{i=1}^n b_i x_i \in R[x]$. Then*

$$g(f_1, \dots, f_n) = 0 \iff g = 0.$$

If $\deg\ g = 1$, *then* $\deg\ g(f_1, \dots, f_n) = 1$

Proof. "$\Leftarrow$": This is obvious.
"$\Rightarrow$": Note that

$$\begin{aligned} g(f_1, \dots, f_n) &= c + \sum_{i=1}^n b_i \left(a_i + \sum_{j=1}^n a_{ij}x_j \right) = c + \sum_{i=1}^n a_i b_i + \sum_{i=1}^n b_i \sum_{j=1}^n a_{ij}x_j \\ &= c + \sum_{i=1}^n a_i b_i + \sum_{j=1}^n \left[\sum_{i=1}^n b_i a_{ij} \right] x_j. \end{aligned}$$

Thus

$$g(f_1, \dots, f_n) = 0 \iff \begin{cases} c + \sum_{i=1}^n a_i b_i = 0, \\ \sum_{i=1}^n b_i a_{ij} = 0 \end{cases}$$

Note that

$$\begin{pmatrix} \sum_{i=1}^n b_i a_{i1} \\ \vdots \\ \sum_{i=1}^n b_i a_{in} \end{pmatrix} = \mathbf{0} \iff (a_{ij})^T \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \mathbf{0} \iff \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \mathbf{0}. \tag{10}$$

We thus also have that $c = 0$, hence $g = 0$. If $\deg\ g = 1$, then $b_i \neq 0$ for some i. Then $\sum_1^n b_i a_{ij} \neq 0$ for some j by (10). $\square$

Lemma 4.3.27.
1. *Let $\varphi, \psi : \mathbb{A}^n \to \mathbb{A}^n$ be polynomial maps and $V \subset \mathbb{A}^n$ a variety such that V^φ is a variety. Then*

$$\left(V^\varphi\right)^\psi = V^{\varphi \circ \psi}.$$

2. *Let $\varphi : \mathbb{A}^n \to \mathbb{A}^n, v \mapsto (f_1(v), \dots, f_n(v))$ be an isomorphism. Let $V = V(g_1, \dots, g_m) \subset \mathbb{A}^n$ be a variety. Then*

$$\left(V^\varphi\right)^{\varphi^{-1}} = V.$$

3. *Let $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ be a polynomial map. Then*

$$\emptyset^\varphi = \emptyset.$$

Proof. Let $V = V(f_1,\dots,f_m)$.
1. Then

$$\left(V^{\varphi}\right)^{\psi} = V(f_1\circ\varphi,\dots,f_m\circ\varphi)^{\psi} = V((f_1\circ\varphi)\circ\psi,\dots,(f_m\circ\varphi)\circ\psi)$$
$$= V(f_1\circ(\varphi\circ\psi),\dots,f_m\circ(\varphi\circ\psi)) = V^{\varphi\circ\psi}.$$

2. From 1. we find

$$\left(V^{\varphi}\right)^{\varphi^{-1}} = V^{\varphi\circ\varphi^{-1}} = V^{\mathrm{id}} = V(f_1\circ\mathrm{id},\dots,f_m\circ\mathrm{id}) = V(f_1,\dots,f_m) = V.$$

□

Proposition 4.3.28. *Let $V\subset\mathbb{A}^n$ be a linear subvariety.*

1. *If $\varphi:\mathbb{A}^n\to\mathbb{A}^n$ given by $f_1,\dots,f_n\in K[x_1,\dots,x_n]$ is an affine change of coordinates, then $V^{\varphi}\subset\mathbb{A}^n$ is a linear subvariety.*
2. *If $V\neq\emptyset$, then there is an affine change of coordinates $\varphi:\mathbb{A}^n\to\mathbb{A}^n$ such that $V^{\varphi}=V(x_{d+1},\dots,x_n)$ for some $1\le d\le n$. Moreover, it thus follows that a linear subvariety is a variety.*
3. *The integer d above is unique, i.e. independent of choice of affine change of coordinates.*

Proof. Let degree 1 polynomials $g_1,\dots,g_m\in K[x_1,\dots,x_n]$ be given such that $V = V(g_1,\dots,g_m)$.
1. Since $\deg\, g_i = 1$ for each i, $\deg\, f_j(g_1,\dots,g_n)=1$ by Lemma 4.3.26, hence $V^{\varphi} = V(g_1(f_1,\dots,f_n),\dots,g_m(f_1,\dots,f_n))$ is a linear subvariety.
2. For the case $n=1$, note $g_i = g_j$ for every i and j, hence $V = V(b+ax)$ where $a\neq 0$. Let $\varphi: v\mapsto a^{-1}v - a^{-1}b$. Then $g\circ\varphi = aa^{-1}x - a^{-1}b + b = x$, hence $V^{\varphi}=V(x_n)$. We prove the result for $n\ge 2$ by induction in m.
Consider first the case $m=1$. Let $V=V(g)$, where $g = b+\sum_1^n a_ix_i\in K[\mathbf{x}]$ is of degree 1. WLOG $a_n\neq 0$. Put $f_i := a_nx_i$ for $i\in\{1,\dots,n-1\}$ and $f_n := \left[-\sum_1^{n-1}a_ix_i\right] + a_n^{-1}x_n - a_n^{-1}b$. Then

$$g(f_1,\dots,f_n) = b+\sum_1^n a_if_i = b+\sum_1^{n-1}a_ia_nx_i - \left[\sum_1^{n-1}a_ia_nx_i\right] + a_na_n^{-1}x_n - a_na_n^{-1}b = x_n.$$

Choosing $\varphi:\mathbb{A}^n\to\mathbb{A}^n, v\mapsto (f_1(v),\dots,f_n(v))$, one finds that $V^{\varphi} = V(g(f_1,\dots,f_n)) = V(x_n) = V(x_{n-1+1})$. It remains to check that φ is invertible. Note that for each $v\in\mathbb{A}^n$, that $\varphi = (a_{ij})v + be_n$, where $a_{ii}=a_n$, $a_{ni} = -a_i$ for $i\in\{1,\dots,n-1\}$, $a_{nn}=a_n^{-1}$.

A is a lower triangular matrix, hence $\det A = \prod_1^n a_{ii} = a_n^{n-2} \neq 0$ MULTIPLE LINEAR ALGEBRA THEOREMS NEED TO BE ADDED, meaning A is invertible.

Suppose the statement is true for some $m \geq 1$. Consider $V = V(g_1, \dots, g_{m+1})$ for degree 1 polynomials $g_1, \dots, g_{m+1} \in K[\mathbf{x}] \setminus 0$. Consider $W = V(g_1, \dots, g_m)$. There is an affine change of coordinates $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ such that $W^\varphi = V(x_{d+1}, \dots, x_n)$ for some $1 \leq d \leq n$. Then

$$V^\varphi = V(x_{d+1}, \dots, x_n) \cap V(g_{m+1} \circ \varphi)) = V(x_{d+1}, \dots, x_n) \cap V(\underbrace{(g_{m+1} \circ \varphi)(x_1, \dots, x_d, \mathbf{0})}_{h}).$$

If $h = 0$, $V^\varphi = V(x_1, \dots, x_d)$. Otherwise $\deg h = 1$, since if not $V(h) = \emptyset$, hence $V^\varphi = \emptyset$, hence $V = \emptyset$leading to a contradiction (here we only rely on the fact that φ is bijective). In this case there following the same procedure as for $m = 1$ is a polynomial map $\psi : \mathbb{A}^d \to \mathbb{A}^d$, such that $V(h) = V(x_d)$. Setting $\phi : \mathbb{A}^n \to \mathbb{A}^n, v \mapsto (\psi(v_1, \dots, v_d), v_{d+1}, \dots, v_n)$, it follows that

$$V^{\varphi \circ \phi} = V(x_{d+1}, \dots, x_n)^\phi \cap V(h \circ \phi) = V(x_d, \dots, x_n),$$

hence $d - 1$ works.

3. Suppose there are affine change of coordinates φ, ϕ such that $V^\varphi = V(x_{d+1}, \dots, x_n)$ and $V^\phi = V(x_{\delta+1}, \dots, x_n)$. Hence setting $\psi := \varphi^{-1} \circ \phi$, one gets $(V^\varphi)^\psi = V^\phi$.

One observation that could make now is that since ψ is an isomorphism, $V^\varphi \simeq V^\phi$. Then

$$K[x_1, \dots, x_d] \simeq \Gamma\left(V^\varphi\right) \simeq \Gamma\left(V^\phi\right) \simeq K[x_1, \dots, x_\delta].$$

This then implies that $\delta = d$ by result about transcendence degrees.

Here we are invoking a lot of theory. One can take a more elementary approach to arguing $d = e$ only using linear algebra. WLOG $d \leq \delta$. Pick $f_1, \dots, f_n \in K[\mathbf{x}]$ such that $\psi(v) = (f_1(v), \dots, f_n(v))$. Write $f_i = b_i + \sum_{j=1}^n a_{ij} x_j$. Then

$$\psi(v) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} v + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \quad \left(v \in \mathbb{A}^n\right).$$

Note that (a_{ij}) is invertible, hence it's rows are linearly independent. Consider the polynomial map,

$$\kappa : \mathbb{A}^n \to \mathbb{A}^{n-d}$$
$$v \mapsto (f_{d+1}(v), \dots, f_n(v)).$$

Note that

$$\kappa(v) = \underbrace{\begin{pmatrix} a_{d+1,1} & \cdots & a_{d+1,n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}}_{A} v + \underbrace{\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}}_{\mathbf{b}} \quad \left(v \in \mathbb{A}^{n-d}\right).$$

Then $A = \kappa - \mathbf{b}$ is a linear map whose kernel is 0. We also have that

$$\{v \in \mathbb{A}^n : v = (v_1, \dots, v_\delta, \mathbf{0})\} = V(x_{\delta+1}, \dots, x_n) = V(x_{d+1}, \dots, x_n)^\psi = \left\{v \in \mathbb{A}^n : Av = -\mathbf{b}\right\}.$$

If $d < \delta$, we get that $e_d, e_{d+1} \in V(x_{d+1}, \dots, x_n)^\psi$. Here is the catch

$$0 = \psi(-e_j) = -\begin{pmatrix} a_{j1} \\ \vdots \\ a_{jn} \end{pmatrix} + \mathbf{b} \iff \begin{pmatrix} a_{j1} \\ \vdots \\ a_{jn} \end{pmatrix} = \mathbf{b}.$$

This implies that $A(e_d - e_{d+1}) = \mathbf{b} - \mathbf{b} = 0$, which contradicts the fact that $\ker A = 0$, since $e_d \neq e_{d+1}$. □

Remark 4.3.29. A further note on V^φ, is that since the pre-image is equal to the image of the inverse of φ, we have that if $V^\varphi = W$ for some variety W, then $V = \varphi(W) = W^{\varphi^{-1}}$.

Definition 4.3.30. The integer k above is called the *dimension* of V, denoted $\dim V$.

Remark 4.3.31. One easily sees that the concept of dimension defined above generalizes the concept of dimension of finite dimensional vector spaces. We shall see later reference that the above notion of dimension can be generalized to any non-empty variety, through the concept of transcendence degree or equivalently Krull-dimension Give reference once these sections are written.

Lemma 4.3.32. *Consider linear subvarieties $V, W \subset \mathbb{A}^n$ such that there exists an affine change of coordinates $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ with $\varphi(V) = W$. Then $V \overset{\varphi}{\simeq} W$ and $\dim V = \dim W$.*

Proof. □

Definition 4.3.33. Let $v, w \in \mathbb{A}^n$ be distinct points. The *line through v and w* is the set

$$L(v, w) := \{t(w - v) + v : t \in K\}$$

Remark 4.3.34. $L(v, w)$ is clearly a line in the sense of Definition 4.1.28

Lemma 4.3.35. *Let $v, w \in \mathbb{A}^n$ be distinct points and $\varphi : \mathbb{A}^n \to \mathbb{A}^n$ an affine change of coordinates. Then*

$$\varphi(L(v,w)) = L(\varphi(v), \varphi(w)).$$

Proof. Pick $A \in M_n(K)$ and $u \in \mathbb{A}^n$ such that $\varphi(x) = Ax + u$ for $x \in \mathbb{A}^n$. Let $t \in K$. The statement follows directly from the following computation

$$t(\varphi(w) - \varphi(v)) + \varphi(v) = t(Aw + u - Av - u) + Av + u = A(t(w - v) + v) + u = \varphi(t(w - v) + v).$$

□

Lemma 4.3.36. *A line $L \subset \mathbb{A}^n$ is a linear subvariety of dimension* 1.

Proof. Lemma 4.1.29 shows that L. WLOG

$$L = \left\{ \begin{pmatrix} a_1 t + b_1 \\ a_2 t + b_2 \\ \vdots \\ a_n t + b_n \end{pmatrix} \in \mathbb{A}^n : t \in K \right\},$$

where $a_1 \neq 0$. From the proof of Lemma 4.1.29 we see that

$$L = V(\{\underbrace{a_1 x_i - a_i x_1 - a_1 b_i + a_i b_1}_{g_i} : i \in \{2, \dots, n\}).$$

Let $A = (a_{ij}) \in M_n(K)$ be the matrix with $a_{11} = 1$, $a_{ii} = a_1^{-1}$ for $i > 1$, $a_{i1} = a_1^{-1} a_i$, $a_{ij} = 0$ for $j > 1$ and $i \neq j$. Since (a_{ij}) is lower triangular, $\det(a_{ij}) = \prod_1^n a_{ii} = a_1^{-(n-1)} \neq 0$, meaning (a_{ij}) is invertible. Set $w = (b_1, \dots, b_n)^T$. Then $\varphi : \mathbb{A}^n \to \mathbb{A}^n, v \mapsto Av + w$ is an affine change of coordinates, given by $f_1 = x_1 + b_1$ and $f_i = a_1^{-1} x_i + a_1^{-1} a_i x_1 + b_i$. We thus find that

$$g_i \circ \varphi = a_1 a_1^{-1} x_i + a_1 a_1^{-1} a_i x_1 + a_1 b_i - a_i x_1 - a_i b_1 - a_1 b_i + a_i b_1 = x_i,$$

for $i > 1$. We thus find that

$$L^\varphi = V(g_2 \circ \varphi, \dots, g_n \circ \varphi) = V(x_2, \dots, x_n),$$

hence $\dim L = 1$. □

Proposition 4.3.37. *Let $v, v' \in \mathbb{A}^2$, L_1, L_2 distinct lines through v and L_1', L_2' distinct lines through v'. There is an affine change of coordinates $\mathbb{A}^2$ such that $\varphi(v) = v'$ and $\varphi(L_i) = L_i'$ for both i.*

Proof. There are distinct $w_1, w_2, w'_1, w'_2 \in \mathbb{A}^2$ such that $L_i = L(v, w_i)$ for $i = 1, 2$ and $L'_i = L(v, w'_i)$ for $i = 1, 2$. We first perform a translation $T : \mathbb{A}^2 \to \mathbb{A}^2, x \mapsto Ix - v$. Then by Lemma 4.3.35,

$$\Lambda_i := T(L_i) = L(0, \underbrace{w_i - v}_{u_i}),$$

and since T is bijective $\Lambda_1 \neq \Lambda_2$. Note also that $u_1 \neq u_2$. Since Λ_1 and Λ_2 are distinct and intersect at 0, $tu_1 \neq su_2$ for each $s, t \in K \setminus 0$ by Proposition 4.1.22, hence u_1 and u_2 are algebraically independent. This means that the matrix $u = (u_{ij}) \in M_2(K)$ is invertible. Then

$$\Lambda'_i := u^{-1}\Lambda_i = L(0, e_i).$$

One notes that $w'_1 - v', w'_2 - v'$ are linearly independent again since $L(0, w'_1 - v')$ intersect only in 0. Thus taking $\psi : \mathbb{A}^n \to \mathbb{A}^n, x \mapsto (w'_{ij} - v'_j)x + v'$. Then

$$\psi(0) = v', \psi(e_i) = \begin{pmatrix} w'_{i1} - v_1 \\ w'_{i2} - v_2 \end{pmatrix} + v' = w'_i - v' + v' = w'_i,$$

hence putting $\varphi = \psi \circ u \circ T$,

$$\varphi(v) = v', \varphi(w_i) = w'_i, \varphi(L_i) = L'_i.$$

□

Lemma 4.3.38. *Let $\mathbb{A}^n(\mathbb{C})$ be equipped with the usual metric induced topology. For any countable set $S \subset \mathbb{A}^n(\mathbb{C})$, $\mathbb{A}^n(\mathbb{C}) \setminus S$ is path-connected.*

Proof. Let $S = \{v_1, \ldots, v_m\}$ we prove the statement by induction in m. for $m = 1$, let $p, q \in \mathbb{A}^n(\mathbb{C}) \setminus S$ be given suppose L is the line parametrized by $[0,1] \to \mathbb{A}^n(\mathbb{C}), t \mapsto (1-t)p + tq$. If $v_1 \notin L$ we are done. If $v_1 \in L$, pick a point $w \in \mathbb{C} \setminus L$. Let $\psi : [0,1] \to \mathbb{C}$ be the composition of the line from p to w with the line from w to q. The line segments parametrized by ψ_i have each exactly one intersection with L at p respectively q. One sees this from the fact that these three line segments are each respectively subsets of $L(p,q), L(p,w)$ and $L(w,q)$ (as lines in $\mathbb{A}^n(\mathbb{C})$ viewed as $\mathbb{A}^{2n}(\mathbb{R})$) and it follows from Proposition 4.1.31 that $\#(L(p,q) \cap L(p,w)) = 1$ and $\#(L(p,q) \cap L(w,q)) = 1$. It thus follows that v is not an element in the line path ψ, hence $\mathbb{A}^n(\mathbb{C})$ is path-connected in the case $m = 1$.

Suppose for each set $S = \{v_1, \ldots, v_m\}$, $p, q \in \mathbb{A}^n(\mathbb{C}) \setminus S$ there is a finite composition of line segments in $\mathbb{A}^n(\mathbb{C})$ connecting p and q.

Let $S = \{v_1, \ldots, v_{m+1}$ and $p, q \in \mathbb{A}^n(\mathbb{C}) \setminus S$, let ψ be the composition of line segments in $A^n(\mathbb{C}) \setminus \{v_1, \ldots, v_m\}$. If $v_{m+1} \notin \psi$ we are done. If $v_{m+1} \in \psi$, let x, y be the respective start and end point of the line segment containing v_{m+1} denoted ρ. We can write this line segment as the composition of a line ρ_1 from x to v_{m+1} and ρ_2 from v_{m+1} to y to Let $0 < \epsilon < \min_{i\in\{1,\dots,m\}} |v_i - v_{m+1}|$. By continuity of ρ_1 and ρ_2 there are points $x' \in B_\epsilon(v_{m+1}) \cap \rho_1$ and $y' \in B_\epsilon(v_{m+1}) \cap \rho_2$ distinct from v_{m+1}. Let ρ_1' be the line segment from x to x' and ρ_2' be the line segment from y' to y. Consider a third point $z \in B_\epsilon(v_{m+1})$ distinct from x', y', v. Let ξ_1 and ξ_2 be the line segment from x' to z respectively z to y'. By convexity of $B_\epsilon(v)$, $\xi_i \subset B_\epsilon(v)$ for $i = 1, 2$. We thus get a path $\rho' := \rho_1' \cup \xi_1 \cup \xi_2 \cup \rho_2'$ from x to y not containing v_{m+1} or indeed any $v_1, \dots, v_m$ by the choice of ϵ. Replacing ρ by ρ' we get a composition of line segments in $\mathbb{A}^n(\mathbb{C}) \setminus S$ from p to q. Thus $\mathbb{A}^n(\mathbb{C}) \setminus S$ is path-connected. $\square$

Proposition 4.3.39. *Let $V \subset \mathbb{A}^n(\mathbb{C})$ be Zariski closed. $\mathbb{A}^n(\mathbb{C}) \setminus V$ is path-connected with respect to the metric induced topology.*

Proof. Let $p, q \in \mathbb{A}^n(\mathbb{C}) \setminus V$. $L(p,q) \cap V = \{w_1, \dots, w_m\}$ by Corollary 4.1.32. By Lemma 4.3.36, $L(p,q) \simeq V(x_2, \dots, x_n) \simeq \mathbb{A}^1(\mathbb{C})$. Note that a polynomial map $\varphi : \mathbb{A}^k(\mathbb{C}) \to \mathbb{A}^l(\mathbb{C})$ is continuous with respect to the usual metric topologies on $\mathbb{A}^k(\mathbb{C})$ and $\mathbb{A}^l(\mathbb{C})$. Hence $L(p,q)$ is homeomorphic to $\mathbb{A}^1(\mathbb{C})$ via a map, ψ, say. Let $S := \psi(L(p,q) \cap V) = \{v_1, \dots, v_m\}$. Then $\mathbb{A}^1(\mathbb{C}) \setminus S$ is path-connected by Lemma 4.3.38, hence $L(p,q) \setminus V = L(p,q) \setminus (L(p,q) \cap V)$ is path-connected, meaning $\mathbb{A}^n(\mathbb{C}) \setminus V \supset L(p,q) \setminus V$ is path-connected. $\square$

4.3.3 The Field of Rational Functions on a Variety & the Local Ring of Rational Functions Defined at a Point

Definition 4.3.40. Given a non-empty variety $V \subset \mathbb{A}^n$, we define the *the field of rational functions on V* to be the field

$$K(V) := Q(\Gamma(V)).$$

The elements of $K(V)$ are called *rational functions on V*.

Remark 4.3.41. This is well-defined by Proposition 4.2.5. The term rational functions is well chosen as $\Gamma(V) \simeq \mathrm{Pol}(V, K)$ (cf. Proposition 4.3.9), hence indeed

$$K(V) = Q(\mathrm{Pol}(V, K)) = \left\{ \frac{f}{g} : f, g \in \mathrm{Pol}(V, K), g \neq 0 \right\}.$$

Definition 4.3.42. For a variety $V \subset \mathbb{A}^n$ and a point $v \in V$, a rational function $f \in K(V)$ is *defined at v* if there are $a, b \in \Gamma(V)$ with $f = \frac{a}{b}$ such that $b(v) \neq 0$.

Remark 4.3.43. If $\Gamma(V)$ is a UFD there is are unique (up to scalar multiplication over K) polynomial functions $a,b \in \Gamma(V)$ with $\gcd(a,b)=1$ such that $f=\frac{a}{b}$. Hence f is defined at v if and only if $b(v)\neq 0$.

Definition 4.3.44. Let $V\subset \mathbb{A}^n$ be a variety and $P\in\mathbb{A}^n$. *The local ring of V at P* is the ring

$$\mathcal{O}_P(V) := \{q \in K(V) : q \text{ defined at } P\}$$

Remark 4.3.45. One readily verifies that $\mathcal{O}_P(V)$ is a subring of $K(V)$ containing $\Gamma(V)$. Indeed, if $\frac{f}{g},\frac{f'}{g'}\in\mathcal{O}_P(V)$, then $gg'(P)\neq 0$

$$\frac{f}{g}+\frac{f'}{g'}=\frac{fg'+f'g}{gg'}\in\mathcal{O}_P(V).$$

Furthermore,

$$\frac{f}{g}\frac{f'}{g'}=\frac{ff'}{gg'}\in\mathcal{O}_P(V).$$

Let $f\in\Gamma(V)$. Then $f=\frac{f}{1}$, and since $1(P)\neq 0$, $f\in\mathcal{O}_P(V)$. It thus also follows that $1,0\in\mathcal{O}_P(V)$.

An alternative way to go about defining the local ring of V at P, is to set $X := \{f\in\Gamma(V) : f(P)\neq 0\}$ and define

$$\mathcal{O}_P(V) := X^{-1}\Gamma(V).$$

One readily verifies that the first definition gives rise to ring that is isomorphic to second ring defined.

Note that X is saturated. Indeed, for $f\in\widehat{X}$ there is an $f'\Gamma(V)$ such that $f'f\in X$. Then $f'(P)f(P)\neq 0$, hence $f(P)\neq 0$.

Definition 4.3.46. Let $V\subset\mathbb{A}^n$ be a variety, $f\in K(V)$. We define *the pole set of f in V* to be the set

$$\mathcal{P}(f) := \{v\in V : f \text{ is not defined at } v\}$$

Definition 4.3.47. Let $f\in K(V)$. Define $J_f := \{g\in K[\mathbf{x}] : (g+I(V))f\in\Gamma(V)\}\subset K[\mathbf{x}]$.

Remark 4.3.48. The above set is an ideal containing $I(V)$. Indeed, if $g,h\in J_f$, then $(g+I(V))f,(h+I(V))f\in\Gamma(V)$, hence

$$((g+h)+I(V))f=(g+I(V))f+(h+I(V))f\in\Gamma(V)\Rightarrow g+h\in J_f,$$

let $r\in K[\mathbf{x}]$. Then

$$(rg+I(V))f=(r+I(V))(g+I(V))f\in\Gamma(V)\Rightarrow rg\in J_f.$$

Suppose $g\in I(V)$. Then $(g+I(V))f=(0+I(V))f=0+I(V)\in\Gamma(V)$, hence $I(V)\subset J_f$.

Lemma 4.3.49. *Let $V \subset \mathbb{A}^n$ be a variety, $f \in K(V)$. $\mathcal{P}(f) = V(J_f)$, hence $\mathcal{P}(f)$ is algebraic.*

Proof. Let $v \in V(J_f)$. Let $(a+I(V)),(b+I(V)) \in \Gamma(v)$ such that $f = \frac{a+I(V)}{b+I(V)}$. Then $b \in J_f$, hence $b(v) = 0$, implying $(b+I(V))(v) = 0$, hence $v \in \mathcal{P}(f)$.
Let $v \in \mathcal{P}(f)$. Let $g \in J_f$, then $h+I(V) := (g+I(V))f \in \Gamma(V)$. If $g \in I(V)$, then clearly $(g+I(V))(v) = (0+I(V))(v) = 0$. Otherwise $f = \frac{h+I(V)}{g+I(V)}$, and since f is not defined at v, $g(v) = (g+I(V))(v) = 0$, hence $v \in V(J_f)$. □

Example 4.3.50. Let $V = V(xw-yz) \subset \mathbb{A}^4$. By a simple application of for instance Eisenstein's criterion $xw-yz$ is indeed irreducible, hence we can consider $\Gamma(V) = K[x,y,z,w]/\langle xw-yz\rangle$. Denote every $a+I(V) \in \Gamma(V)$ by $\overline{a}$. Define $f := \frac{\overline{x}}{\overline{z}} = \frac{\overline{y}}{\overline{w}}$. Note the following fact. Let $a,b \in K[x,y,z,w]$ be given such that $f = \frac{\overline{a}}{\overline{b}}$. Then $az-bx \in I(V) = \langle xw-yz\rangle$. We can thus find $q \in K[x,y,z,w]$ with $az-bx = q(xw-yz)$, hence $z(a+qy) = x(qw+b)$, implying $z \mid qw+b$. One therefor finds that $b = sz-qw$ for some $s \in K[x,y,z,w]$.
A feature of the rational function f is that given a representation of f there is a point $P \in V$ where f is defined at which the denominator vanishes. Indeed, suppose for a contradiction, $f = \frac{\overline{a}}{\overline{b}}$ for some $a,b \in K[x,y,z,w]$ with $\overline{b}(P) \neq 0$ for every $P \in V \setminus \mathcal{P}(f)$. Note that $P_{\alpha,\beta} = (0,0,\alpha,\beta) \in V \setminus \mathcal{P}(f)$ for every $\alpha,\beta \in K$ with $\alpha \neq 0$ or $\beta \neq 0$. Hence in particular $b(0,0,\alpha,\beta) \neq 0$ for every $(\alpha,\beta) \in \mathbb{A}^2$ with $\beta \neq 0$, hence by Proposition write proposition later! $b(0,0,z,w) \in K[w]$. A symmetric argument shows that $b(0,0,z,w) \in K[z]$, hence b is constant. However note that the fact noted at the beginning implies $b(0,0,0,0) = 0$. Then $b = 0$, leading to a contradiction with b not vanishing on points at which f is defined.
One can show that $J_f = \langle z,w\rangle$, which implies that f is defined exactly at the points where $\overline{z}$ or $\overline{w}$ do not vanish. Clearly $z,w \in J_f$. Let $b \in J_f$, then setting $\overline{a} = \overline{b}f$, $f = \frac{\overline{a}}{\overline{b}}$. The fact proven in the beginning of this section shows the other inclusion.

Example 4.3.51. Let $V := V(y^2 - x^2(x+1)) \subset \mathbb{A}^2$. Since the generating polynomial is irreducible, $I(V) = \langle y^2 - x^2(x+1)\rangle$ by HNS. Set $\overline{a} := a + I(V)$ for every $a+I(V) \in \Gamma(V)$. Let $f = \frac{\overline{y}}{\overline{x}}$. We determine $\mathcal{P}(f)$. Note that

$$f = \frac{\overline{y}}{\overline{x}} = \frac{\overline{y}^2}{\overline{xy}} = \frac{\overline{x^2(x+1)}}{\overline{xy}} = \frac{\overline{x(x+1)}}{\overline{y}} \Rightarrow x,y \in J_f.$$

If we can prove that $(0,0) \in \mathcal{P}(f)$ it follows that $\mathcal{P}(f) = V(x,y) = \{(0,0)\}$. Let $\frac{\overline{a}}{\overline{b}} = f$. Then

$$\overline{b}\overline{y} = \overline{a}\overline{x} \Rightarrow \overline{b}(0,c) = \overline{b}(0,c)c = \overline{a}(0,c)0 = 0,$$

for every $c \in K \setminus 0$, hence $b(0,y)$ has infinitely many roots, meaning $b(0,y) = 0$ for a representative $b \in K[x,y]$ of $\overline{b}$. This means that b has no terms not divisible by x. It follows that $\overline{b} = \overline{qx}$ for some $q \in K[x,y]$, hence $\overline{b}(0,0) = \overline{q}(0,0)0 = 0$.
Secondly, let's determine $\mathcal{P}\left(f^2\right)$. Then

$$f^2 = \frac{\overline{y}^2}{\overline{x}^2} = \frac{\overline{x^2(x+1)}}{\overline{x}^2} = \frac{x+1}{1},$$

hence $1 \in J_{f^2}$, meaning $\mathcal{P}\left(f^2\right) = V(J_{f^2}) = V(1) = \emptyset$ by WNS.

Proposition 4.3.52. *Let $V \subset \mathbb{A}^n$ be a variety. Then*

$$\Gamma(V) = \bigcap_{P \in V} \mathcal{O}_P(V).$$

Proof. As $\Gamma(V) \subset \mathcal{O}_P(V)$ for every $P \in V$, the first inclusion is obvious.
Let $f \in \bigcap_{P \in V} \mathcal{O}_P(V)$. Then f is defined at every $P \in V$, hence $V(J_f) = \mathcal{P}(f) = \emptyset$. By the weak Nullstellensatz, we then have that $1 \in J_f$, hence $f = (1 + I(V))f \in \Gamma(V)$. $\square$

Definition 4.3.53. Let $V \subset \mathbb{A}^n$ be a variety, $P \in V$ and $f \in \mathcal{O}_P(V)$. Pick $a, b \in \Gamma(V)$ with $f = \frac{a}{b}$ and $b(P) \neq 0$. We define *the value of f at P* to be

$$f(P) := \frac{a(P)}{b(P)}$$

Remark 4.3.54. This is well-defined. Indeed, if there additionally are $a', b' \in \Gamma(V)$ with $f = \frac{a'}{b'}$ and $b'(P) \neq 0$, then

$$\frac{a}{b} = \frac{a'}{b'} \Rightarrow ab' = a'b \Rightarrow a(P)b'(P) = a'(P)b(P) \Rightarrow \frac{a(P)}{b(P)} = \frac{a'(P)}{b'(P)}.$$

Definition 4.3.55. Let $V \subset \mathbb{A}^n$ be a variety $P \in V$. *The maximal ideal of V at P*. Is the set

$$\mathfrak{m}_P(V) := \{f \in \mathcal{O}_P(V) : f(P) = 0\}$$

Remark 4.3.56. Consider the map

$$\begin{aligned} \mathrm{ev}_P : \mathcal{O}_P(V) &\to K \\ f &\mapsto f(P) \end{aligned}$$

This is easily verfied to be a K-algebra homomorphism. Note that

$$\mathfrak{m}_P(V) = \ker \ \mathrm{ev}_P,$$

hence $\mathfrak{m}_P(V)$ is an ideal.

One can furthermore show that $\mathfrak{m}_P(V)$ is the maximal ideal containing every proper ideal of $\mathcal{O}_P(V)$, hence $\mathcal{O}_P(V)$ will be a local ring (hence the name given to this ring is well chosen). The following proposition will be sufficient to prove this.

Proposition 4.3.57. *Consider $V \subset \mathbb{A}^n$ a variety and $P \in V$. Let $f \in \mathcal{O}_P(V)$.*

$$f \in \mathcal{O}_P(V)^* \iff f(P) \neq 0,$$

Hence,

$$\mathfrak{m}_P(V) = \left\{ f \in \mathcal{O}_P(V) : f \notin \mathcal{O}_P(V)^* \right\}$$

Proof. Consider the set $X := \{f \in \Gamma(V) : a(P) \neq 0\}$. Note that X is saturated and $\mathcal{O}_P(V) = X^{-1}\Gamma(V)$ (cf. Remark 4.3.45). Let $f = \frac{a}{b}$ with $b(P) \neq 0$. Then

$$0 \neq f(P) = \frac{a(P)}{b(P)} \iff a(P) \neq 0 \iff a \in X = \widehat{X} \iff f = \frac{a}{b} \in \mathcal{O}_P(V)^*.$$

The last bi-implication is due to Lemma 2.8.68. □

Corollary 4.3.58. *Consider $V \subset \mathbb{A}^n$ a variety and $P \in V$. $\mathfrak{m}_P(V)$ is the maximal ideal containing every proper ideal of $\mathcal{O}_P(V)$, hence $\mathcal{O}_P(V)$ is local. In addition $\mathcal{O}_P(V)$ is Noetherian.*

Proof. By the above proposition, $\mathcal{O}_P(V) \setminus \mathcal{O}_P(V)^*$ is an ideal, hence by Proposition 2.8.59 $\mathcal{O}_P(V)$ is local where $\mathfrak{m}_P(V)$ is the unique maximal ideal.
Hilbert's basis theorem shows that $K[\mathbf{x}]$ is Noetherian, by Lemma 2.4.60 $\Gamma(V)$ is Noetherian. Let $I \subset \mathcal{O}_P(V)$. Let $J := I \cap \Gamma(V) \subset \Gamma(V)$. Then by Theorem 2.4.58 there are $f_1, \ldots, f_m \in \Gamma(V)$ such that $J = \langle f_1, \ldots, f_m \rangle$. Let $f \in I$ and pick $f = \frac{a}{b}$ with $b(P) \neq 0$. Then $bf \in \Gamma(V)$, hence $bf = \sum_1^m \lambda_i f_i$ for suitable $\lambda_1, \ldots, \lambda_m \in \Gamma(V)$, hence $f = \sum_1^m \frac{\lambda_i}{b} f_i$. Then I is generated by $f_1, \ldots, f_m$ in $\mathcal{O}_P(V)$. □

Corollary 4.3.59. *Consider $V \subset \mathbb{A}^n$ a variety and $P \in V$. Let $I \subset \mathcal{O}_P(V)$ be a proper ideal. Then $I \subset \mathbf{ker}\,(\mathbf{ev}_P : \mathcal{O}_P(V) \to K)$, hence in addition $I \cap \Gamma(V) \subset \mathbf{ker}\,(\mathbf{ev}_P : \Gamma(V) \to K)$*

Proof. By the above corollary $I \subset \mathfrak{m}_P(V) = \mathbf{ker}\,(\mathbf{ev}_P : \mathcal{O}_P(V) \to K)$. One readily verifies that $\mathbf{ker}\,(\mathbf{ev}_P : \mathcal{O}_P(V) \to K) \cap \Gamma(V) = \mathbf{ker}\,(\mathbf{ev}_P : \Gamma(V) \to K)$. □

Lemma 4.3.60. *Let V be an affine variety and $P \in V$. There is a one-to-one correspondence between prime ideals in $\mathcal{O}_P(V)$ and prime ideals in $\Gamma(V)$ contained in* $\mathbf{ker}\,(\mathbf{ev}_P : \Gamma(V) \to K)$

Proof. Let $I \subset \mathcal{O}_P(V)$ be prime. Then $I \cap \Gamma(V) \subset \Gamma(V)$ is prime by Proposition 2.8.24. By Corollary 4.3.59 $I \cap \Gamma(V) \subset \ker (\mathrm{ev} : P : \Gamma(V) \to K)$.

Let $J \subset \Gamma(V)$ be a prime ideal whose elements all vanish on P. Then $X^{-1}J$ is a proper ideal of $\mathcal{O}_P(V)$. Let $\frac{a}{x}, \frac{b}{x} \in \mathcal{O}_P(V)$ with $\frac{ab}{xy} \in X^{-1}I$, hence for some $\frac{c}{z} \in X^{-1}I$, $\frac{ab}{xy} = \frac{c}{z}$. Then $abz \in I$, hence $ab \in I$ or $z \in I$. Since $z(P) \neq 0$, $z \notin I$. Then $ab \in I$, meaning $a \in I$ or $b \in I$. It thus follows that $\frac{a}{x} \in X^{-1}I$ or $\frac{b}{y} \in X^{-1}I$. □

Lemma 4.3.61. *Let V be an affine variety and $P \in V$. There is a one-to-one correspondence between prime ideals in $\Gamma(V)$ contained in $\ker \mathrm{ev}_P$ and subvarieties W of V containing P.*

Proof. Let $I \subset \Gamma(V)$ be prime with $I \subset \ker \mathrm{ev}_P$. By Proposition 4.2.32 there is a subvariety $W \subset V$ such that $I = I_V(W)$. Let $f \in I(W)$, then $f(P) = \mathrm{ev}_P(f + I(V)) = 0$, hence $P \in V(I(W)) = W$.

Let $W \subset V$ be a subvariety containing P. Then $I_V(W)$ is prime and $\mathrm{ev}_P(f + I(V)) = 0$ for every $f + I(V) \in I_V(W)$. □

Proposition 4.3.62. *Let V be an affine variety and $P \in V$. There is a one-to-one correspondence between prime ideals in $\mathcal{O}_P(V)$ and subvarieties of V containing P.*

Proof. This follows directly from the two above lemmas. □

Proposition 4.3.63. *Let V be an affine variety. Denote $h + I(V)$ by $\overline{h}$ for $h \in K[\mathbf{x}]$. Consider $f = \frac{\overline{a}}{\overline{b}} \in K(V)$. Let $U = \{P \in V : f$ defined at $P\}$. Consider the function $\mathrm{ev}_\bullet(f) : U \to K$. f is uniquely determined by this function.*

Proof. Let $g = \frac{\overline{c}}{\overline{d}} \in K(V)$ be given such that $\mathrm{ev}_\bullet(g) = \mathrm{ev}_\bullet(f)$. Let $P \in U$. Then $\overline{a}(P)\overline{d}(P) = \overline{c}(P)\overline{b}(P)$ hence $(ad - cb)(P) = \overline{a}(P)\overline{d}(P) - \overline{c}(P)\overline{b}(P) = 0$. We thus see that $(ad - cb)(v) = 0$ for every $v \in V$, hence $ad - cb \in I(V)$, meaning $\overline{a}\overline{d} - \overline{c}\overline{b} = 0$, hence $f = g$. □

Proposition 4.3.64. *Let $\varphi : V \to W$ a polynomial map and $P \in V$. $\widetilde{\varphi} : \Gamma(W) \to \Gamma(V)$ extends uniquely to a K-algebra homomorphism $\widetilde{\varphi} : \mathcal{O}_{\varphi(P)}(W) \to \mathcal{O}_P(V)$. Furthermore $\widetilde{\varphi}(\mathfrak{m}_{\varphi(P)}(W)) \subset \mathfrak{m}_P(V)$.*

Proof. Let $X = \{a \in \Gamma(W) : a(\varphi(P)) \neq 0\}$. Let $a \circ \varphi \in \widetilde{\varphi}(X) = \{\widetilde{\varphi}(a) = a \circ \varphi \in \widetilde{\varphi}(X) : a \in X\}$. Then $a \circ \varphi(P) = a(\varphi(P)) \neq 0$, hence the first statement follows from Lemma 2.8.72.

Let $f \in \mathfrak{m}_{\varphi(P)}(W) = \{f \in \mathcal{O}_{\varphi(P)}(W) : f(\varphi(P)) = 0\}$. Then

$$\widetilde{\varphi}(f)(P) = (f \circ \varphi)(P) = f(\varphi(P)) = 0 \Rightarrow \widetilde{\varphi}(f) \in \mathfrak{m}_P(V).$$

□

Corollary 4.3.65. *Let V, W be varieties, $P \in V$ and $\varphi : V \to W$ an isomorphism. Then $\mathcal{O}_{\varphi(P)}(W) \simeq \mathcal{O}_P(V)$.*

Proof. This follows from the above proposition and Corollaries 4.3.16, 2.8.75. $\square$

Corollary 4.3.66. *Let $V \subset \mathbb{A}^n$ be a subvariety and $P \in V$. Let $\varphi : \mathbb{A}^n \to \mathbb{A}^n, v \mapsto Av + w$ be an affine change of coordinates. Then $\widetilde{\varphi} : \mathcal{O}_{\varphi(P)}(\mathbb{A}^n) \to \mathcal{O}_P(\mathbb{A}^n)$ is an isomorphism. Furthermore, $\widetilde{\varphi}$ induces an isomorphism between $\mathcal{O}_{\varphi(P)}(V) \simeq \mathcal{O}_P(V^{\varphi})$.*

Proof. The first statement follows from the above corollary. One notes that $V^{\varphi} \overset{\varphi'}{\simeq} V$, where φ' is the restriction of φ to V. Hence $\Gamma(V) \overset{\widetilde{\varphi'}}{\simeq} \Gamma(V^{\varphi}), v + I(V) \mapsto \varphi(v) + I(V^{\varphi})$. The above corollary shows that $\varphi' : \mathcal{O}_{\varphi(P)}(V) \to \mathcal{O}_P(V^{\varphi}), \frac{a}{b} \mapsto \frac{\widetilde{\varphi'}(a)}{\widetilde{\varphi'}(b)}$ is an isomorphism. $\square$

Proposition 4.3.67. *Let $P = \mathbf{0} \in \mathbb{A}^n$. Set $\mathcal{O} := \mathcal{O}_P(\mathbb{A}^n)$ and $\mathfrak{m} := \mathfrak{m}_P(\mathbb{A}^n)$. Let $I := \langle x_1, \ldots, x_n \rangle \subset K[x_1, \ldots, x_n]$. Then $I\mathcal{O} = \mathfrak{m}$, hence $I^r\mathcal{O} = \mathfrak{m}$.*

Proof. Firs note that for $f = \sum_1^n \frac{a_i}{b_i} x_i \in I\mathcal{O}$,

$$f(P) = \sum_1^n 0 \frac{a_i(P)}{b_i(P)} = 0 \Rightarrow f \in \mathfrak{m}.$$

Hence $I\mathcal{O}$ is a proper ideal. Let $f = \frac{a}{b} \in \mathfrak{m}$. Then $a(P) = 0$, hence $a \in \langle x_1, \ldots, x_n \rangle = I$ (cf. Proposition 2.9.38). $\square$

Proposition 4.3.68. *Let $V \subset \mathbb{A}^n$ be an affine variety, set $I := I(V) \subset K[x_1, \ldots, x_n]$, pick $P \in V$ and let $J \subset K[\mathbf{x}]$ be an ideal containing I. Let $J' = J/I$. Then $\mathcal{O}_P(\mathbb{A}^n)/J\mathcal{O}_P(\mathbb{A}^n) \simeq \mathcal{O}_P(V)/J'\mathcal{O}_P(V)$ as K-algebras. In particular $\mathcal{O}_P(\mathbb{A}^n)/I\mathcal{O}_P(\mathbb{A}^n) \simeq \mathcal{O}_P(V)/0\mathcal{O}_P(V) = \mathcal{O}_P(V)$.*

Proof. Consider the surjective ring homomorphism, $\sigma : \mathcal{O}_P(\mathbb{A}^n) \twoheadrightarrow \mathcal{O}_P(V)$ given by restriction, i.e $\frac{a}{b} \mapsto \frac{a+I}{b+I}$. Note that for $j\frac{a}{b} \in J\mathcal{O}_P(\mathbb{A}^n)$,

$$\sigma(j\frac{a}{b}) = (j + I)\frac{a + I}{b + I} \in (J/I)\mathcal{O}_P(V) = J'\mathcal{O}_P(V).$$

Conversely if $\sigma(\frac{g}{h}) \in J'\mathcal{O}_P(V)$, then $\frac{g+I}{h+I} = (j + I)\frac{a+I}{b+I}$, $gb + I = jha + I$, hence $gb + I \in J'\mathcal{O}_P(V)$, hence $gb \in J$, implying $\frac{g}{h} = \frac{bg}{bh} \in J\mathcal{O}_P(\mathbb{A}^n)$. This means

$$\varsigma : \mathcal{O}_P(\mathbb{A}^n)/J\mathcal{O}_P(\mathbb{A}^n) \to \mathcal{O}_P(V)/J'\mathcal{O}_P(V), f + J\mathcal{O}_P(\mathbb{A}^n) \mapsto f + J'\mathcal{O}_P(V),$$

is an isomorphism of K-algebras. $\square$

Lemma 4.3.69. *I feel that this lemma should be added earlier* *Consider ideal* $I, J \subset K[x_1, \dots, x_n]$. I, J *are comaximal iff* $V(I) \cap V(J) = \emptyset$

Proof. This follows immediately from the weak Nullstellensatz:

$$I + J = \langle 1 \rangle \iff \emptyset = V(I+J) = V(I) \cap V(J).$$

□

4.3.4 Rational Functions and DVR's

Example 4.3.70. Let $V = \mathbb{A}^1$. Then $K(V) = K(X)$. Let $a \in V$. Then $\mathcal{O}_a(V)$ is a non-field integral domain which is local and Noetherian. We prove that $\mathfrak{m}_a(V) = \langle x - a \rangle$. Indeed, $\mathrm{ev}_a(x-a) = 0$, proving the first inclusion. If $f \in \mathfrak{m}_a(V)$, then $f(a)$, hence $x - a \mid f$. The maximal ideal of $\mathcal{O}_a(V)$ is thus principal ideal with uniformizing parameter $x - a$.

Definition 4.3.71. We define the local ring at infinity to be the ring

$$\mathcal{O}_\infty := \left\{ \frac{f}{g} \in K(X) : \deg\, f \leq \deg\, g \right\}.$$

Remark 4.3.72. One readily verifies that this is a subring of $K(X)$ by noting that if $a, b, c, d \in K[x]$ with $\deg\, a \leq \deg\, b$ and $\deg\, c \leq \deg\, d$, then $\deg\, ac = \deg\, a + \deg\, c \leq \deg\, b + \deg\, d$ and $\deg\, ad + cb \leq \max(\deg\, ad, \deg\, cb) \leq bd$. Furthermore, $1, 0 \in \mathcal{O}_\infty$ Change def of degree?, since $1 = \frac{1}{1}$ and $0 = \frac{0}{1}$. It is a subring of $K(X)$, hence it is an integral domain. It is not a field since $\frac{1}{x} \in \mathcal{O}_\infty$ but $x \notin \mathcal{O}_\infty$. One can check that the units of $\mathcal{O}_\infty$ are the fractions $\frac{f}{g}$ with $\deg\, f = \deg\, g$. Indeed, $\frac{g}{f}$ is the inverse of such elements. Conversely, if $\frac{f}{g}$ is invertible, then $\frac{g}{f} \in \mathcal{O}_\infty$.

Proposition 4.3.73. $\mathcal{O}_\infty$ *is a DVR.*

Proof. We prove that $t := \frac{1}{x}$ is a uniformizing parameter. Suppose $\frac{1}{t} = \frac{a}{b}\frac{c}{d}$. Then

$$1 + \deg\, a + \deg\, c = \deg\, t + \deg\, a + \deg\, d = \deg\, b + \deg\, d.$$

Then $\deg\, a = \deg b$ or $\deg\, c = \deg\, d$, hence $\frac{a}{b}$ is a unit or $\frac{c}{d}$ is a unit. Consequently, t is irreducible. Let $\frac{f}{g} \in \mathcal{O}_\infty \setminus 0$ with $n = \deg\, f$ and $m := \deg\, g$. Set $u = \frac{fx^{m-n}}{g}$. This has a mutual inverse $\frac{g}{fx^{m-n}}$ in $\mathcal{O}_\infty$. Then

$$\frac{f}{g} = \frac{fx^{m-n}}{g} \frac{1}{x^{m-n}}.$$

The uniqueness really just follows from t being irreducible. Since if $ut^{m-n} = vt^l$ for some other unit v and $l \geq 0$, then $m - n - l = 0$ and hence $u = v$. By Proposition 2.8.81, $\mathcal{O}_\infty$ is a DVR. □

Proposition 4.3.74. *The only DVR's with quotient field $K(X)$ (recall K is assumed algebraically closed) containing K are $\mathcal{O}_\infty$ and $\mathcal{O}_a(\mathbb{A}^1) = \left\{\frac{f}{g}(x-a)^n : \frac{f}{g} \in K(x), (x-a) \nmid g, n \geq 0\right\}$, where $a \in K$.*

Proof. $K \subset R \subsetneq K(x)$ be a DVR. Let $\mathfrak{m} = \left\langle \frac{f}{g} \right\rangle$ be the maximal ideal of R with $\gcd(f,g) = 1$.

Suppose first that $K[x] \subset R$. Then $I := K[x] \cap \mathfrak{m} = \langle h \rangle$ (recall that $K[x]$ is a PID) is a prime ideal in $K[x]$ by Proposition 2.8.24. Since $\frac{f}{g}$ is irreducible, it is non-zero. Then $f = g\frac{f}{g} \in \mathfrak{m} \cap K[x]$, meaning $h \neq 0$. Then $h = x-a$ for some $a \in K$ using the assumption that K is algebraically closed. We prove that $R = \mathcal{O}_a(\mathbb{A}^1)$. Let $\frac{\lambda}{\mu} \in R \setminus 0$ with $\gcd(\lambda,\mu) = 1$. Suppose for a contradiction that $h \mid \mu$. Then $\mu \in \mathfrak{m}$, but then $\lambda = \mu\frac{\lambda}{\mu} \in \mathfrak{m} \cap K[x] = \langle h \rangle$, contracting the assumption $\gcd(\lambda,\mu) = 1$. Then $\mu(a) \neq 0$, which implies $\frac{\mu}{\lambda} \in \mathcal{O}_a(\mathbb{A}^1)$, hence $R \subset \mathcal{O}_a(\mathbb{A}^1)$. In particular we get that $\frac{1}{g} \in \mathcal{O}_a(\mathbb{A}^1)$. This means that $\mathfrak{m} \subset \mathcal{O}_a(\mathbb{A}^1)\mathfrak{m} = \mathcal{O}_a(\mathbb{A}^1)f \subset \mathcal{O}_a(\mathbb{A}^1)h$. It follows from Proposition 2.8.85 that $\mathcal{O}_a(\mathbb{A}^1) = R$.

Suppose now that $K[x] \not\subset R$. In particular we must then have that $x-a \notin R$ for any $a \in K$. Then $x-a \in K(x) \setminus R$, which implies that $\frac{1}{x-a} \in \mathfrak{m}$ by Lemma 2.8.84. Note that for $a,b \in K$,

$$\frac{x-a}{x-b} = 1 + \frac{b-a}{x-b} \in R.$$

This means that $\frac{x-b}{x-a} \in R^*$. Now, let $\frac{h}{k} \in R \setminus 0$ with $\gcd(h,k) = 1$. Suppose for a contradiction that $\deg\, h > \deg\, k$. Then we can write $\frac{h'}{k}h''$ with $h', h'' \in K[x]$ such that $\deg\, h' = \deg\, k$ and $\deg\, h'' \geq 1$. But then

$$h'' = \frac{k}{h''}\frac{h}{k} \in R.$$

Then for some $\alpha \in K$ and $l \in K[x]$, $(x-\alpha) = \frac{h''}{l} \in R$, but then $K[x] \subset R$, leading to a contradiction. We thus conclude that $R \subset \mathcal{O}_\infty$. By irreducibility of $\frac{f}{g}$ it is clear that since $\frac{f}{g} \mid \frac{1}{x}$, $\frac{f}{g} = u\frac{1}{x}$ for some unit $u \in R$. We thus get that $\mathfrak{m} \subset \mathcal{O}_\infty \frac{1}{x}$. It follows from Proposition 2.8.85 that $\mathcal{O}_\infty = R$. □

4.3.5 Ideals with a Finite Number of Zeroes

Theorem 4.3.75. *Let $I \subset K[x_1,\ldots,x_n]$ and suppose $\#V(I) < \infty$. Denote the points of $V(I)$ by $P_1,\ldots,P_m$. Set $\mathcal{O}_i := \mathcal{O}_{P_i}(\mathbb{A}^n)$. Then*

$$K[\mathbf{x}]/I \simeq \prod_1^m \mathcal{O}_i/I\mathcal{O}_i$$

Proof. Set $I_i := I(\{P_i\}) = \langle x_1 - P_{i1}, \ldots, x_n - P_{in}\rangle$. We can construct

$$\sigma_i : K[\mathbf{x}]/I \to \mathcal{O}_i/I\mathcal{O}_i$$
$$f + I \mapsto f + I\mathcal{O}_i$$

We thus get a ring homomorphism

$$\sigma : K[\mathbf{x}]/I \to \prod_1^m \mathcal{O}_i/I\mathcal{O}_i$$
$$f + I \mapsto (\sigma_1(f+I), \ldots, \sigma_m(f+I))$$

We aim to prove that σ is an isomorphism. We need to make few constructions before we are able to do so. Note that by HNS, $\mathbf{rad}(I) = I(V(I)) = I(P_1, \ldots, P_m) = \bigcap_1^m I_i$. Note also that $V(I_i) \cap V(\bigcap_{j \neq i} I_j) = \emptyset$, hence by Lemma 4.3.69 I_i and $\bigcap_{j\neq i} I_j$ are comaximal. Thus by Lemma 2.8.33 and Lemma 2.8.37,

$$\bigcap_1^m I_i^d = \left(\prod_1^m I_i\right)^d = \left(\bigcap_1^m I_i\right)^d \subset I,$$

for some $d \geq 0$. For each $i \in \{1, \ldots, m\}$ pick $F_i \in K[\mathbf{x}]$ such that $F_i(P_j) = 0$ and $F_i(P_i) = 1$ for $j \neq i$ (such a polynomial exist do to Corollary 4.1.43). Set $E_i := 1 - (1 - F_i^d)^d$ for each i. Note that

$$E_i + \langle F_i^d \rangle = 1 - (1 - F_i^d)^d + \langle F_i^d \rangle = 1 - 1 + \langle F_i^d \rangle = 0 + \langle F_i^d \rangle \Rightarrow E_i \in \langle F_i^d \rangle \subset \bigcap_{j \neq i} I_j^d,$$

hence $E_i E_j \in \bigcap_1^m I_i^d \subset I$. Furthermore,

$$1 - \sum_1^m E_j + I_i^d = (1 - E_i) - \sum_{j \neq i} E_j + I_i^d = 0 + I_i^d,$$

for each i, hence $1 - \sum_1^m E_j \in \bigcap_1^m I_j^d$. Note also that

$$E_i^2 - E_i = E_i(E_i - 1) = E_i(1 - (1 - F_i^d)^d - 1) = -E_i(1 - F_i^d)^d \in I_i^d \bigcap_{j \neq i} I_j^d = \bigcap_1^m I_j^d \subset I.$$

Set $e_i := E_i + I$. Then $e_i e_j = 0$ for each $i \neq j$, $e_i^2 = e_i$ for each i and $\sum_1^m e_i = 1$.
Claim: If $G \in K[\mathbf{x}]$ and $G(P_i) \neq 0$. Then there is a $t \in K[\mathbf{x}]/I$ such that $tg = e_i$ where $g := G + I$.
Proof of Claim: WLOG $G(P_i) = 1$. Set $H := 1 - G$ and $h := H + I$. Since $H(P_i) = 0$, $H_i \in I_i$, hence $H_i^d E_i \in I_i^d \bigcap_{j \neq i} I_j^d = \bigcap_1^m I_j^d \subset I$. Note that

$$(1 - H) \sum_0^{d-1} E_i H^k = E_i - E_i H^d.$$

Considering the image of the left- and right-hand side in $K[\mathbf{x}]/I$ we therefor get that

$$g\underbrace{\left(-\sum_0^{d-1} e_i h^k\right)}_{t} = e_i.$$

Equipped with the above claim we can proceed by proving σ is injective. Suppose $f := F + I \in K[\mathbf{x}]/I$ is given such that $f \in I\mathcal{O}_i$ for each i. For each i we can then pick a $G_i \in K[\mathbf{x}]$ such that $G_i(P_i) \neq 0$ and $G_i F \in I$. Set $g_i = G_i + I$. Then there is a $t_i g_i = e_i$. Then

$$f = f\sum_1^m e_i = \sum_1^m t_i g_i f = 0 \Rightarrow \ker\sigma = 0.$$

Before we prove that σ is surjective we record a few facts. Since $E_i(P_i) \neq 0$, E_i is a unit in $\mathcal{O}_i$, hence e_i is a unit in $\mathcal{O}_i/I\mathcal{O}_i$. This in particular means,

$$\sigma_i(e_i)\sigma_i(e_j) = \sigma_i(e_i e_j) = 0 \Rightarrow \sigma_i(e_j) = 0.$$

This means

$$\sigma_i(e_i) = \sigma_i(e_i) + \sigma_i\left(\sum_{j\neq i} e_j\right) = \sigma_i\left(\sum_1^m e_j\right) = \sigma_i(1) = 1.$$

Consider an arbitrary element $z = \left(\frac{a_1}{s_1}, \ldots, \frac{a_m}{s_m}\right) \in \prod_1^m \mathcal{O}_i/I\mathcal{O}_i$. For a suitable t_i, $s_i t_i = e_i$. Then

$$\frac{a_i}{s_i} = \frac{a_i t_i}{s_i t_i} = \frac{a_i t_i}{e_i} = a_i t_i \sigma_i(e_i)^{-1} = a_i t_i.$$

Then we get that

$$\sigma_i\left(\sum_1^m a_j t_j e_j\right) = \sigma_i(a_i t_i) = \frac{a_i}{s_i} \Rightarrow$$
$$\sigma\left(\sum_1^m a_j t_j e_j\right) = \left(\sigma_1\left(\sum_1^m a_j t_j e_j\right), \ldots, \sigma_m\left(\sum_1^m a_j t_j e_j\right)\right) = \left(\frac{a_1}{s_1}, \ldots, \frac{a_m}{s_m}\right) = z,$$

meaning σ is surjective. □

Corollary 4.3.76. *Let $I \subset K[x_1, \ldots, x_n]$ and suppose $\#V(I) < \infty$. Denote the points of $V(I)$ by $P_1, \ldots, P_m$. Set $\mathcal{O}_i := \mathcal{O}_{P_i}(\mathbb{A}^n)$. Then*

$$\dim_k\ K[\mathbf{x}]/I = \dim_k\left(\prod_1^m \mathcal{O}_i/I\mathcal{O}_i\right) = \sum_1^m \dim_k\ \mathcal{O}_i/I\mathcal{O}_i$$

Corollary 4.3.77. *In the setup above suppose $m = 1$ then*

$$K[\mathbf{x}]/I \simeq \mathcal{O}_P(\mathbb{A}^n)/I\mathcal{O}_P(\mathbb{A}^n).$$

Corollary 4.3.78. *Suppose $R \supset K$ and $\dim_k\ R < \infty$. Then $R \simeq \prod_1^m R_i$ where R_i are local.*

Proof. Since R is finite dimensional over K it is also ring finite over K generated by some linear basis $\{a_1,\dots,a_n\}$ of R over K, hence $R \simeq K[x_1,\dots,x_n]/I$ (cf. Proposition 2.10.3). Then by a suitable corollary of HNS,

$$V(I) \leq \dim_K\ K[\mathbf{x}]/I = \dim_K\ R < \infty.$$

It follows by the theorem of this subsection that $R \simeq K[\mathbf{x}]/I \simeq \prod_1^m R_i$, where R_i is local for each i. $\square$

4.4 Local Properties of Affine Plane Curves

Fix again an algebraically closed field K. Sometimes we will need that $\mathbf{char}\ K = 0$ (we will have to make some considerations with derivatives, so to make things more simple we makes this assumption about characteristic). We first define an equivalence relation of polynomials in $K[x,y]$.

Definition 4.4.1. We say that a pair of polynomials $f,g \in K[x,y]$ are *(factor) equivalent* (the "factor"-part is non-standard) if $f = \lambda g$ for some $\lambda \in K\setminus 0$. (This is the definition of $\sim$ from Definition 2.7.9 in the special case $V = K[x,y]_{\geq 1}$).

Definition 4.4.2. *An affine plane curve* C is an equivalence class under factor equivalence, i.e. an element of $K[x,y]/\sim$. The *degree* of C (denoted $\mathbf{deg}\ C$) is the degree of a representative of C (This is notion is clearly independent of choice of representative). A degree 1 curve is called a *line*. The irreducible factors of any representative of C up to multiplication by a scalar are called *components* of C. A component with multiplicity 1 in the factorization of a representative of C is called a *simple* components. A non-simple component is called a *multiple* component.

Remark 4.4.3. We already now that $I(V(C)) = f_1\cdots f_m$ where $f_1,\dots,f_m$ are the components of C. We thus lose information about multiplicities. For an irreducible polynomial $f \in K[x,y]$, $V(f)$ is a variety. We define $\Gamma(f) := \Gamma(V(f))$, $K(f) := K(V(f))$, $\mathcal{O}_P(f) := \mathcal{O}_P(V(f))$. We know that each curve in $\mathbb{A}^2$ arises as $V(f)$ for some $f \in K[x,y]$. It thus follows that there is a bijection between algebraic curves in $\mathbb{A}^2$ and algebraic curves in $(K[x,y]_{\geq 1}\setminus 0)/\sim$, so the choice of terminology is well-chosen.

Definition 4.4.4. let f be a curve and $P = (a,b) \in f$. P is called *simple* if $f_x(P) \neq 0$ or $f_y(P) \neq 0$. The *tangent line* at a simple point is the line

$$f_x(P)(x-a) + f_y(P)(y-b).$$

If P is not simple it is called *multiple/singular*. If every point on a curve is simple, then the curve is called *non-singular*.

Example 4.4.5. Each of the described curves will be in an algebraically closed field of characteristic 0 (think $\mathbb{C}$).

1. A *parabola*, $f = y - ax^2 - bx - c$, $a,b,c \in K$ with $a \neq 0$ is a non-singular curve since $f_y = 1$

2. Consider the curve $f = y^2 - x^3 + x$, an example of an *elliptic curve*. In this instance f is non-singular. Indeed, $f_x = -3x^2 + 1$, $f_y = y$. If $f_y(a,b) = 0$, then $0 = f(x,b) = -x^3 + x = -x(x^2 - 1) = x(x-1)(x+1)$, has solution $a \in \{0, \pm 1\}$. Then $f_x(a,b) \in \{1, -2\}$. On the other hand, if $f_x(a,b) = 0$, then $a = \pm\frac{1}{\sqrt{3}}$. Then $0 = f(\pm a, y)$ has solutions $b \in \{\pm\sqrt{-a^3 + a}, \pm\sqrt{a^3 - a}\}$, hence $b \neq 0$. It follows that $f_y(b) \neq 0$.

3. Consider the curve $f = y^2 - x^3$. It has its only singular point at $(0,0)$. Indeed, $f_x = -3x^2$ and $f_y = 2y$. Therefor $-3a^2 = f_x(a,b) = 0$ and $2b = f_y(a,b) = 0$ if and only if $(a,b) = (0,0)$.

4. Consider the curve $f = y^2 - x^3 - x^2$. It has its only singular point at $(0,0)$. Indeed, $f_x = -3x^2 - 2x = -3x(x + 2/3)$ and $f_y = 2y$. Thus the only common zeroes of f_x and f_y are $(0,0)$ and $(-2/3, 0)$. The only common zero of f_x and f_y lying on f is thus $(0,0)$.

Definition 4.4.6. For a curve f write $f = \sum_m^n f_i$, where $f_m, \ldots, f_n$ are homogeneous polynomials with degree corresponding to their index and $f_m \neq 0$. We call the value m the *multiplicity of f at* $(0,0)$, we denote it $m_0(f)$. If $m = 2$ we call $(0,0)$ a *double point*, if $m = 3$ we call $(0,0)$ a *triple point*, etc.

Lemma 4.4.7. *Let f be a curve. Then $m_0(f) > 0$ if and only if $(0,0) \in f$.*

Proof. $m_0(f) > 0$ is equivalent to f having constant term zero which is equivalent to $(0,0) \in f$. □

Lemma 4.4.8. *Let f be a curve with $(0,0) \in f$. Then $m_0(f) = 1$ if and only if $(0,0)$ is a simple point in f. In this case f_1 is the tangent line at 0.*

Proof. "$\Rightarrow$:" If $m_0(f) = 1$. Then $f_1 = ax + by$ where $a \neq 0$ or $b \neq 0$. Then $(f_1)_x = a \neq 0$ or $(f_1)_y = b \neq 0$, hence $(0,0)$ is a simple point of f.

"$\Leftarrow$": Let $f_1 = ax + by$ the degree 1 homogeneous polynomial in the term expansion of f. We write $f = f_1 + g$ where $\deg\ g > 1$. Given a term t in g we thus have that $\deg\ t_x \geq 1$ resp. $\deg\ t_y \geq 1$ or $t_x = 0$ resp. $t_y = 0$. Thus any term in g_x and g_y is divisible by x or y. Then $a = f_x(0,0) \neq 0$ or $b = f_y(0,0) \neq 0$, hence $m_0(f) = 1$. Writing $f_1 = ax + by \in K[x,y]$, we find that the degree 0 homogeneous polynomial terms in f_x is exactly $(f_1)_x = a$ and $(f_1)_y = b$, hence $f_x(0,0)x + f_y(0,0)y = ax + by = f_1$. □

Remark 4.4.9. In the case $m_0(f) > 1$, f_m has multiple tangent lines corresponding to those obtained from its factorization into linear forms (cf Corollary 2.9.110). Each of these tangent lines have a multiplicity in the factorization of f_m. If the multiplicity of each tangent line is 1 and $m > 1$ we call $(0,0)$ an *ordinary multiple point* of f. An ordinary double point is called a *node*.

Lemma 4.4.10. *Consider a curve and its factorization into irreducible factors $f = \prod_1^r p_i^{e_i}$. Then $m_0(f) = \sum_1^m m_0(p_i)e_i$.*

Proof. Indeed, write $p_i = \sum_{m_i}^{n_i} p_{ij}$, as the representation of p_i as a sum of homogeneous polynomials where $p_{im_i} \neq 0$. We get that

$$\sum_m^n f_i = f = \prod_1^m p_i^{e_i} = \prod_1^m \left(\sum_{m_i}^{n_i} p_{ij}\right)^{e_i} = \prod_1^m p_{im_i}^{e_i} + \underbrace{\cdots}_{\text{strictly higher degree terms}} \Rightarrow f_m = \prod_1^m p_{im_i}^{e_i},$$

hence $m_0(f) = m = \deg \prod_1^m p_{im_i}^{e_i} = \sum_1^m m_i e_i = \sum_1^m m_0(p_i)e_i$ □

Definition 4.4.11. Let $P = (a,b) \in \mathbb{A}^2 \setminus 0$. Let $T_P : \mathbb{A}^2 \to \mathbb{A}^2$ be translation by P. We extend all above to definitions pertaining to a $(0,0)$ and a curve f to involve an arbitrary point P by considering $f^T = f(x+a, y+b)$. Then $m_P(f) := m_0(f^T)$. Writing $f^T = \sum_{m_P(f)}^n g_i$. We define the tangent lines at P to be $l_i^{T^{-1}}$, where l_i are the linear factors of $g_{m_P(f)}$. Moreover, we assert that $m_P(0) = \infty$.

Remark 4.4.12. Note that $f_x(P) = f_x^T(P)$ and $f_y(P) = f_y^T(0)$, so the choice of definition of tangent line is sensible, since then P is simple if and only if $m_P(f) = 1$ and in this case the tangent line at P is exactly $g_{m_P(f)}^{T^{-1}}$.

Example 4.4.13. Show some examples of multiple points and tangent lines for a few different points for a few different curves

Proposition 4.4.14. *If a curve f of degree d has a point $P \in f$ of multiplicity d. Then*

$$f = \prod_1^d l_i,$$

where l_i are (possibly non-distinct) lines.

Proof. Suppose $P=(a,b)$ is a point of multiplicity d. Then $f(x+a,y+b)$ is a non-zero homogeneous polynomial of degree d, since the lowest degree term in the homogeneous polynomial expansion of $f(x+a,x+b)$ is d (as $m_P(f)=d$) and $\deg f(x+a,x+b)\le \deg f=d$. It follows from Corollary 2.9.110 that $f(x+a,x+b)=\prod_1^d l_i$, where l_i are lines and hence $f=f(x+a,x+b)^{T^{-1}}=\prod_1^d l_i(x-a,x-b)$, note that $l_i(x-a,x-b)$ are lines, since otherwise $f=0$. ☐

Proposition 4.4.15. *Let P be a double point on a curve f. P is a node if and only if*

$$f_{xy}(P)^2\neq f_{xx}(P)f_{yy}(P).$$

Proof. By Lemma 2.9.146 it is sufficient to prove the result for $P=(0,0)$. We translate the statement into one of linear algebra, by noting that since f has a double point, $f_2=l_1l_2$ where $l_1=ax+by$, $l_2=cx+dy$. Furthermore, the constant term of f_{xx},f_{yy},f_{xy} is respectively $(l_1l_2)_{xx}=2ac,(l_1l_2)_{yy}=2bd$ and $(l_1l_2)_{xy}=ad+bc$. We thus notice that

$$f_{xy}(P)^2-f_{xx}(P)f_{yy}(P)=(ad+bc)^2-4adbc=(ad-bc)^2=\left(\det\begin{pmatrix}a&b\\c&d\end{pmatrix}\right)^2.$$

Note that l_1 and l_2 are distinct (i.e. not factor equivalent) if and only if (a,b) and (c,d) are linearly independent. Hence P is a node if and only if

$$\det\begin{pmatrix}a&b\\c&d\end{pmatrix}\neq 0\iff f_{xy}(P)^2-f_{xx}(P)f_{yy}(P)\neq 0.$$

☐

Proposition 4.4.16. *Suppose* $\operatorname{char} K=0$. *For a point P on a curve f, $m_P(f)$ is the smallest integer such that for some i,j with $i+j=m_P(f)$,*

$$\left(\frac{\partial^{m_P(f)}}{\partial x^i\partial y^j}f\right)(P)\neq 0.$$

The lowest degree homogeneous polynomial of f at P in terms of these partial derivatives in the following way

$$f_{m_P}=\sum_{i,j:i+j=m}\frac{1}{i!j!}\left(\frac{\partial^m}{\partial x^i\partial y^j}f\right)(P)x^iy^j.$$

Proof. Again we may assume WLOG that $P=(0,0)$, since it again follows from Lemma 2.9.146 that $\left(\frac{\partial^{m_P(f)}}{\partial x^i\partial y^j}f\right)(P)=\left(\frac{\partial^{m_P(f)}}{\partial x^i\partial y^j}f^{T_P}\right)(0,0)$. We prove the result by induction in the multiplicity at P. (we define the 0'th partial derivative at any variable to be the identity map). If $m_P(f)=0$, then $P \notin f$, hence $f(P)\neq 0$. If this is isn't satisfactory, when $m_P(f)=0$, then $f_x(P)\neq 0$ or $f_y(P)\neq 0$. Suppose the statement is true for every polynomial of multiplicity $m\geq 0$. Consider a curve f of multiplicity $m+1$ at P. Then f_x or f_y is curve of multiplicity m and the result follows by applying the induction hypothesis. Note here that we implicitly use that $\mathbf{char}\, K=0$. Indeed, from this assumption it follows that $\deg((f_{m+1}))_x=m$ or $\deg\,(f_{m+1})_y=m$. An alternative way (perhaps a more algorithmic approach) is to write $f=\sum_m^n f_i$ in the homogeneous polynomial expansion. Then a term in f_m is of the form $a_{ij}x^iy^j$ where $i+j=m$. Then $\frac{\partial^m}{\partial x^i y^j}a_{ij}x^iy^j=a_{ij}j!i!$ and $\frac{\partial^m}{x^k y^l}a_{k,l}x^ky^l=0$, meaning $\left(\frac{\partial^m}{\partial x^i\partial y^j}f\right)(P)=i!j!a_{ij}$. Hence,

$$f_m=\sum_{i,j:i+j=m}\frac{1}{i!j!}\left(\frac{\partial^m}{\partial x^i\partial y^j}f\right)(P)x^iy^j.$$

□

Proposition 4.4.17. *Let $l_1,\dots,l_n$ be lines all vanishing at $(0,0)$ in $\mathbb{A}^2$ and $r_1,\dots,r_n$ a sequence of positive integers. Set $m:=\sum_1^n r_i$ and $f_m=\prod_1^n l_i^{r_i}$, pick a form f_{m+1} of degree $m+1$ such that $\gcd(f_m,f_{m+1})=1$. Then f_m+f_{m+1} is an irreducible curve with tangent lines $l_1,\dots,l_n$ with respective multiplicities r_i*

Proof. This follows immediately from Proposition 2.9.111. □

Definition 4.4.18. For a polynomial map $\varphi=(\varphi_1,\dots,\varphi_m):\mathbb{A}^n\to\mathbb{A}^m$ and a point $P\in\mathbb{A}^n$ we define the *jacobian of φ at P* to be the matrix

$$J_PT:=\left(\frac{\partial\varphi_i}{\partial x_j}(P)\right)=\begin{pmatrix}\frac{\partial\varphi_1}{\partial x_1}(P) & \cdots & \frac{\partial\varphi_1}{\partial x_n}(P)\\ \vdots & \ddots & \vdots\\ \frac{\partial\varphi_m}{\partial x_1}(P) & \cdots & \frac{\partial\varphi_m}{\partial x_n}(P)\end{pmatrix}\in M_{m\times n}(K)$$

Proposition 4.4.19. *Let $\varphi=(\varphi_1,\varphi_2):\mathbb{A}^2\to\mathbb{A}^2$ be a polynomial map such that f^φ is non-constant. Let f be a curve and Q a point. Set $P=\varphi(Q)$. We then have the following:*

1. *$m_Q(f^\varphi)\geq m_P(f)$.*
2. *the above is true with equality if $J_Q\varphi$ is invertible.*

Proof. 1. We consider first the case $Q=P=(0,0)$. Then $\varphi_i(0,0)=0$ for $i=1,2$, hence the constant term of φ_i is 0 for $i=1,2$. Then the degree of any monomial under composition with φ does not decrease. Thus writing $f=\sum_m^n$ and $f^\varphi=\sum_l^k g_i$ as sums of homogeneous polynomials we have that $m_0(f^{\varphi}) = \deg g_l \geq \deg f_m = m_0(f)$. Note that any translation T is a polynomial such that h^T is a non-constant whenever h is non-constant. Suppose now φ is an arbitrary polynomial map restricted to the assumption. Let $\varphi' = T_P^{-1}\varphi T_Q$ Note that $\varphi'(0,0) = (T_P^{-1}\varphi T_Q)(0,0) = T_P^{-1}(P) = (0,0)$, hence

$$m_Q(f^{\varphi}) = m_0(f^{\varphi T_Q}) = m_0(f^{T_P T_P^{-1}\varphi T_Q}) = m_0((f^{T_P})^{\varphi'}) \geq m_0(f^{T_P}) = m_P(f).$$

2. Consider again first the case of $Q = P = (0,0)$. We have that $\frac{\partial\varphi_1}{\partial x}(Q), \frac{\partial\varphi_2}{\partial y}(Q) \neq 0$ or $\frac{\partial\varphi_1}{\partial y}(Q), \frac{\partial\varphi_2}{\partial x}(Q) \neq 0$. Meaning the jacobian of φ at $(0,0)$ is given by

$$\begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix},$$

where $\lambda_1 = \alpha_1 x + \beta_1 y$ and $\lambda_2 = \alpha_2 x + \beta_2 y$ are the linear homogeneous forms of φ_1 resp. φ_2 such that $\alpha_1, \beta_2 \neq 0$ or $\alpha_2, \beta_1 \neq 0$. There is a natural invertible, linear map on linear forms in 2 variables given by

$$J_0\varphi : V(1,2) \to V(1,2)$$
$$l \mapsto l(\alpha_1 x + \beta_1 y, \alpha_2 x + \beta_2 y).$$

In particular we thus have that $J_0\varphi(l) = 0 \iff l = 0$. Write $\varphi_i = \lambda_i + \mu_i$, where μ_i is 0 or of degree>1. By Corollary 2.9.110 $f_i = \prod_1^{m_0(f)+i} l_j$ for suitable (non-zero!) linear forms $l_j := c_{1j}x + c_{2j}y$. Then

$$l_j(\varphi_1, \varphi_2) = \underbrace{J_0\varphi(l_j)}_{=:l'_j \neq 0} + c_{1j}\mu_1 + c_{2j}\mu_2.$$

We thus get that

$$f_i^{\varphi} = \prod_1^{m_0(f)+i} l'_j + h$$

where h is 0 or a polynomial of degree>1. It thus follows that the lowest degree term of f^{φ} is equal to the lowest degree term $f_{m_0(f)}$, which by the above (maybe overly thorough) considerations (that therefor should be reorganized and rewritten in smaller lemmas at a later point) that $m_0(f) = m_0(f^{\varphi})$. Let P, Q be arbitrary, and define φ' as in 1. It is easy to check that

$$J_v T_w = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

for any $v, w \in \mathbb{A}^n$ By the chain rule for polynomial maps we find that

$$J_{(0,0)}\varphi' = J_P T_P^{-1}\varphi T_Q = (J_{\varphi(Q)} T_P^{-1})(J_Q\varphi)(J_{(0,0)}T_Q) = J_Q\varphi.$$

Hence $J_{(0,0)}\varphi'$ is invertible, meaning

$$m_Q(f^{\varphi}) = m_0((f^{T_P}))^{\varphi'}) = m_0(f^{T_P}) = m_P(f).$$

□

Example 4.4.20. The converse of 2. in the above proposition is not true in general. Consider $\varphi = (x^2, y)$, $f = y - x^2$, $P = Q = (0,0)$. We note that $f^{\varphi} = y - x^4$, hence $m_0(f) = m_0(f^{\varphi})$. But the jacobian of φ at $(0,0)$ is equal to

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

This result maybe is not in the right place

Lemma 4.4.21. *Let* $\operatorname{char} K = p > 0$ *(and* K *algebraically closed). If* $f \in K[x_1, \ldots, x_n]$ *is non-constant with* $f_{x_i} = 0$ *for each* i*, then* $f = h^p$ *for some polynomial* h*. In particular* f *is not irreducible.*

Proof. By lemma 2.9.148 $f = g(x_1^p, \ldots, x_n^p)$ for some $g = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v \in K[x_1, \ldots, x_n]$. Then set $h = \sum_{v \in \mathbb{N}^n : a_v \neq 0} {a_v^{1/p}}^p \mathbf{x}^v$. Then using Freshman's dream we get that

$$h^p = \left(\sum_{v \in \mathbb{N}^n : a_v \neq 0} a_v^{1/p} \mathbf{x}^v \right)^p = \sum_{v \in \mathbb{N}^n : a_v \neq 0} a_v \mathbf{x}^{(pv_1, \ldots, pv_n)} = g(x_1^p, \ldots, x_n^p) = f.$$

□

Proposition 4.4.22. *An irreducible curve* f *has only finitely many multiple points.*

Proof. A point of f is multiple at a point P if and only if $f_x(P) = f_y(P) = 0$. It is thus sufficient to prove that $V(f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$ is finite. It is again sufficient to prove that $\gcd(f, f_x) = 1$ or due to Theorem 4.2.13. We prove the result in the $\operatorname{char} K = 0$ case first. WLOG f_x is non-constant and $\deg f_x < \deg f$, hence $f \nmid f_x$, meaning $\gcd(f, f_x) = 1$, proving the statement in this case. For the $\operatorname{char} K = p > 0$, we use the assumption that f is irreducible in conjunction with the contrapositive of the above lemma to see that f_x and f_y cannot both be identically 0. Hence WLOG $f_x \neq 0$, hence $\deg f_x < \deg f$, meaning $\gcd(f_x, f) = 1$. □

4.4.1 Aside on Hypersurfaces and Tangent Spaces

Definition 4.4.23. Let $f \in K[x_1, \ldots, x_n]$ be a non-constant polynomial and $P \in \mathbb{A}^n$. We define *the multiplicity of* f *at* P to be the degree of the lowest degree term in f^{T_P}, where $T_P : \mathbb{A}^n \to \mathbb{A}^n$ is translation by P. If $m_P(f) = 1$, then we can write f^{T_p} as a sum of homogeneous polynomial $\sum_1^n f_i$, where $f_1 = \sum_1^n a_i x_i \neq 0$. We define the *tangent hyperplane* at P to be the vanishing set of f_1. A simple point of a hypersurface is a point at which all the partial derivatives vanish.

Remark 4.4.24. It is easy to see that a point is simple if and only if $m_P(f) = 1$.

Example 4.4.25. Let's examine $f = x^2 + y^2 - z^2$ have tangent hyperplanes at $(0,0)$. Can we write $f = l_1 l_2$ for homogeneous planes $l_1 := a_1x+b_1y+c_1y, l_2 = a_2x+b_2y+c_2z$. We have that

$$l_1l_2 = a_1a_2x^2 + b_1b_2y^2 + c_1c_2z^2 + (a_1b_2 + b_1a_2)xy + (a_1c_2 + c_1a_2)xz + (b_1c_2 + c_1b_2)yz.$$

To satisfy the identity we thus have that

$$\begin{cases} a_1a_2 = 1 \\ b_1b_2 = 1 \\ c_1c_2 = -1 \\ a_1b_2 + b_1a_2 = 0 \\ a_1c_2 + c_1a_2 = 0 \\ b_1c_2 + c_1b_2 = 0 \end{cases} \iff \begin{cases} a_1 = \frac{1}{a_2} \\ b_1 = \frac{1}{b_2} \\ c_1 = \frac{-1}{c_1} \\ \frac{a_1}{b_1} + \frac{b_1}{a_1} = 0 \\ -\frac{a_1}{c_1} + \frac{c_1}{a_1} = 0 \\ -\frac{b_1}{c_1} + \frac{c_1}{b_1} = 0 \end{cases} \iff \begin{cases} a_2 = \frac{1}{a_1} \\ b_2 = \frac{1}{b_1} \\ c_2 = \frac{-1}{c_1} \\ \frac{a_1^2+b_1^2}{b_1a_1} = 0 \\ \frac{a_1^2-c_1^2}{a_1c_1} = 0 \\ \frac{b_1^2-c_1^2}{b_1c_1} = 0 \end{cases}.$$

It is thus sufficient to find $(a_1, b_1, c_1) \in V(X^2 + Y^2) \cap V(X^2 - Z^2) \cap V(Y^2 - Z^2)$ such that $a_1, b_1, c_1 \neq 0$. However for any given point in the intersection, $a_1^2 = -b_1^2$ and $c_1^2 = b_1^2$, implying $b_1^2 = c_1^2 = a_1^2 = b_1^2$, hence $c_1 = 0$. It is therefor not possible to write f as a product of linear homogeneous polynomials. The theory, therefor does not extend naively to a theory of the behavior of hypersurfaces.

Example 4.4.26. Proposition 4.4.22 is not true for general irreducible hypersurfaces. Indeed, consider a curve f with a multiple point $P = (a,b)$. Let $g = zf \in K[x,y,z]$. Then $V(g)$ has infinitely many multiple points since (a,b,c) is a multiple point of g for each $c \in K \setminus 0$.

Definition 4.4.27. Let $V \subset \mathbb{A}^n$ be an affine variety, $P \in V$. We define *the tangent space of V at P* to the linear subspace

$$T_P(V) := \{v \in \mathbb{A}^n : \forall g \in I(V), (\nabla g)(P) \cdot v = 0\}$$

Lemma 4.4.28. *Let $f \in K[x_1, \ldots, x_n]$ be irreducible, $P \in \mathbb{A}^n$, set $V := V(f)$. Then*

$$T_P(V) = \{v \in \mathbb{A}^n : (\nabla f)(P) \cdot v = 0\}.$$

Proof. If f is irreducible $I(V)=\langle f\rangle$ by HNS. It is clear that $T_P(V)\subset\{v\in\mathbb{A}^n:(\nabla f)(P)\cdot v=0\}$. Note

$$(\nabla gf)_i=(gf)_{x_i}=gf_{x_i}+g_{x_i}f\Rightarrow(\nabla gf)_i(P)=g(P)f_{x_i}(P)\quad\forall g\in K[\mathbf{x}],P\in V.$$

Therefor, if $v\in\mathbb{A}^n$ such that $(\nabla f)(P)\cdot v=0$, then $(\nabla h)(P)\cdot v=0$ for any $h\in\langle f\rangle$, implying the converse inclusion. □

Remark 4.4.29. One readily sees that if P is a multiple point, then $T_P(V)=\mathbb{A}^n$. If P is simple then $\dim\ T_P(V)<n$ it seems that one show that for simple points $\dim\ T_P(V)=\dim\ V$. The tools for this is not developed at this point in fultons curves, so let's leave it

4.4.2 Multiplicities & Local Rings of Rational Function

For a pair of a point and a curve $P\in f$ and $g\in K[x,y]$, $\overline{g}:=g+I(f)\in\Gamma(f)\subset\mathcal{O}_P(f)$.

Lemma 4.4.30. *Let $P\in f$ where f is an irreducible curve. If P is a simple point of f, then $\mathcal{O}_P(f)$ is a DVR.*

Proof. WLOG $P=(0,0)$. Since An affine transformation has invertible jacobian, hence by Proposition 4.4.19 we don't lose generality either by assuming the tangent line of f at P is y (recall that the tangent line has (affine) dimension 1). Then $l=x$ is a line passing through P which is not tangent to f at P. If we can prove that $\mathfrak{m}_P(f)=\langle\overline{x}\rangle$ in $\mathcal{O}_P(f)$, then we are done by definition of a DVR, since the local ring of any non-trivial variety at any point is a non-field integral domain that is Noetherian and local (cf. Corollary 4.3.58). By Proposition 4.3.67 $\mathfrak{m}_P(\mathbb{A}^2)=\mathcal{O}_P(\mathbb{A}^2)\langle x,y\rangle$. By Propostion 4.3.68, $\mathcal{O}_P(\mathbb{A}^2)/\langle f\rangle\mathcal{O}_P(\mathbb{A}^2)\simeq\mathcal{O}_P(f)$, hence $\mathfrak{m}_P(f)=\overline{\mathfrak{m}_P(\mathbb{A}^2)}=\langle\overline{x},\overline{y}\rangle$. Now write $f=gy-hx^2$, where $g=1+g'$, where $g'=0$ or $\deg\ g'>0$ (we can do this since y is the linear term of f) and where $h\in K[x]$. Since g has constant term 1, we note that $g(P)\neq 0$. Note also that we obtain a factor of x^2 by the fact that no linear term of f is divisible by x. Then

$$\overline{yg}=\overline{x^2h}\Rightarrow\overline{y}=\frac{\overline{x^2h}}{\overline{g}}\in\langle\overline{x}\rangle\Rightarrow\mathfrak{m}_P(f)=\langle\overline{x},\overline{y}\rangle=\langle\overline{x}\rangle.$$

□

Remark 4.4.31. To spell out the assumption of $P=(0,0)$. Note that f is simple at P if and only if f^{T_P} is simple at $(0,0)$ and that $\mathcal{O}_{(0,0)}(f^{T_P})\simeq\mathcal{O}_P(f)$ by Corollary 4.3.65, since $f^{T_P}\overset{T_P^{-1}}{\simeq}f$.

Definition 4.4.32. Let P be a simple point on an irreducible curve f. Then $\mathbf{ord}_P^f$ is the order function on $K(f)$ induced by the DVR $\mathcal{O}_P(f)$.

Remark 4.4.33. It readily follows that $\mathbf{ord}_P^f \circ \varphi^{-1} = \mathbf{ord}_0^{f^\varphi}$, where φ is the affine change of coordinates mapping P to 0 and the tangent of f at $(0,0)$ to Y, again since $\mathcal{O}_{(0,0)}(f^\varphi) \simeq \mathcal{O}_P(f)$ Perhaps add little diagram. Note then that given a non-tangent line l of f passing through P we have that $\mathbf{ord}_P^f(\bar{l}) = \mathbf{ord}_P^f(\varphi^{-1}(\overline{x+by})) = \mathbf{ord}_0^{f^\varphi}(\overline{x+by}) = 1$. If l is tangent to f at P, then

$$\mathbf{ord}_P^f(\bar{l}) = \mathbf{ord}_0^{f^\varphi}(\bar{y}) = \mathbf{ord}_0^{f^\varphi}(\bar{x}^2) + \mathbf{ord}_0^{f^\varphi}\left(\frac{\bar{h}}{\bar{g}}\right) \geq 2 > 1.$$

Theorem 4.4.34. *Let $P \in f$ where f is an irreducible curve. Then for sufficiently large n,*

$$m_P(f) = \mathbf{dim}\ \mathfrak{m}_P(f)^n/\mathfrak{m}_P(f)^{n+1},$$

where we recall that $\mathfrak{m}_P(f)$ is the maximal ideal of $\mathcal{O}_P(f)$.

Proof. WLOG $P = (0,0)$.Set $R := \mathcal{O}_P(f)$ and $\mathfrak{m} := \mathfrak{m}_P(f)$. For each n we get an exact sequence

$$0 \longrightarrow \mathfrak{m}^n/\mathfrak{m}^{n+1} \longrightarrow R/\mathfrak{m}^{n+1} \longrightarrow R/\mathfrak{m}^n \longrightarrow 0$$

We thus note that

$$\mathbf{dim}\ \mathfrak{m}^n/\mathfrak{m}^{n+1} = \mathbf{dim}\ R/\mathfrak{m}^{n+1} - \mathbf{dim}\ R/\mathfrak{m}^n,$$

hence if we can prove that $\mathbf{dim}\ R/\mathfrak{m}^n = nm_P(f)+s$ for some $s \in \mathbb{Z}$ for every $n \geq m_P(f)$ we get for such n that

$$\mathbf{dim}\ \mathfrak{m}^n/\mathfrak{m}^{n+1} = (n+1)m_P(f) + s - (nm_P(f) + s) = m_P(f).$$

Set $I := \langle x, y\rangle$. Fix $n \geq m_P(f)$. Then $\mathfrak{m}^n = I^n R$ by Proposition 4.3.67. $V(I^n) = \{P\}$, hence

$$K[x,y]/\langle I^n, f\rangle \simeq \mathcal{O}_P(\mathbb{A}^2)/(I^n, f)\mathcal{O}_P(\mathbb{A}^2) \simeq R/I^n R \simeq R/\mathfrak{m}^n$$

by Theorem 4.3.75, Proposition 4.3.68. We will be using the fact that

$$K[x,y]/I^d \simeq K_{\leq d-1}[x,y],$$

repeatedly ref. Recall that $\mathbf{dim}\ K_{\leq d-1}[x,y] = \frac{d(d+1)}{2}$ (cf. Example 2.9.120. Since the lowest degree term of f is of degree $m := m_P(f)$ we have that $f \in I^m$. Hence for

$g \in K[x,y]$, $g \in I^{n-m}$ if and only if $fg \in I^n$. Let $\sigma : K[x,y]/I^n \twoheadrightarrow K[x,y]/\langle f, I^n\rangle$ and $\tau : K[x,y]/I^{n-m} \hookrightarrow K[x,y]/I^n, g + I^{n-m} \mapsto fg + I^n$. We then get an exact sequence

$$0 \longrightarrow K[x,y]/I^{n-m} \xhookrightarrow{\tau} K[x,y]/I^n \xrightarrow{\sigma} K[x,y]/\langle f, I^n\rangle \longrightarrow 0$$

Then

$$\begin{aligned}
\dim\ K[x,y]/\langle f, I^n\rangle = \dim\ K[x,y]/I^n - \dim\ K[x,y]/I^{n-m} &= \frac{n(n+1)}{2} - \frac{(n-m)(n-m+1)}{2} \\
&= \frac{n(n+1)}{2} - \frac{n^2 - nm + n - nm + m^2 + m}{2} = \frac{n(n+1)}{2} + nm - \frac{n(n+1)}{2} - \frac{m(m+1)}{2} \\
&= nm - \underbrace{\frac{m(m+1)}{2}}_{s}.
\end{aligned}$$

□

Remark 4.4.35. Consider an arbitrary curve f with components $f_1, \dots, f_m$ such that $f = \prod_i^m f_i^{e_i}$. Then the multiplicity f at a point $P \in f$ depends only on the local rings of the components at P and the positive integers $e_1, \dots, e_m$.

Theorem 4.4.36. *Let $P \in f$ where f is curve. P is a simple point of f if and only if $\mathcal{O}_P(f)$ is a DVR. In this case, if $l = ax + by + c$ is a line through P such that l is not tangent to f at P, then $l + \Gamma(f) \in \mathcal{O}_P(f)$ is a uniformizing parameter of $\mathcal{O}_p(f)$.*

Proof. The first implication follows from Lemma 4.4.30. If $\mathcal{O}_P(f)$ is a DVR, then by ref and the prior theorem for sufficiently large n,

$$m_P(f) = \dim\ \mathfrak{m}_P(f)^n / \mathfrak{m}_P(f)^{n+1} = 1.$$

The remaining statement was already proven in Lemma 4.4.30. □

Remark 4.4.37. For a local ring R, the function

$$\begin{aligned}
\chi_R : \mathbb{N} &\to \mathbb{N} \\
n &\mapsto \dim\ R/\mathfrak{m}
\end{aligned}$$

is called the *Hilbert-Samuel polynomial of R*.

Example 4.4.38. 1. Let $R = \mathcal{O}_P(\mathbb{A}^2)$, $P \in \mathbb{A}^2$. Then $\chi_R(n) = \dim\ R/\mathfrak{m}^n = \dim\ K[x,y]/\langle x,y\rangle^n = \frac{n(n+1)}{2}$ for each n.

2. We determine χ for $R := \mathcal{O}_P(\mathbb{A}^k)$ at a point $P \in \mathbb{A}^k$. We then have that

$$\chi_R(n) = \dim\ R/\mathfrak{m}^n = \dim\ K[x_1,\dots,x_k]/\langle x_1,\dots,x_k\rangle^n = \dim\ K_{\leq n-1}[x_1,\dots,x_k] = \binom{k+n-1}{k}$$

$$= \frac{(n-1+k)!}{(n-1)!k!} = \frac{1}{k!}\prod_{h=n-1}^{n-1+k} h = \frac{1}{k!}n^k + \cdots$$

The n'th Hilbert-Samuel polynomial for R is thus a polynomial over $\mathbb{Q}$ of degree k with leading coefficient $\frac{1}{r!}$.

Proposition 4.4.39. *Let $f \in K[x_1,\dots,x_k]$ define an irreducible hypersurface containing $P = 0 \in \mathbb{A}^k$. Let $f_{m_P(f)}$ be the polynomial of lowest degree terms in f. Set $R := \mathcal{O}_P(V(f))$ and denote its maximal ideal by $\mathfrak{m}$. For sufficiently large n, $\chi(n)$ is a polynomial n of degree $k-1$ with leading coefficient $m_P(f)/(k-1)!$*

Proof. DO AT SOME POINT □

Remark 4.4.40. Given a local ring R we are enclined to define $m(R)$ (the *Hilbert-Samuel multiplicity of R*) as $\chi(1)$.

Definition 4.4.41. A point simple point P on a curve f is called a *flex* if $\mathrm{ord}_P^f(l) \geq 3$, where l is tangent of f at P. It is called *ordinary* if the inequality is true with equality, otherwise it is called *higher*.

Example 4.4.42. 1. Let $f_n = y - x^n$, $n \geq 1$ and $P = (0,0)$.

For $n = 1$, $l_1 = y - x$ is tangent to f_1 at P. x is a non-tangent line to f_1 at P. Hence

$$\overline{l_1} = 0 \Rightarrow \mathrm{ord}_P^{f_1}(\overline{l_1}) = \infty,$$

and hence P is a higher flex.

For $n \geq 2$, $l_n = y$ is the tangent of f_n at P, hence x is a non-tangent to f_n at P, meaning $\overline{x}$ is a uniformizing parameter of $\mathcal{O}_P(f_n)$. We thus get that

$$\mathrm{ord}_P^{f_n}(\overline{l}) = \mathrm{ord}_P^{f_n}(\overline{x}^n) = n,$$

so for $n = 2$, P is not a flex, for $n = 3$, P is an ordinary flex and for $n \geq 4$, P is a higher flex.

2. Let $f = y - \sum_2^n a_i x^i$, where the polynomial in x is non-zero. Then $P = (0,0)$ is a flex of P if and only if $a_2 = 0$. Indeed, y is the tangent of f at P and $\overline{x}$ is a uniformizing parameter of $\mathcal{O}_P(f)$, hence

$$\mathrm{ord}_P^f(\overline{y}) = \mathrm{ord}_P^f\left(\sum_2^n a_i \overline{x}^i\right) = \min(\{i : a_i \neq 0\}).$$

Hence if $a_2 = 0$, $\mathrm{ord}_P^f(\overline{y}) \geq 3$; otherwise $\mathrm{ord}_P^f(\overline{y}) = 2$.

Proposition 4.4.43. *Let P be a point on an irreducible curve f. Then*

$$\dim\ \mathfrak{m}_P(f)^n/\mathfrak{m}_P(f)^{n+1} = n+1, \quad (0 \le n < m_P(f)),$$

hence P is simple if and only if $\dim\ \mathfrak{m}_P(f)/\mathfrak{m}_P(f)^2 = 1$*; otherwise* $\dim\ \mathfrak{m}_P(f)/\mathfrak{m}_P(f)^2 = 2$.

Proof. With the notation used in Theorem 4.4.34 we see that $R/\mathfrak{m}^k \simeq K[x,y]/\langle f, I^k\rangle$ for each $k \ge 0$. In particular for $k \le m$, $\langle f, I^k\rangle = I^k$ since $f \in I^m$, hence $R/\mathfrak{m}^k \simeq K[x,y]/I^k$. For each $n \in \{0,\ldots,m-1\}$ we get an exact sequence

$$0 \longrightarrow \mathfrak{m}^n/\mathfrak{m}^{n+1} \longrightarrow R/\mathfrak{m}^{n+1} \longrightarrow R/\mathfrak{m}^n \longrightarrow 0$$

Hence,

$$\dim\ \mathfrak{m}^n/\mathfrak{m}^{n+1} = \dim\ R/\mathfrak{m}^{n+1} - \dim\ R/\mathfrak{m}^n = \frac{(n+1)(n+2)}{2} - \frac{n(n+1)}{2} = \frac{(n+1)(n+2-n)}{2} = n+1.$$

□

Example 4.4.44. Set $f_1 := x^2 - y^3, f_2 := y^2 - z^3 \in K[x,y,z]$, $P = (0,0,0)$ and $V := V(f_1,f_2) \subset \mathbb{A}^3$. Determine $\dim\ \mathfrak{m}_P(V)/\mathfrak{m}_P(V)^2$.

4.4.3 Intersection Numbers

For this subsection, we begin by fixing curves f and g together with a point $P \in \mathbb{A}^2$. We allow that f or g be non-zero constants. We will see that allowing this will make definitions simpler and that it will inconsequential for counting intersections of curves.

Definition 4.4.45. f and g *intersect properly* at P, if f and g have no common components passing through P. I.e. for f and g to intersect properly at P we require that for any common factor h of f and g, $h(P) \neq 0$

Definition 4.4.46. If P is simple for both f and g such that f and g have distinct tangents at P we say that f and g intersect *transversally* at P.

Definition 4.4.47. An *intersection number of f and g at P* is a value $I(P, f \cap g) \in \mathbb{Z} \cup \{\infty\}$ satisfying the following 7 axioms:

1. When f and g intersect properly at P, $I(P, f\cap g) \ge 0$; otherwise $I(P, f \cap g) = \infty$.

2. $I(P,f\cap g)=0 \iff P\notin F\cap G$ and $I(P,f\cap g)$ depends only on the components of f and g passing through P, i.e if $f_1,\dots,f_m$ and $g_1\dots,g_l$ are these components, then

$$I(P,f\cap g)=I\left(P,\prod f_i\cap\prod g_i\right)$$

3. If T is an affine change of coordinates on $\mathbb{A}^2$ with $T(Q)=P$, then

$$I(P,f\cap g)=I(Q,f^T\cap g^T).$$

4. $I(P,f\cap g)=I(P,g\cap f)$.

5. $I(P,f\cap g)\geq m_P(f)m_P(g)$ with equality holding if and only if f and g have no tangents in common. In particular $I(P,f\cap g)=1$ if and only if f and g intersect transversally.

6. If $f=\prod_1^m f_i^{r_i}$ and $g=\prod_1^l g_i^{s_i}$, then

$$I(P,f\cap g)=\sum_1^m\sum_1^l r_i s_j I(P,f_i\cap g_j).$$

7. $I(P,f\cap g)$ depends only on the image of g in $\Gamma(f)$, i.e. for every $h\in K[x,y]$,

$$I(P,f\cap g)=I(P,f\cap g+hf).$$

The intuition for the above definition is that we want to be able count the number of intersections f and g. We detect an intersection if $P\in f\cap g$. We want this way of counting to be sensitive to the fact that intersections may occur with multiplicity. For example if P vanishes on two components of f and on one component of g, then an intersection in a sense actually occurs twice. To be more concrete take $f=x$ and $g=y^2-x$. Then g intersects f only at $(0,0)$, however $g(0,y)$ has a double root at $(0,0)$, so this intersection should really be counted twice. This is indeed the case: $I((0,0),f\cap g)=I((0,0),f\cap g+f)=I((0,0),x\cap y^2=2I((0,0),x\cap y)=2$. We want to describe this number in terms of rational functions on the affine plane at each P. To do this we need some technical lemmas.

Lemma 4.4.48. *Set $m:=m_P(f)$ and $n:=m_P(g)$. Suppose f and g intersect properly, have no common tangents and pass through P. Then for $I:=\langle x,y\rangle$, we have*

$$I^t\subset\langle f,g\rangle\mathcal{O}_P(\mathbb{A}^2),$$

For sufficiently large $t\geq 0$.

Proof. Note that $\mathbf{gcd}(f,g) = 0$, hence $f \cap g = \{Q_1, \dots, Q_k, P\}$. Then there is an $h \in K[x,y]$ such that $h(Q_i) = 0$ and $h(P) = 1$. Then $xh, yh \in I(f \cap g)$, hence for a suitably large N, $(xh)^N, (yh)^N \in \langle f, g\rangle$ (cf. HNS). In $\mathcal{O}_P(\mathbb{A}^2)$, h is a unit, since it doesn't pass through P, hence $x^N, y^N \in \langle f, g\rangle$, implying the result for $t \geq 2N$ □

Lemma 4.4.49. *With the same setup as the above lemma we have that*

$$I^t \subset \langle f, g\rangle \mathcal{O}_P(\mathbb{A}^2),$$

whenever $t \geq m + n - 1$

Proof. Let $l_1, \dots, l_m$ and $\lambda_1, \dots, \lambda_n$ be the tangents of f resp. g. for $i > m$ define $l_i := l_m$ and for $j > n$ define $\lambda_j := \lambda_n$ Define $a_{ij} := \left(\prod_1^i l_i\right)\left(\prod_1^j \lambda_i\right)$ for every $i \in \{0, \dots, m\}$ and $j \in \{0, \dots, n\}$. Then $\{a_{ij} : i + j = t\}$ forms a basis for $V_K(t, 2) = I^t$ by do exercise!. Hence it is sufficient to prove that $a_{ij} \in \langle f, g\rangle \mathcal{O}_P(\mathbb{A}^2)$ for every i, j with $i + j \geq m + n - 1$. If we are given i, j such that $i + j \geq m + n - 1$. Then $i \geq m$ or $j \geq n$. By the lemma above there is a $T \geq 0$ such that $I^T \subset \langle f, g\rangle$. Hence if we can prove that $a_{ij} \in \langle f, r\rangle$ where $\mathbf{deg}\ r \geq T$ in the case where $i \geq m$ and symmetrically that $a_{ij} \in \langle g, s\rangle$ when $j \geq n$ we are done. Indeed in the case $i \geq m$ we would have that $a_{ij} = bf + cr \in \langle f, g\rangle$. The other case follows by an identical argument. Suppose $i \geq m$. Then $a_{ij} = a_{m0} b$ for some homogeneous b of degree $t := i + j - m$. Writing $f = A_{m0} + f'$ where $\mathbf{deg}\ f' \geq m + 1$ we get that

$$a_{ij} = bf - bf'$$

where $bf' = 0$ or the terms of bf' are of degree$\geq i + j + 1$. In the second case write all terms in bf' as an appropriate linear combination of elements in $\{a_{kh} : k + h \geq m'\}$. Then $a_{ij} = cf - ch'$ where $h' = 0$ or has terms of degree $\geq m' + 1$. Repetition of this argument yields an expression $a_{ij} = uf + r$ with $r \in I^T$. □

Lemma 4.4.50. *With the same setting as above, the map* $\rho : K[x,y]/I^m \times K[x,y]/I^n \to K[x,y]/I^{n+m}, (\lambda + I^m, \mu + I^n) \mapsto \lambda f + \mu g + I^{n+m}$ *is injective if and only if* f *and* g *have distinct tangents at* P*.*

Proof. "$\Rightarrow$": If f and g had a common tangent, l say, write $f_m = l f'_{m-1}$ and $g_n = l g'_{n-1}$. Let F and G, denote term of degree$> m$ resp. n. Then $f'_{m-1} \notin I^m$ and $g'_{n-1} \notin I^n$ and

$$g'_{n-1} f + f'_{m-1} g = f'_{m-1} g_n + g'_{n-1} F - f'_{m-1} g_n - f'_{m-1} G = g'_{n-1} F - f'_{m-1} G \in I^{m+n}.$$

implying $\rho(g'_{n-1} + I^m, -f'_{m-1} + I^n) = 0$, which means ρ is not injective. "$\Leftarrow$": Suppose the tangents are distinct. Suppose then that $af + bg \in I^{n+m}$. Suppose for a contradiction that $m_P(a) := r < m$ or $m_P(b) := s < n$. Write $a = a_r + \dots$ and $b = b_s + \dots$.

The terms in $a_r f_m$ and $b_s g_n$ must cancel, hence $r+m = s+n$ and $a_r f_m = -b_s g_n$. Then, since f_m and g_n have no common factors. Then $f_m \mid a_r$ and $g_n \mid b_s$, leading to a contradiction. This means $r \geq m$ and $s \geq n$, hence $a \in I^m$ and $b \in I^n$. $\square$

Remark 4.4.51. We describe an algorithm $J(\_,\_\cap\_)$ for computing intersection numbers given that an intersection number exists for every $P \in \mathbb{A}^2$ and every pair of curves f and g. We first describe it in the special case $P = (0,0)$:

(I) If $P \neq (0,0)$, return $J((0,0), f^{T_P} \cap g^{T_P})$.

(II) If f and g do not intersect properly, then return ∞.

(III) If $P \notin f \cap g$, return 0.

(IV) Set $r := \deg_x f$ and $s := \deg_x g$. If $r \leq s$, proceed, otherwise swap the role of f and g in the next steps

(V) If $r = 0$, write $f = yh$ for a suitable unique $h \in K[x,y]$ and return $m_P(g(x,0)) + J(P, h \cap g)$.

(VI) If $r > 0$ scale f and g such that $f(x,0)$ and $g(x,0)$ become monic, and set $h := g - x^{s-r} f$. Return $J(P,h)$.

We will prove that the algorithm will always terminate in the theorem below.

Theorem 4.4.52. **dim** $\mathcal{O}_P(\mathbb{A}^2)/\langle f,g\rangle$ *is the unique intersection number of* f *and* g *at* P.

Proof. **Uniqueness:** We prove that if there is an intersection number $I(P, f \cap g)$ for each $P \in \mathbb{A}^2$ and each pair of curves f,g, then the algorithm above returns that intersection value. Trivially, if there were two candidates for intersection numbers $I(\_,\_\cap\_)$, $I'(\_,\_\cap\_)$, then

$$I(\_,\_\cap\_) = J(\_,\_\cap\_) = I'(\_,\_\cap\_).$$

By property 3, if $P \neq (0,0)$, then $I(P, f \cap g) = I((0,0), f^{T_P} \cap g^{T_P})$, hence, if the algorithm terminates correctly in the $P = (0,0)$-case so will it in the general case. So assume $P = (0,0)$. If $I(P, f \cap g) = \infty$, then f and g do not intersect properly by property 1, hence $J(P, f \cap g) = \infty$. Suppose $I(P, f \cap g) < \infty$. Suppose $I(P, f \cap g) = 0$. Then $P \notin f \cap g$ by property 2, hence $J(P, f \cap g) = 0$. Suppose $I(P, f \cap g) > 0$. Suppose $J(P, A \cap B) = I(P, A \cap B)$, whenever $I(P, A \cap B) < n$ for some $n \geq 0$. If $I(P, f \cap g) = n$, set $r := \deg_x f$ and $s := \deg_x g$, where WLOG $r \leq s$ and $\deg\ 0 = 0$.

When $\mathbf{r = 0}$, we land in (V). We get that $y \mid f$, since $P = (0,0) \in f$, hence we can write $f = yh$ for a unique $h \in K[x,y]$. Then

$$n = I(P, f \cap g) = I(P, y \cap g) + I(P, h \cap g).$$

Write $g = \lambda + yk$ for some unique $\lambda \in K[x]$ and $k \in K[x,y]$. Note that $\lambda \neq 0$ for otherwise g and f would not intersect properly. We also have that $m_P(\lambda) > 0$ since $\lambda(P) = \mathrm{ev}_P(\lambda + yk) = g(P) = 0$. Then

$$I(P, y \cap g) = I(P, y \cap \lambda + yk) = I(P, y \cap \lambda) = m_P(y)m_P(\lambda) = m_P(\lambda) > 0.$$

For the second equality we use property 7. For the third equality we use property 5. We thus have that

$$I(P, h \cap g) = n - m_P(\lambda) < n.$$

Note that $\lambda = g(x,0)$. Then by induction hypothesis

$$I(P, f \cap g) = m_P(g(x,0)) + I(P, h \cap g) = m_P(g(x,0)) + J(P, h \cap g) = J(P, f \cap g).$$

When $r > 0$ we land in (VI). WLOG f and g are monic. We set $a_0 := \frac{1}{\mathrm{lc}(g)} g$, $b_0 := \frac{1}{\mathrm{lc}(f)} f$ and define

$$h_i := a_{i-1} - x^{\deg_x\ a_{i-1} - \deg_x\ b_{i-1}} b_{i-1}.$$

We pick a_i to be the polynomial with largest x-degree of b_{i-1} and h_i scaled by the inverse of its leading coefficient and let b_i be the polynomial with smaller x-degree of the two scaled by the inverse of it leading coefficient. We thus note that $\deg_x\ h_1 < \deg_x\ g$ and that $\deg_x\ h_i < \deg_x h_{i+1}$. Hence for some $N \geq 1$, $\deg_x\ b_N = 0$ with $\deg_x\ a_N \geq \deg\ b_N$. We repeatedly use property 7 and 4 to find that

$$I(P, f \cap g) = I(P, a_0 \cap b_0) = \ldots = I(P, a_N \cap b_N) = J(P, a_N \cap b_N),$$

where the last equality follows from case $r = 0$. Applying the algorithm to P, f and g clearly yields

$$J(P, f \cap g) = J(P, a_0 \cap b_0) = \ldots = J(P, a_N \cap b_N) = I(P, f \cap g),$$

finishing the proof of uniqueness Check if can be written better later!.

Existence: We first show that property 2 is fulfilled. Note that $\mathcal{O}_P(\mathbb{A}^2)/\langle f, g \rangle = 0$ if and only if f or g is a unit which is equivalent to f or g not passing through P. Note that if $f_1, \ldots, f_m$ and $g_1, \ldots, g_l$ are the components of f resp. g. passing through O, then in $\mathcal{O}_P(\mathbb{A}^2)$, $\langle f, g\rangle = \langle \prod_1^m f_i, \prod_1^l g_i\rangle$ since every other component is a unit in $\mathcal{O}_P(\mathbb{A}^2)$. Property 4 is obviously also fulfilled. As for property 7 note that clearly $\langle f, g\rangle = \langle f, g + hf\rangle$ for every curve h. For an isomorphism φ an affine change of coordinates on $\mathbb{A}^2$ taking Q to P is an isomorphism it follows that

$$\mathcal{O}_P(\mathbb{A}^2)/\langle f, g\rangle \simeq \mathcal{O}_{T(Q)}(\mathbb{A}^2)/\langle f^T, g^T\rangle$$

by Corollary 4.3.66, which shows that property 3 holds. From this point forward we therefor need only consider the case where $P = (0,0)$ and P passes through every component of f and g. We now check that property 1 is satisfied. If f and g have no common components then $f \cap g = \{P_1, \ldots, P_n\}$ is finite, hence if $P_i = (a_i, b_i)$,

$$\dim\ \mathcal{O}_{P_i}(\mathbb{A}^2)/\langle f, g\rangle = \dim\ K[x,y]/\langle f, g\rangle < \infty$$

by Theorem 4.3.75 and a Corollary of HNS. If f and g have a common component h, then $\langle f, g\rangle \subset \langle h\rangle$, hence we get that $\dim\ \mathcal{O}_P(\mathbb{A}^2)/\langle f, g\rangle \geq \mathcal{O}_P(\mathbb{A}^2)/\langle h\rangle$ since we have a K-algebra surjection

$$\mathcal{O}_P(\mathbb{A}^2)/\langle f, g\rangle \longrightarrow\!\!\!\!\!\rightarrow \mathcal{O}_P(\mathbb{A}^2)/\langle h\rangle$$

Note that $\mathcal{O}_P(\mathbb{A}^2)/\langle h\rangle \simeq \mathcal{O}_P(h)$ by Proposition 4.3.68. Note also that $\mathcal{O}_P(h) \supset \Gamma(f)$ and that $\dim\ \Gamma(f) = \infty$ by a corollary of HNS, hence $\dim\ \mathcal{O}_P(\mathbb{A}^2)/\langle f, g\rangle \geq \dim\ \Gamma(f) = \infty$.

We move on to prove that property 6. holds. It is sufficient to prove that $I(P, f \cap gh) = I(P, f \cap g) + I(P, f \cap h)$ since 6 then follows by induction arguments. When f and gh have a common component, the result is trivial. So suppose this is not the case. We get an injective K-algebra map $\mathcal{O}_P(\mathbb{A}^2)/\langle f, g\rangle \to \mathcal{O}_P(\mathbb{A}^2)/\langle f, gh\rangle, \lambda + \langle f, g\rangle \mapsto \lambda h + \langle f, gh\rangle$. It is well-defined K-algebra homomorphism, since if $\lambda = af + bg$, then $h\lambda = afh + bgh \in \langle f, gh\rangle$. Suppose $z \in \mathcal{O}_P(\mathbb{A}^2)$ is given such that $zh \in \langle f, gh\rangle$. Then for suitable $b, c \in \mathcal{O}_P(\mathbb{A}^2)$,

$$hz = bf + cgh$$

For some $\mu \in K[x,y] \setminus 0$, $\mu b, \mu c \in K[x,y]$, hence

$$\mu zh = \mu bf + \mu cgh \Rightarrow \mu bf = (\mu z - \mu cg)h.$$

Then since $\gcd(f, h) = 1$, we get that $f \mid \mu z - \mu cg$ in $K[x,y]$, hence $\mu z - \mu cg = qf$, meaning

$$z = \frac{q}{\mu} f + cg \in \langle f, g\rangle$$

We thus get a short exact sequence

$$0 \longrightarrow \mathcal{O}_P(\mathbb{A}^2)/\langle f,g\rangle \hookrightarrow \mathcal{O}_P(\mathbb{A}^2)/\langle f,gh\rangle \twoheadrightarrow \mathcal{O}_P(\mathbb{A}^2)/\langle f,h\rangle \longrightarrow 0\ ,$$

hence $\dim\ \mathcal{O}_P(\mathbb{A}^2)/\langle f,gh\rangle = \dim\ \mathcal{O}_P(\mathbb{A}^2)/\langle f,g\rangle + \dim\ \mathcal{O}_P(\mathbb{A}^2)/\langle f,h\rangle$ which is what we wanted to show.

We know just need to prove that property 5 holds. Set $m := m_P(f)$, $n := m_P(g)$ and $I = \langle x,y\rangle \subset K[x,y]$. Consider the diagram

$$\begin{tikzcd}
K[x,y]/I^m \times K[x,y]/I^n \arrow[r, "\rho"] & K[x,y]/I^{m+n} \arrow[r, "\sigma"] & K[x,y]/\langle I^{n+m},f,g\rangle \arrow[r] \arrow[d, "\alpha"] & 0 \\
& \mathcal{O}_P(\mathbb{A}^2)/\langle f,g\rangle \arrow[r, "\pi"] & \mathcal{O}_P(\mathbb{A}^2)/\langle I^{n+m},f,g\rangle \arrow[r] & 0
\end{tikzcd}$$

Where $\rho(\lambda+I^m,\mu+I^n) = \lambda f+\mu g+I^{m+n}$, σ and π are the canonical surjections. Note that $V(I^{n+m},f,g) \subset V(I) = \{P\}$, hence α is the isomorphism $K[x,y]/\langle I^{n+m},f,g\rangle \simeq \mathcal{O}_P(\mathbb{A}^2)/\langle I^{n+m},f,g\rangle$ (cf. Theorem 4.3.75). Clearly $\sigma\circ\rho = 0$, hence the top sequence is exact

$$\dim\ K[x,y]/I^m + \dim\ K[x,y]/I^n = \dim\,(\ker\ \rho) + \dim\,(\mathrm{im}\ \rho) = \dim\,(\ker\ \rho) + \dim\,(\ker\ \sigma),$$

hence $\dim\ K[x,y]/I^m + \dim\ K[x,y]/I^n \geq \dim\ (\ker\ \sigma)$, which is true with equality if and only if ρ is injective. Furthermore,

$$\dim\ K[x,y]/I^{m+n} = \dim\ (\mathrm{im}\ \sigma) + \dim\ (\ker\ \sigma) = \dim\ K[x,y]/\langle I^{n+m},f,g\rangle + \dim\ \ker\ \sigma,$$

hence $\dim\ K[x,y]/\langle I^{m+n},f,g\rangle = \dim\ K[x,y]/I^{m+n} - \dim\ \ker\ \sigma$. It thus follows that

$$\begin{aligned}
I(P,f\cap g) &= \dim\ \mathcal{O}_P(\mathbb{A}^2)/\langle f,g\rangle \geq \dim \mathcal{O}_P(\mathbb{A}^2)/\langle I^{m+n},f,g\rangle \\
&= \dim\ K[x,y]/\langle I^{m+n},f,g\rangle = \dim\ K[x,y]/\langle I^{m+n}\rangle - \dim\ (\ker\ \sigma) \\
&\geq \dim\ K[x,y]/\langle I^{m+n}\rangle - \dim\ K[x,y]/I^m - \dim\ K[x,y]/I^n \\
&= mn
\end{aligned}$$

which holds with equality if and only if ρ is injective and π is an isomorphism. The last equality follows from verifying the computation,

$$\frac{(m+n)(m+n+1)}{2} - \frac{m(m+1)}{2} - \frac{n(n+1)}{2} = mn.$$

That π is an isomorphism is ensured by the assumption that f and g have no common tangents due to Lemma 4.4.49. Indeed, $m+n \geq m+n-1$, meaning $\langle I^{n+m},f,g\rangle = \langle f,g\rangle$. That ρ is injective also follows from the assumption that f and g have no common tangents due to Lemma 4.4.50. $\square$

Example 4.4.53. 1. Over $\mathbb{C}$, let $e = (x^2+y^2)^2+3x^2y-y^3$, $f = (x^2+y^2)^3-4x^2y^2$. Then

$$f-e(x^2+y^2) = y(\underbrace{(x^2+y^2)(y^2-3x^2)-4x^2y}_{g}) = yg.$$

Note now that

$$g+3e = y(\underbrace{5x^2-3y^2+4y^3+4x^2y}_{h}) = yh.$$

The tangent lines of e are two copies of y, $(\sqrt{3}x+y)$ and $\sqrt{3}x-y$ and the tangent lines of h are two copies of x, hence e and h have no tangent lines in common. Then for $P=(0,0)$.

$$\begin{aligned}I(P,e\cap f) &= I(P,e\cap f-e(x^2+y^2)) = I(P,e\cap yg) = I(P,e\cap y)+I(P,e\cap g)\\ &= I(P,e\cap y)+I(P,e\cap g+3e) = I(P,e\cap y)+I(P,e\cap yh)\\ &= 2I(P,e\cap y)+I(P,e\cap h) = 2I(P,y\cap e-y^4-2x^2y^2-3x^2y+y^3)+m_P(e)m_P(h)\\ &= 2I(P,y\cap x^4)+6 = 8I(P,y\cap x)+6 = 8+6 = 14.\end{aligned}$$

Note that instead of religiously following the steps of the algorithm presented it is often better to intuitively use the properties of the intersection numbers.

2. Consider $a = y-x^2$, $b = y^2-x^3+x$. Then $I((0,0),a\cap b) = m_0(a)m_0(b) = 1$. We then see that a and b intersect once in $(0,0)$. Note that $a\cap b = \{c\in\mathbb{C} : c^4-c^3+c\} = \{(c,c^2) : c\in V(x,x^3-x^2+1)\}$. Hence we get an intersection at $(0,0)$ and another real intersection (c,c^2)

$$c = \frac{1}{3}\left(1-\sqrt[3]{\frac{2}{25-\sqrt{69}}}-\sqrt[3]{\frac{1}{2}(25-3\sqrt{69})}\right),$$

together with a pair of distinct non-real solutions. We thus have that a and b have 4 intersections each of multiplicity 1.

3. Consider a as before and $c = y^2-x^3$. It is clear that $a\cap c = \{(\alpha,\alpha^2) : \alpha V(x,x-1)\} = \{(0,0),(1,1)\}$.

$$\begin{aligned}I((0,0),a\cap c) &= I((0,0),a\cap c-ay) = I((0,0),a\cap x^3+x^2y) = I((0,0),a\cap x^2(y+x))\\ &= 2I((0,0),a\cap x)+I((0,0),a\cap y+x) = 2m_0(a)m_0(x)+m_0(a)m_0(y+x) = 3.\end{aligned}$$

We also have that $a(x+1,y+1) = -x^2+y-2x$ and $b(x+1,y+1) = y^2+2y+1-x^3-3x^2-3x-1 = -x^3+y^2-3x^2+2y-3x$. These polynomials clearly intersect transversally, hence $I((1,1),a\cap b) = I((0,0),a(x+1,y+1)\cap b(x+1,y+1)) = 1$. a and b therefor have 4 intersections; 3 at $(0,0)$ and 1 at $(1,1)$.

Proposition 4.4.54. *If P is a simple point of f, then $I(P,f\cap g)=\text{ord}_P^f(g)$.*

Proof. We note that when writing f as a product of it's components $p_1,\dots,p_n$. Then $m_P(f)=\sum_1^n m_P(p_i)$, hence WLOG, P is passed through by only one p_i. We thus have that it is sufficient to consider the case where f is irreducible by property 2. and 6. We thus get that $\text{ord}_P^f(g)=\dim\ \mathcal{O}_P(f)/\langle g\rangle$ by Lemma 2.8.95. We note that $\langle f\rangle\subset\langle f,g\rangle\subset K[x,y]$. Note also that

$$\langle f,g\rangle/\langle f\rangle=\langle g\rangle/\langle f\rangle\subset\Gamma(f).$$

Then

$$\mathcal{O}_P(\mathbb{A}^2)/\langle f,g\rangle\simeq\mathcal{O}_P(f)/\langle g\rangle,$$

by Proposition 4.3.68. We thus conclude that

$$\text{ord}_P^f(g)=\dim\ \mathcal{O}_P(f)/\langle g\rangle=\dim\ \mathcal{O}_P(\mathbb{A}^2)/\langle f,g\rangle=I(P,f\cap g),$$

where the second equality is due to Lemma 2.8.95 □

Remark 4.4.55. Deduce only using 1-7

Proposition 4.4.56. *If $\gcd(f,g)=1$, then $\sum_{P\in\mathbb{A}^2}I(P,f\cap g)=\dim\ K[x,y]/\langle f,g\rangle$*

Proof. Note that $f\cap g$ only pass through finitely many $P\in\mathbb{A}^2$ by Theorem 4.2.13. It follows from Corollary 4.3.76 that

$$\sum_{P\in\mathbb{A}^2}I(P,f\cap g)=\dim\ K[x,y]/\langle f,g\rangle$$

□

Corollary 4.4.57. *A line l is tangent to f if and only if $I(P,f\cap l)>m_P(f)$.*

Proof. This follows immediately from property 5. Indeed, l is a tangent of f if and only if l and f has a tangent in common, which by 5. is equivalent to

$$I(P,f\cap l)>m_P(f)m_P(l)=m_P(f).$$

□

Proposition 4.4.58. *If P is a simple point on f, then*

$$I(P,f\cap g+h)\geq\min(I(P,f\cap g),I(P,f\cap h)).$$

Proof. Indeed, it follows from Proposition 4.4.54 that

$$I(P, f \cap g + h) = \mathrm{ord}_P^f(g+h) \geq \min(\mathrm{ord}_P^f(g), \mathrm{ord}_P^f(h)) = \min(I(P, f \cap g), I(P, f \cap h)).$$

□

Example 4.4.59. Show bound is not true when point is multiple

Proposition 4.4.60. *Let f be a curve and l a line that is not a component of f. If $l = \{(a+tb, c+td) : t \in K\}$. Define $g := f(a+bz, c+dz) \in K[x]$. For suitable distinct $\alpha_i \in K$, $r_i \geq 1$ we may write*

$$g = \alpha \prod_1^m (z - \alpha_i)^{r_i}.$$

Then there is a one-to-one correspondence between roots of g and points in $l \cap f$ given by $\alpha_i \mapsto P_i := (a + b\alpha_i, c + d\alpha_i)$. Moreover $I(P_i, f \cap l) = r_i$ and

$$\sum_{P \in \mathbb{A}^2} I(P, l \cap f) \leq \deg\ f.$$

Proof. The one-to-one correspondence is established in Proposition 4.1.22. WLOG $l = x$ and $P_i = (0,0)$. Then

$$f(0,z) = g = z^{r_i} \underbrace{\prod (z - \lambda_j)^{r_j}}_{h}.$$

We can therefor write $f = y^{r_i} h + h'$ where $h'(0,y) = 0$. Note that $h' = xh''$ $h(P_i) \neq 0$ by assumption. Then

$$I(P_i, l \cap f) = I(P_i, x \cap f - xh'') = I(P_i, x \cap y^{r_i} h) = I(P_i, x \cap y^{r_i}) + I(P_i, x \cap h) = r_i I(P_i, x \cap y) = r_i.$$

It then readily follows that.

$$\sum_P I(P, l \cap f) = \sum_1^m e_i = \deg\ g \leq \deg\ f.$$

The last inequality follows from the fact evaluating f in linear polynomials cannot increase the degree (when we use the convention that $\deg\ 0 = 0$). □

Definition 4.4.61. Let f be a curve with only one tangent l at a double point P. f is said to have an *(ordinary) cusp* at P if $I(P, f \cap l) = 3$.

Lemma 4.4.62. *Let f be a curve with a tangent l at a double point P. Then $I(P, f \cap l) \geq 3$.*

Proof. It follows from Corollary 4.4.57 that

$$I(P, f \cap l) > m_P(f) \geq 2.$$

$\square$

Lemma 4.4.63. *Suppose* char $K \neq 2,3$. *Let* $P = (0,0)$ *and* $l = y$ *be the only tangent of* f *at a double point* P. *Then* P *is a cusp if and only if* $f_{xxx}(P) \neq 0$

Proof. "$\Rightarrow$": $f = y^2 + h$, where every term in h has degree greater than 2. Then $3 = I(P, y \cap f) = I(P, y \cap h) \geq m_P(h) = 3$, which means y is not tangent to h. Then x^3 must divide the lowest degree term of h, hence $f_{xxx}(P) = h_{xxx}(P) = (cx^3 + \dots)_{xxx} = 6c \neq 0$.
"$\Leftarrow$": If $f_{xxx}(P) = 0$, then y is tangent to h (using the same notation as for the first implication). Then $I(P, l \cap f) = I(P, l \cap h) > m_P(h) = 3$ $\square$

Remark 4.4.64. Since every technique used above is indifferent to an affine change of coordinates we get the above result for general tangents and points.

Proposition 4.4.65. *If* P *is a cusp on* f, *then only one component of* f *passes through* P.

Proof. $\square$

Definition 4.4.66. A point P on a curve f is called a *hypercusp* if $m_P(f) > 1$, f has a single tangent l and $I(P, f \cap l) = m_P(f) + 1$.

Lemma 4.4.67.

Proof. $\square$

Proposition 4.4.68.

Proof. $\square$

4.4.4 The Dimension of an Affine Variety

Definition 4.4.69. Let $V \subset \mathbb{A}^n$ be an affine variety. We define the *dimension* of V to be

$$\dim V = \operatorname{trdeg}_K K(V)$$

Remark 4.4.70. Note that $\mathbf{dim}$ is just the functor $\mathbf{trdeg}_K \circ \boldsymbol{Q} \circ \Gamma$. It is thus an integer invariant. Hence if $V \simeq W$, then $\mathbf{dim}\ V = \mathbf{dim}\ W$. Note that $K(\mathbb{A}^n) = K(x_1,\dots,x_n)$, hence $\mathbf{dim}\ \mathbb{A}^n = n$. If V is a linear subvariety, then $V \simeq V(x_{d+1},\dots,x_n) \simeq \mathbb{A}^d$ for some d, hence $\mathbf{dim}\ V = d$, hence this generalizes the notion of dimension in definition 4.3.30. By Lemma 2.10.70 every variety has finite dimension.

Lemma 4.4.71. *If $V \subset W$ is a subvariety of an affine variety W, then $\mathbf{dim}\ V \le \mathbf{dim}\ W$. Furthermore, $\mathbf{dim}\ V = \mathbf{dim}\ W \iff V = W$.*

Proof. Note that $\Gamma(W) \to \Gamma(V)$ is given by restriction of a polynomial function on W to V, which is clearly surjective. Let $\widetilde{\varphi} : K(W) \twoheadrightarrow K(V)$ be the surjection between rational functions on W to a rational functions on V induced by this map. Since this is a surjective K-algebra homomorphism, $\mathbf{dim}\ W = \mathbf{trdeg}\ K(W) \ge \mathbf{trdeg}\ K(V) = \mathbf{dim}\, V$ by Lemma 2.10.43. Suppose $n := \mathbf{dim}\ V = \mathbf{dim}\ W$. Let $f \in I(V)$. Take $\overline{g} \in \gamma(W)$ such that $\overline{g}\big|_V = \overline{f} \in \Gamma(V)$. There is a monic, non-zero $H \in K[x_1,\dots,x_n][y]$ ($x_1,\dots,x_n$ are polynomial variables over K) such that $H(\overline{g}) = 0$. Then $H(\overline{f}) = 0$, and the minimal monic polynomial in $K[x_1,\dots,x_n][y]$ vanishing on $\overline{f}$ is y, hence $H = y$, meaning $g \in I(W)$. It follows that $I(V) = I(W)$, hence $V = W$ by Lemma 4.1.42. ☐

4.4.5 Finite Polynomial Maps

Definition 4.4.72. Let $V \subset \mathbb{A}^n$ and $W \subset \mathbb{A}^m$ be affine varieties and $\varphi : V \to W$ a polynomial map. φ is said to be *finite* if $\widetilde{\varphi} : \Gamma(W) \to \Gamma(V)$ is finite.

Lemma 4.4.73. *Let $V \subset \mathbb{A}^n$, $W \subset \mathbb{A}^m$ and $Z \subset \mathbb{A}^l$ be affine varieties and $\varphi : V \to W$, $\psi : W \to Z$ be finite polynomial maps. Then $\psi\varphi$ is finite*

Proof. This follows from functoriality of $\Gamma(\bullet)$ and Lemma 2.10.60. ☐

Lemma 4.4.74. *Let $V \subset \mathbb{A}^n$ be a d-dimensional, affine variety. Then there is a finite polynomial map $\pi = (y_1,\dots,y_d) : V \to \mathbb{A}^d$ where* $\mathbf{deg}\ y_i = 1$.

Proof. By Corollary 2.10.71 there is a finite K-algebra homomorphism $\iota : K[y_1,\dots,y_d] \hookrightarrow \Gamma(V)$, where $y_1,\dots,y_n \in \Gamma(v)$ is a Noether normalization of $\Gamma(V)$. By the Noether normalization theorem for infinite fields we may pick $y_i = \sum_1^n a_{ij} x_j$ for suitable $a_{ij} \in K$. Then upon defining

$$\begin{aligned} \pi : V &\to \mathbb{A}^d \\ v &\mapsto (y_1(v),\dots,y_d(v)) \end{aligned}$$

we see that $\widetilde{\pi} = \iota$, hence π is finite. ☐

4.4.6 A Second Approach to an ENS

Theorem 4.4.75. *Consider non-constant $f_1,\dots,f_m \in K[x_1,\dots,x_n]$ such that $V(f_1,\dots,f_m) = \emptyset$ and $d_1 \geq \dots \geq d_m$. Suppose $m \leq n$. Then there are $\lambda_1,\dots,\lambda_m \in K[\mathbf{x}]$ such that*

1. $\sum_1^m \lambda_i f_i = 1,$
2. $\mathbf{deg}\ \lambda_i f_i \leq \prod_1^m d_i \leq D^m \quad (i \in \{1,\dots,m\}, D := \max(d_1,\dots,d_m).$

Proof. Using WNS we can find $\mu_1,\dots,\mu_m \in K[x_1,\dots,x_n]$ such that

$$\sum_1^n \mu_i f_i = 1.$$

Consider the polynomial map

$$\varphi : \mathbb{A}^n \times \mathbb{A}^1 \to \mathbb{A}^n \times \mathbb{A}^m$$
$$(v,a) \mapsto (v, af_1(v),\dots,af_m(v)).$$

. We prove that $X = Y$. It is clear that $X \subset Y$. Let $(v,w) \in Y$. WLOG $f_1(v) \neq 0$. Set $a = \frac{w_1}{f_1(v)}$. For each i we $w_i f_1(v) = w_1 f_i(v)$, hence $w_i = af_i(v)$. It follows that

$$(v,w) = (v,af_1(v),\dots,af_m(v)) = \varphi(v,a) \in X.$$

Since $\varphi^{-1}(X) = \mathbb{A}^n \times \mathbb{A}^1$ is a variety and $X = \mathbf{im}\ \varphi$ it follows that X is a variety by Lemma 4.3.4. Then $\phi : \mathbb{A}^n \times \mathbb{A}^1 \to X$ is a polynomial map with inverse

$$\phi^{-1} : X \to \mathbb{A}^n \times \mathbb{A}^1$$
$$(v,w) = (v,af_1(v),\dots,af_m(v)) \mapsto \left(v, a\sum_1^m g_i(v)f_i(v)\right) = (v,a).$$

Then ϕ is finite by Lemma **??** and $\mathbf{dim}\ X = \mathbf{dim}\ \mathbb{A}^n \times \mathbb{A}^1 = n+1$. Then it follows by Lemma 4.4.74 that there is a finite affine change of coordinates

$$\pi : X \to \mathbb{A}^{n+1}$$

Defined by linear forms $l_1 + \sum_1^m a_{1j}y_j, l_2 + \sum_2^m a_{2j}y_j,\dots,l_m + a_{mm}y_m, l_{m+1},\dots,l_{n+1} \in K[\mathbf{y},\mathbf{y}]$ where $l_i \in K[\mathbf{x}]$ are linear forms for each i. Then $\psi = \pi \circ \phi$ is a finite polynomial map by Lemma 4.4.73. This polynomial map is defined by $h_1 := l_1 + \sum_1^m a_{1j}zf_j, h_2 := l_2 + \sum_2^m a_{2j}zf_j, \cdots, h_m := l_m + a_{mm}zf_m, h_{m+1} := l_{m+1},\dots,h_{n+1} := l_{n+1} \in K[\mathbf{x}][z]$. Set $L := K(z)$. When viewing $h_i \in L[\mathbf{x}]$, we see that $\mathbf{deg}\ h_i = d_i$ for $i \leq m$ and $\mathbf{deg}\ h_i = 1$ for $i > m$. By finiteness the map

$$\widetilde{\psi} : K[\mathbf{x}][z] \to K[\mathbf{x}][z], f \mapsto f(h_1,\dots,h_{n+1})$$

is finite. Then $K[\mathbf{x}][z]$ is finitely generated as a $K[h_1,\ldots,h_n,h_{n+1}]$-module, hence $K[\mathbf{x}][z] \supset K[h_1,\ldots,h_n,h_{n+1}]$ is integral. Then there is a monic polynomial $P_z \in K[Y_1,\ldots,Y_{n+1},t] \setminus 0$ of minimal degree in t such that $P_z(h_1,\ldots,h_{n+1},z) = 0$ (cf. Remark 2.10.14). By Perron's theorem there is a $Q \in L[Y_1,\ldots,Y_{n+1}] \setminus 0$ such that

a. $Q(h_1,\ldots,h_{n+1}) = 0$

b. $\deg\ Q(Y_1^{d_1},\ldots,Y_m^{d_m},Y_{m+1},\ldots,,Y_{n+1}) \le \prod_1^m d_i$

After scaling Q by an appropriate power of z, we get a polynomial P such that

a. $P(h_1,\ldots,h_{n+1},z) = 0$

b. $\deg_{\mathbf{Y}}\ P(Y_1^{d_1},\ldots,Y_m^{d_m},Y_{m+1},\ldots,,Y_{n+1},z) \le \prod_1^m d_i$

Then $P_z \mid P$, hence $\deg_{\mathbf{Y}}\ P_z(Y_1^{d_1},\ldots,Y_m^{d_m},Y_{m+1},\ldots,,Y_{n+1},z) \le \prod_1^m d_i$. Write $P_z(Y_1,\ldots,Y_{n+1},z) = z^N + \sum_0^{N-1} b_i z^i$, $b_i \in K[\mathbf{Y}]$. Set $b_N := 1$ and write

$$b_i(h_1,\ldots,h_{n+1})z^i = \mu_i z^N + \sum_{j\neq N} \nu_{ij} t^j,$$

for suitable $\mu_i, \nu_j \in K[\mathbf{x}]$. Then

$$0 = P_z(h_1,\ldots,h_{n+1},z) = z^N + \sum_0^{N-1} b_i(h_1,\ldots,h_n)t^i = (1 + \sum_0^{N-1} \mu_i)z^N + \sum_{j\neq N} \left[\sum_0^{N-1} \nu_{ij}\right] z^j$$

which implies that $1+\sum_0^{N-1} \mu_i = 0$ hence $1 = -\sum_0^{N-1} \mu_i$. Note that $\deg_z\ b_i(h_1,\ldots,h_{n+1})z^i \ge N$, implies $\deg_z\ b_i(h_1,\ldots,h_{n+1}) > 0$. Now note that each $h_j = \omega_j + \kappa_j z$, where $\omega_j, \kappa_j \in K[\mathbf{x}]$ and $\kappa_j \in \langle f_1,\ldots,f_m\rangle$, hence $\mu_i \in \langle f_1,\ldots,f_m\rangle$, hence for suitable $\lambda_1,\ldots,\lambda_m \in K[\mathbf{x}]$,

$$1 = -\sum_0^{N-1} \mu_i = \sum_1^m \lambda_i f_i.$$

Secondly using Lemma 2.9.44,

$$\begin{aligned} \deg\ \lambda_i f_i \le \max_{1,\ldots,N-1}\ \deg\ \mu_i &\le \deg_{\mathbf{x}}\ P_z(h_1,\ldots,h_{n+1},z) \\ &\le \deg_{\mathbf{Y}}\ P_z(Y_1^{d_1},\ldots,Y_m^{d_m},Y_{m+1},\ldots,Y_{n+1},z) \le \prod_1^m d_i. \end{aligned}$$

□

Remark 4.4.76. Suppose $m > n$. Then we still get the bound by considering $f_1,\ldots,f_m$ as elements of $K[x_1,\ldots,x_n,\ldots,x_m]$.

5 Projective & Multiprojective Algebraic Geometry

5.1 Projective Algebraic Sets & Projective Varieties

5.1.1 Basic Definitions

Definition 5.1.1. Let $S \subset K[x_1, \ldots, x_{n+1}]$ then we define *the projective vanishing of S over K* to be the set

$$V^{\mathbb{P}}(S) := V(S) := \{[v] \in \mathbb{P}^n : [v] \text{ is a zero for every } f \in S\}$$

Definition 5.1.2. A set $X \subset \mathbb{P}^n$ is said to be a *projective algebraic set* if

$$X = V(S)$$

for some $S \subset K[x_1, \ldots, x_{n+1}]$. X is called a *(projective) hypersurface* if $X = V(F)$ for some homogeneous $F \in K[\mathbf{x}]$. If $\mathbf{deg}\ F = 1$ we call it a *(projective) hyperplane.*

Remark 5.1.3. One notes that the hyperplanes $V(x_i)$ are the i'th hyperplanes at infinity.

Proposition 5.1.4. *Let K be an infinite field, $S \subset K[x_1, \ldots, x_{n+1}]$ and $I := \langle S \rangle = \langle f_1, \ldots, f_m \rangle \subset K[x_1, \ldots, x_{n+1}]$. Write*

$$f_i = \sum_{0}^{d_i} f_{ij},$$

where f_{ij} are forms of degree j. Then

$$V(S) = V(I) = V(\{f_{ij} : 1 \le i \le m, 0 \le j \le d_i\}).$$

Proof. The first equality is trivial and the second equality follows from Lemma 2.9.123. □

Remark 5.1.5. It thus follows that if $X \subset \mathbb{P}^n$ is a projective algebraic set, we may assume that $X = V(F_1, \ldots, F_m)$ for homogeneous polynomials $F_1, \ldots, F_m \in K[x_1, \ldots, x_{n+1}]$. As a consequence if given a homogeneous ideal $I \subset K[\mathbf{x}]$, then $V^{\mathbb{A}}(I) \setminus 0 \neq \emptyset$ implies $V^{\mathbb{P}}(I) \neq \emptyset$. Indeed, we may write $I = \langle F_1, \ldots, F_m \rangle$ for suitable homogeneous polynomials $F_1, \ldots, F_m \in K[\mathbf{x}]$. Meaning if $v \in \mathbb{A}^{n+1} \setminus 0$ is a zero of each F_i, then so is $[v]$.

Lemma 5.1.6. *If $K[x_1, \ldots, x_{n+1}] \supset M \supset M'$, then $V^{\mathbb{P}}(M) \subset V^{\mathbb{P}}(M')$.*

Proof. The proof is almost identical to the proof of Lemma 4.1.5. □

Lemma 5.1.7. *We collect the following results which are projective analogues of Lemma 4.1.10*

(i) Let A be some indexing set. Consider a family of algebraic sets $\{X_\alpha\}_{\alpha\in A}$ in $\mathbb{A}^n$. Then $\bigcap_\alpha X_\alpha$ is an algebraic set.

(ii) Consider algebraic sets $X, Y \subset \mathbb{A}^n$. Then $X \cup Y$ is algebraic. It follows by induction that $\bigcup_1^k X_i$ is algebraic for any finite sequence of algebraic sets $X_1, \ldots, X_k$ in $\mathbb{A}^n$.

Proof. 1. Writing $X_\alpha = V(I_\alpha)$ for homogeneous I_α, we prove the statement in the same manner as for the affine case, i.e. by proving $\bigcap_\alpha X_\alpha = V(\bigcup_\alpha I_\alpha) = V(\langle\bigcup_\alpha I_\alpha\rangle) = V(\sum_\alpha I_\alpha)$. The proof of this is identical to the one given in Lemma 4.1.10 1.
2. For homogeneous ideals $I, J \subset K[\mathbf{x}]$. We prove that $V(I) \cup V(J) = V(IJ)$ for which the proof is identical to that of Lemma 4.1.10 2. □

Example 5.1.8. 1. $V^{\mathbb{P}}(0) = \mathbb{P}^n$. Indeed $0([v]) = 0$ for every $[v] \in \mathbb{P}^n$.

2. $V^{\mathbb{P}}(1) = \emptyset$. Indeed $1(v) = 1 \neq 0$ for every $v \in \mathbb{A}^{n+1} \setminus 0$.

3. $V^{\mathbb{P}}(x_1 - x_{n+1}a_1, \ldots, x_n - x_{n+1}a_n) = \{[(a_1, \ldots, a_n, 1)]\}$. Indeed, set $v = a_1, \ldots, a_n, 1) \in \mathbb{A}^{n+1}$. Then evaluating each of the n polynomials in λv for $\lambda \in K \setminus 0$, we get 0. Conversely if $w_i = w_{n+1}a_i$ for each i, note that necessarily $w_{n+1} \neq 0$, hence $(w_1, \ldots, w_{n+1}) = w_{n+1}v$, hence $[w] = [v]$. After permutation of indices we see that any point in $\mathbb{P}^n$ is a projective algebraic set, hence any finite subset of $\mathbb{P}^n$ is algebraic

Remark 5.1.9. The system

$$\tau_{\mathcal{Z}} := \left\{\mathbb{P}^n \setminus X : X \subset \mathbb{P}^n \text{ is an algebraic set}\right\}$$

(analogously to the affine case) defines a topology on $\mathbb{P}^n$ which we also call *Zariski topology.*

Definition 5.1.10. For a set $X \subset \mathbb{P}^n$ we form

$$I^{\mathbb{P}}(X) := I(X) := \{f \in K[x_1, \ldots, x_n] : f([v]) \text{ for every } [v] \in X\}$$

which we call the *(homogeneous) ideal of X*

Remark 5.1.11. The above set is a homogeneous ideal. That it is an ideal is trivial. Again by Lemma 2.9.123 it is homogeneous. We also have that $I(X)$ is generated by a finite set of homogeneous polynomials due to Lemma 2.9.125.

Lemma 5.1.12. *If $X, Y \subset \mathbb{P}^n$ are algebraic such that $X \subset Y$, then $I(X) \supset I(Y)$.*

Proof. The proof is identical to the affine case $\square$

Lemma 5.1.13. *Let $M \subset K[x_1, \dots, x_{n+1}]$ and $X \subset \mathbb{P}^n$. Then we have the following*

1. *$I(V(M)) \supset M$.*

2. *$V(I(X)) \supset X$.*

3. *$V(I(V(M))) = V(M)$. Hence if X is algebraic $X = V(I(X))$.*

4. *$I(V(I(X))) = I(X)$. Hence if M is an ideal of some algebraic set, $M = I(V(M))$*

Proof. The proof is identical to the affine case. $\square$

Lemma 5.1.14. *Let $X \subset \mathbb{A}^n$. Then $I(X)$ is radical.*

Proof. The proof is identical to the affine case. $\square$

Lemma 5.1.15. *Let $X, Y \subset \mathbb{A}^n$ be algebraic subsets. Then*

$$X = Y \iff I(X) = I(Y).$$

Proof. The proof is identical to the affine case. $\square$

Example 5.1.16. 1. $I(\emptyset) = K[x_1, \dots, x_{n+1}]$

2. If $\#K = \infty$, $I(\mathbb{P}^n) = 0$. Indeed, if $f([v]) = 0$ for every $[v]$, then $f_0([v]) = 0$, hence $f(0) = 0$. It thus follows that $f(v) = 0$ for every $v \in \mathbb{A}^{n+1}$, hence $f = 0$.

3. $I(\{[v_1, \dots, v_n, 1]\}) = \langle x_1 - x_{n+1}v_1, \dots, x_n - x_{n+1}v_n \rangle$.

Definition 5.1.17. An algebraic set is *reducible* if can be written as the union of two strictly smaller algebraic set. It is called *irreducible* or *a (projective) variety*, if it is not reducible.

Lemma 5.1.18. *An algebraic set $V \subset \mathbb{P}^n$ is a variety if and only if $I(V)$ is prime if and only if $K[x_1, \dots, x_{n+1}]/I(V)$ is an integral domain.*

Proof. "$\Rightarrow$": Suppose $I(V)$ is not prime. Then there are homogeneous $F_1, F_2 \in K[x_1, \dots, x_{n+1}] \setminus I(V)$ such that $F_1F_2 \in I(V)$, by Lemma 2.9.127. Then $V \cap V(F_i) \subsetneq V$ and

$$V = V \cap V(F_1F_2) = (V \cap V(F_1)) \cup (V \cap V(F_2)).$$

"$\Leftarrow$": Conversely if $V = V_1 \cup V_2$ where $V_i \subsetneq V$, then $I(V_i) \supsetneq I(V)$. Pick $F_i \in I(V_i) \setminus I(V)$, then $F_1F_2 \in I(V_1 \cup V_2) = I(V)$, hence $I(V)$ is not prime. $\square$

Again using that $K[x_1, \dots, x_{n+1}]$ is noetherian, we can prove that any projective algebraic set has unique decomposition into a union of projective varieties.

5.1.2 The Cone of a Projective Algebraic Set

Definition 5.1.19. Let $X \subset \mathbb{P}^n$ be algebraic. We define *the cone over X* to be the set

$$C(X) := \{v \in \mathbb{A}^{n+1} : [v] \in X \text{ or } v = 0\}.$$

Lemma 5.1.20. *Let $V \subset \mathbb{P}^n$ be a non-empty algebraic set. Then*

$$I^{\mathbb{A}}(C(V)) = I^{\mathbb{P}}(V).$$

Proof. "$\subset$": Let $f \in I^{\mathbb{A}}(C(V))$ and $[v] \in V$. Then $\lambda v \in C(V)$ for every $\lambda \in K \setminus 0$, meaning $f(\lambda v) = 0$. Then $f([v]) = 0$, hence $f \in I^{\mathbb{P}}(V)$.
"$\supset$": Let $f \in I^{\mathbb{P}}(V)$ and $v \in C(V)$. If $[v] \in V$, then $f([v]) = 0$, hence in particular $f(v) = 0$. If $v = 0$, note that writing $f = \sum_0^d f_i$, we get that $f_i(w) = 0$ for a $[w] \in V$, hence $f_i = 0$. Then trivially $f(0) = 0$. □

Lemma 5.1.21. *If $I \subset K[x_1, \ldots, x_{n+1}]$ is homogeneous, then $V^{\mathbb{P}}(I) \neq \emptyset$ implies $C(V^{\mathbb{P}}(I)) = V^{\mathbb{A}}(I)$*

Proof. "$\subset$": Let $v \in C(V^{\mathbb{P}}(I))$ and $f = \sum_0^d f_i \in I$. Then if $[v] \in V^{\mathbb{P}}(I)$, $f([v]) = 0$, hence $f(v) = 0$. If $v = 0$, then for a given $[w] \in V^{\mathbb{P}}(I)$, $f_i([w]) = 0$, hence $f_i = 0$, meaning $f_i(0) = 0$. In any case $v \in V^{\mathbb{A}}(I)$.
"$\supset$": Let $v \in V^{\mathbb{A}}(I)$ and $f = \sum_0^d f_i \in I$. If $v = 0$, then $v \in C(V^{\mathbb{P}}(I))$ trivially. So suppose $v \neq 0$. Note that $f_i(v) = 0$ for each i since $f_i \in I$, hence $f_i(\lambda v) = 0$ for each $\lambda \in K \setminus 0$. It thus follows that $f([v]) = 0$, hence $v \in C(V^{\mathbb{P}}(I))$. □

Lemma 5.1.22. *If $I \subset K[x_1, \ldots, x_{n+1}]$ is homogeneous and not contained in $\langle x_1, \ldots, x_{n+1}\rangle$, then $V^{\mathbb{P}}(I) \neq \emptyset$ if and only if $C(V^{\mathbb{P}}(I)) = V^{\mathbb{A}}(I)$. In the case $I \not\subset \langle x_1, \ldots, x_{n+1}\rangle$, $V^{\mathbb{A}}(I) = \emptyset$. In the case $I = \langle x_1, \ldots, x_{n+1}\rangle^d$, then $C(V^{\mathbb{P}}(I)) = \{0\} = V^{\mathbb{A}}(I)$.*

Proof. "$\Rightarrow$": Follows immediately from the prior lemma.
"$\Leftarrow$": If $V^{\mathbb{P}}(I) = \emptyset$, then $C(V^{\mathbb{P}}(I)) = \{0\}$ and $V^{\mathbb{A}}(I) \setminus 0 = \emptyset$ (cf. Remark 5.1.5). Note also that "$\supset$" in the prior lemma holds true in any case. Then $V^{\mathbb{A}}(I) \in \{\emptyset, \{0\}\}$. If $V^{\mathbb{A}}(I) = \{0\}$, then $I \subset I(V^{\mathbb{A}}(I)) = \langle x_1, \ldots, x_n\rangle$, so we conclude that $V^{\mathbb{A}}(I) = \emptyset \subsetneq \{0\} = C(V^{\mathbb{P}}(I))$.
If $I = \langle x_1, \ldots, x_{n+1}\rangle^d$, then $v \in V^{\mathbb{A}}(I) \iff v = 0$, hence $V^{\mathbb{P}}(I) = \emptyset$ □

Proposition 5.1.23. *Each irreducible component of a cone $C(V^{\mathbb{P}}(I))$ is itself a cone.*

Proof. **Claim:** Let $W \subset V^{\mathbb{P}}(I)$ be a component of $V^{\mathbb{P}}(I)$. Then $C(W) \subset C(V^{\mathbb{P}}(I))$ is a component. Suppose $V := V^{\mathbb{P}}(I) = V_1 \cup V_2$ with $V_i \subsetneq V$. Then there is a $[v_i] \in V$ such that $[v_i] \notin V_i$. Then $v_i \in C(V) \setminus C(V_i)$.
Let U be a component of $C(V)$. By uniqueness of decomposition into components $U = C(W)$ for some component of V. □

5.1.3 The Projective Nullstellensatz

Lemma 5.1.24. *Let $I \subset K[x_1,\dots,x_{n+1}]$ be an homogeneous ideal. The following are equivalent.*

1. $V^{\mathbb{P}}(I) = \emptyset$.
2. $V^{\mathbb{A}}(I) \subset \{0\}$.
3. $\mathrm{rad}(I) = I^{\mathbb{A}}(V^{\mathbb{A}}(I)) \supset \langle x_1,\dots,x_{n+1}\rangle$.
4. $\langle x_1,\dots,x_{n+1}\rangle^d \subset I$ *for a suitably large* $d \geq 0$

Proof. "1. $\iff$ 2.": This follows from the prior lemma. "2. $\iff$ 3.": This follows from HNS. "3. $\iff$ 4.": This follows from Lemma 2.8.33. □

Theorem 5.1.25. *(Projective Nullstellensatz/PNS)*
Let $I \subset K[x_1,\dots,x_{n+1}]$ be a homogeneous ideal. Then

1. $V^{\mathbb{P}}(I) = \emptyset$ *if and only if* $V(d, n+1) \subset I$ *for some* $d \geq 0$.
2. *If* $V^{\mathbb{P}}(I) \neq \emptyset$, *then* $I^{\mathbb{P}}(V^{\mathbb{P}}(I)) = \mathrm{rad}(I)$.

Proof. 1. Is an immediate consequence of the above lemma.
2. By Lemma 5.1.21 $V^{\mathbb{A}}(I) = C(V^{\mathbb{P}}(I))$ and by Lemma 5.1.20

$$I^{\mathbb{P}}(V^{\mathbb{P}}(I)) = I^{\mathbb{A}}(C(V^{\mathbb{P}}(I))) = I^{\mathbb{A}}(V^{\mathbb{A}}(I)) = \mathrm{rad}(I),$$

where the last equality is just HNS. □

write relevant corollaries to the PNS

5.1.4 Rational Functions and Local Rings of Projective Varieties

Definition 5.1.26. For a projective variety, $V \subset \mathbb{P}^n$ we define the *homogeneous coordinate ring of* V to be the integral domain (recall that $I(V)$ is prime)

$$\Gamma^h(V) := \Gamma(V) := K[\mathbf{x}]/I^{\mathbb{P}}(V).$$

Remark 5.1.27. In general it is not true that the homogeneous coordinate ring can be thought of as a subring of $\mathbf{Fun}(V,K)$, indeed consider for example $f = \sum_0^d f_i \in K[x_1,\dots,x_n]$, then for a $v \in \mathbb{A}^{n+1} \setminus 0$, $f(\lambda v) = \sum_0^d \lambda^i f_i(v)$. Note then that the function $\lambda \mapsto f(\lambda v)$ is constant if and only if $f_i = 0$ for $i > 0$. Hence for some $\lambda, \lambda' \in K \setminus 0$, $f(\lambda v) \neq f(\lambda' v)$, meaning $f([v])$ is not well-defined unless f is constant. What then can be said is that for $V \neq \emptyset$, $\{c + I(V) : c \in K\} \simeq \{([v] \mapsto c) \in \mathbf{Fun}(V,K) : c \in K\}$, which is trivial.

Definition 5.1.28. For a projective variety, $V \subset \mathbb{P}^n$ we define the *homogeneous function field of* V to be the field

$$K^h(V) := Q(\Gamma^h(V).$$

Remark 5.1.29. It is also clear that the homogeneous function field is not isomorphic to some subring of $Q(\mathrm{Fun}(V,K))$. We can however make an amendment to this consider $\phi = \frac{f+I(V)}{g+I(V)} \in K^h(V)$, where $f+I(V), g+I(V)$ are homogeneous of degree $d \geq 0$. Then for every $v \in V$ with $(g+I(V))(v) = g(v) \neq 0$ and every $\lambda \in K \setminus 0$,

$$\frac{f(\lambda v)}{g(\lambda v)} = \frac{\lambda^d f(v)}{\lambda^d g(v)} = \frac{f(v)}{g(v)}.$$

We thus get that $\phi([v]) := \frac{f(v)}{g(v)}$ is well-defined. It is also independent of choice of representative of ϕ, since evaluation of $\frac{f+I(V)}{g+I(V)}$ is independent of representatives.

Inspired by the above remark we introduce the following definition.

Definition 5.1.30. For a projective variety, $V \subset \mathbb{P}^n$ we define the *function field of* V to be

$$K(V) := \left\{ z \in K^h(V) : \exists d \geq 0 f, g \in V(d, n+1), g \neq 0, z = \frac{f+I(V)}{g+I(V)} \right\}.$$

Remark 5.1.31. This is a subfield of $K^h(V)$ and a K-algebra. Indeed, consider $f, g \in V(d, n+1)$ and $\lambda, \mu \in V(e, n+1)$. Then $f\lambda, g\mu, f\mu, g\lambda$ and there sums are in $V(d+e, n+1)$. This implies that $\frac{f+I(V)}{g+I(V)} + \frac{\lambda+I(V)}{\mu+I(V)}, \frac{f+I(V)}{g+I(V)} \frac{\lambda+I(V)}{\mu+I(V)} \in K(V)$. If $f \neq 0$, then clearly $\left(\frac{f+I(V)}{g+I(V)}\right)^{-1} = \frac{g+I(V)}{f+I(V)}$. We see that an element of this ring can be viewed as a function $[v] \mapsto \frac{f}{g}(v)$ away from the points where $g([v]) = 0$. We therefor also call these elements *rational functions on* V.

Definition 5.1.32. Let $V \subset \mathbb{P}^n$ be a projective variety, $\phi \in K(V)$. We define *the pole set of* ϕ to be the set

$$\mathcal{P}(\phi) := \{P \in \mathbb{P}^n : \phi \text{ not defined at } P\}$$

Definition 5.1.33. For $\phi \in K(V)$, we define $J_\phi := \{g \in K[\mathbf{x}] : (g+I(V))\phi \in \Gamma^h(V)\}$, which we know to be an ideal already (cf. Remark 4.3.48).

Remark 5.1.34. The ideal is also homogeneous. Indeed, write $g = \sum_0^d g_i \in J_\phi$ and $\phi = \frac{\alpha}{\beta}$ where $\alpha, \beta \in \Gamma^h(V)$ are homogeneous of degree δ. Note that $\sum_0^d (g_i + I(V))\alpha$ is a decomposition into a sum of forms of degree $\delta + i$ for $i \leq d$. We have that $(g+I(V))\phi = \sum_0^e h_i$ for suitable homogeneous $h_i \in \Gamma^h(V)$. Then

$$\sum_0^d (g_i + I(V))\alpha = \sum_0^e h_i \beta$$

hence $e = d$ and $(g_i + I(V))\alpha = h_i\beta$ for each i, implying $(g_i + I(V))\phi \in \Gamma^h(V) \Rightarrow g_i \in J_\phi$.

Lemma 5.1.35. *Let V be a variety and $\phi \in K(V)$. Then $\mathcal{P}(\phi) = V(J_\phi)$.*

Proof. The proof is identical to the affine case. $\square$

Definition 5.1.36. Let $V \subset \mathbb{P}^n$ be a projective variety. Consider $\phi = \frac{\alpha}{\beta} \in K(V)$, where α and β are homogeneous of degree $d \geq 0$ and a point $P \in V$. We say that ϕ *is defined at P* if P is not a zero of g.
In continuation of this we define *the local ring of V at P* to be

$$\mathcal{O}_P(V) := \{z \in K(V) : z \text{ is defined at } P\}.$$

Remark 5.1.37. One readily verifies that this is a subring of $K(V)$ containing K and that $\mathfrak{m}_P(V) := \mathcal{O}_P(V) \setminus \mathcal{O}_P(V)^* = \{z \in \mathcal{O}_P(V) : z = \frac{\alpha}{\beta}, \alpha(P) = 0\}$ is an ideal, meaning $\mathcal{O}_P(V)$ is a local ring.

Definition 5.1.38. For a variety V with a point $P \in V$ we define *evaluation at P* to be the function

$$\begin{aligned} \mathrm{ev}_P : \mathcal{O}_P(V) &\to K \\ \phi &\mapsto \phi(P) \end{aligned}$$

Remark 5.1.39. One readily verifies that this is a K-algebra homomorphism.

5.1.5 (Projective) change of coordinates

Definition 5.1.40. Consider an affine change of coordinates $\varphi : \mathbb{A}^{n+1} \to \mathbb{A}^{n+1}, v \mapsto Av$, i.e. an affine change of coordinates that maps 0 to 0. We define *the projective change of coordinates* induced by φ to be the function

$$\begin{aligned} \varphi : \mathbb{P}^n &\to \mathbb{P}^n \\ [v] &\mapsto [Av] \end{aligned}$$

Remark 5.1.41. The above function is well-defined: Since A is invertible $Av \neq 0$, hence $[Av]$ is an element of $\mathbb{P}^n$. For any $[v] \in \mathbb{P}^n$, $[v] = L(v, 0) \setminus 0$. Hence by Lemma 4.3.35 and using A being invertible,

$$A(L(\lambda v, 0)) \setminus 0 = A(L(v, 0) \setminus 0) = L(Av, 0) \setminus 0.$$

Note that the inverse of φ is the affine change of coordinates induced by A^{-1}.

Lemma 5.1.42. *If $V \subset \mathbb{P}^n$ is a projective variety and $\varphi : \mathbb{P}^n \to \mathbb{P}^n, [v] \mapsto [Av]$ is a projective change of coordinates, then $\varphi^{-1}(V)$ is a projective variety.*

Proof. One readily verifies that if $V = V(F_1,\dots,F_m)$ for homogeneous $F_1,\dots,F_m$, and φ is induced by linear forms $\varphi_1,\dots,\varphi_{n+1}$, then $F_i(\varphi_1,\dots,\varphi_{n+1})$ is homogeneous and $\varphi^{-1}(V) = V(F_1(\varphi_1,\dots,\varphi_{n+1}),\dots,F_m(\varphi_1,\dots,\varphi_{n+1}))$ is therefor a projective algebraic set. Suppose $\varphi^{-1}(V) = V_1 \cup V_2$. Then $V = \varphi(V_1)\cup\varphi(V_2) = (\varphi^{-1})^{-1}(V_1)\cup(\varphi^{-1})^{-1}(V_2)$, hence WLOG $V = \varphi(V_1)$. Then $\varphi^{-1}(V) = V_1$. □

Lemma 5.1.43. *For a projective variety V and a projective change of coordinates φ induced by linear forms φ_i, we get induced isomorphisms K-algebras*

$$\begin{aligned}\widetilde{\varphi}:\Gamma^h(V)&\to\Gamma^h(\varphi^{-1}(V))\\ f+I(V)&\mapsto f(\varphi_1,\dots,\varphi_{n+1})+I(\varphi^{-1}(V))\end{aligned}$$

(It has inverse $f+I(\varphi^{-1})\mapsto f(\varphi_1^{-1},\dots,\varphi_{n+1}^{-1})+I(V)$). We thus get an induced isomorphism

$$\widetilde{\varphi}:K(V)\to K(\varphi^{-1}(V))$$

and if $\varphi(Q)=P$ an isomorphism

$$\widetilde{\varphi}:\mathcal{O}_P(V)\to\mathcal{O}_Q(\varphi^{-1}(V))$$

Proof. $\widetilde{\varphi}:\Gamma^h(V)\to\Gamma^h(\varphi^{-1}(V))$ is well-defined for the same reason it is in the affine case (i.e. $f(\varphi_1,\dots,\varphi_{n+1})\in I(\varphi^{-1}(V))$ for every $f\in I(V)$). The inverse is $\widetilde{\varphi^{-1}}$. We get an induced isomorphism $K^h(V)\to K^h(\varphi^{-1}(V))$ □

Example 5.1.44. Let $V=\mathbb{P}^1$ and set $t=\frac{x}{y}\in K(V)$. Then $K(V)=K(t)\simeq K(z)$ (i.e. $K(t)$ can be viewed as the quotient field of the polynomial ring in 1 variable over K, since t is algebraically independent over K). It remains to check that $K(V)\subset K(t)$. It is sufficient to check that $\frac{f}{g}\in K(V)$ where f,g are homogeneous of degree $d\ge 1$ are elements of $K(t)$. Consider $\frac{\alpha x+\beta y}{\lambda x+\mu y}\in K(V)$ where $(\alpha,\beta),(\lambda,\mu)\neq(0,0)$. Then

$$\begin{aligned}\frac{\alpha x+\beta y}{\lambda x+\mu y}&=\frac{\alpha x}{\lambda x+\mu y}+\frac{\beta y}{\lambda x+\mu y}\\ &=\left(\frac{\lambda}{\alpha}+\frac{\mu}{\alpha}t^{-1}\right)^{-1}+\left(\frac{\lambda}{\beta}t+\frac{\mu}{\beta}\right)^{-1}\in K(t)\end{aligned}$$

Since f and g can be written as product of linear forms it follows that $\frac{f}{g}\in K(t)$. Using Proposition 4.3.74 we obtain the following one-to-one correspondence

$$\begin{aligned}\mathbb{P}^1=K\cup\{\infty\}&\to\{\text{DVRs containing } K \text{ with quotient field } K(t)\}\\ P&\mapsto\begin{cases}\mathcal{O}_P(\mathbb{A}^1) & \text{if } P\in K\\ \mathcal{O}_\infty & \text{if } P=\infty\end{cases}\end{aligned}$$

Definition 5.1.45. A *projective linear subvariety* is a projective algebraic set of the form $V(L_1,\dots,L_m)$ where $L_1,\dots,L_m \in K[x_1,\dots,x_{n+1}]\backslash 0$ are linear homogeneous polynomials.

Lemma 5.1.46. *The pre-image of a projective linear subvariety under a projective change of coordinates is again a projective linear subvariety.*

Proof. This follows from Lemma 5.1.42 the fact that a non-zero linear homogeneous polynomial evaluated in non-zero linear homogeneous polynomial is itself a non-zero linear homogeneous polynomials. □

Proposition 5.1.47. *Let $V := V(L_1,\dots,L_m)$ be a projective linear subvariety. Then $\varphi^{-1}(V) = V(x_{d+2},\dots,x_{n+1})$ for some projective change of coordinates $\varphi : \mathbb{P}^n \to \mathbb{P}^n$, $-1 \le d \le n-1$ (in particular V is a projective variety). This d is independent of choice of change of coordinates.*

Proof. In the projective we have the ability to handle the case where $V = \emptyset$. Indeed setting $d = -1$ we are done since $V^{\mathbb{P}}(x_1,\dots,x_{n+1}) = \emptyset$. In this case any projective change of coordinates will do the trick, by bijectivity.
Suppose $V \neq \emptyset$. In the $m = 1$, we use the same construction as in the proof of Proposition 4.3.28. Note that this yields an invertible linear transformation and hence a projective change of coordinates. We then get that $\varphi^{-1}(V) = V(x_{n+1})$ and there for $d = n-1$ works. We proceed by induction in m. So consider $V = V(L_1,\dots,L_{m+1})$ and $W = V(L_1,\dots,L_m)$. Then there is an affine change of coordinates $\varphi : \mathbb{P}^n \to \mathbb{P}^n$ induced by some invertible linear transform of the same name such that $\varphi^{-1}(W) = V(x_{d+2},\dots,x_{n+1})$. Then

$$V^{\varphi} = V(x_{d+2},\dots,x_{n+1}) \cap V(h),$$

where $h := (L_{m+1} \circ \varphi)(x_1,\dots,x_{d+1},\mathbf{0})$. As with the affine case (cf. Proposition 4.3.28), $h = 0$, in which case we are done, or $\deg h = 1$ and h is homogeneous. In the second case as there is linear transform $\phi : \mathbb{A}^{d+1} \to \mathbb{A}^{d+1}$ inducing a projective change of coordinates $\psi : \mathbb{P}^d \to \mathbb{P}^d$ such that $\psi^{-1}(V(h)) = V(x_{d+1})$. Taking

$$\begin{aligned} \phi : \mathbb{A}^{n+1} &\to \mathbb{A}^{n+1} \\ v &\mapsto (\psi(v_1,\dots,v_{d+1}), v_{d+2},\dots,v_{n+1}) \end{aligned}$$

it follows that $(\varphi \circ \phi)^{-1}(V) = V(x_{d+1},\dots,v_{n+1})$. The uniqueness follows from the same linear algebra as it did in the affine case. □

Remark 5.1.48. This integer invariant d, will be denoted $\mathbf{dim}\ V$ and will be referred to as the *dimension of* V *(over* K*)*. Note that upon writing $L_i = \sum_1^{n+1} a_{ij}x_j$, the dimension of a projective variety under this definition is the same as $\mathbf{dim}\,\mathbf{ker}\,(a_{ij})-1$. In fact $V^{\mathbb{P}}(L_1,\ldots,L_m)$ is the projection of $(\mathbf{ker}\,(a_{ij}))\setminus 0$ into $\mathbb{P}^{n+1}$. One therefor sees that a d-dimension linear subvariety is just the projective span of some $[v_1],\ldots,[v_{d+1}]$ where $v_1,\ldots,v_{d+1}$ are linearly independent, since for a suitable projective change of coordinates, A,

$$V(L_1,\ldots,L_m) = A(V(x_{d+2},\ldots,x_{n+1}) = A(\text{Span}([e_1],\ldots,[e_{d+1}])) = \text{Span}([Ae_1],\ldots,[Ae_{d+1}])$$

by lemma 5.1.50. Conversely, if $v_1,\ldots v_{d+1} \in \mathbb{A}^{n+1}\setminus 0$ are linearly independent, then $\text{Span}([v_1],\ldots,[v_{d+1}])$ is also a linear subvariety of dimension d.

Definition 5.1.49. Let $[v] \neq [w] \in \mathbb{P}^n$. We define *the line through* $[v]$ *and* $[w]$ to be

$$L([v],[w]) := \text{Span}([v],[w]).$$

If $[v_1],\ldots,[v_n] \in \mathbb{P}^n$ such that $[v_i] \notin \text{Span}([v_1],\ldots,\widehat{[v_i]},\ldots,[v_n])$, we define the *hyperplane through* $[v_1],\ldots,[v_n]$ to be $H([v_1],\ldots,[v_n]) := \text{Span}([v_1],\ldots,[v_n])$

Lemma 5.1.50. *Let* $[v_1],\ldots,[v_m] \in \mathbb{P}^n$ *and* $\varphi : \mathbb{P}^n \to \mathbb{P}^n$ *a projective change of coordinates. Then* $\varphi(\text{Span}([v_1],\ldots,[v_m])) = \text{Span}([\varphi(v_1)],\ldots,[\varphi(v_m)])$

Proof. Set $A \in M_{n+1}(K)$ to be the linear transform inducing ϕ. Then for any $(\lambda_1,\ldots,\lambda_m) \in K^m \setminus 0$,

$$\left[\sum_1^m \lambda_i(Av_i)\right] = \left[A\left(\sum_1^m \lambda_i v_i\right)\right] = \varphi\left(\left[\sum_1^m \lambda_i v_i\right]\right),$$

and the result follows. ☐

Lemma 5.1.51. *Consider* $[v_1],\ldots,[v_m] \in \mathbb{P}^n$ *with* $v_1,\ldots,v_m$ *linearly independent over* K*. Then* $\text{Span}([v_1],\ldots,[v_m])$ *is linear subvariety of dimension* $m-1$ *if* $m \leq n$*, of dimension* n *otherwise. In particular, we have that a line between distinct points* $[v],[w] \in \mathbb{P}^n$ *is a projective linear variety of dimension* 1*. And that the hyperplane through* $[v_1],\ldots,[v_m]$ *is of dimension* n

Proof. In the case $m \geq n+1$ the result is obvious since then $\text{Span}([v_1],\ldots,[v_m]) = \mathbb{P}^n$. In other case, we complete the basis of $\mathbb{A}^{n+1}$, with $v_{m+1},\ldots,v_{n+1}$ such that $v_1,\ldots,v_{n+1}$ spans $\mathbb{A}^{n+1}$. We set $A = (v_{ij})^T$ and using the above lemma find that

$$A^{-1}\text{Span}([v_1],\ldots,[v_m]) = \text{Span}([A^{-1}v_1],\ldots,[A^{-1}v_m]) = \text{Span}([e_1],\ldots,[e_m]) = V(x_{m+1},\ldots,x_{n+1}),$$

hence $\text{Span}([v_1],\ldots,[v_m])$ is a linear subvariety and its dimension is $m-1$. ☐

Lemma 5.1.52. *Let $[v_1],\dots,[v_{n+1}],[w_1],\dots,[w_{n+1}] \in \mathbb{P}^n$ such that $[v_1],\dots,[v_{n+1}$ resp. $[w_1],\dots,[w_{n+1}]$ do not lie on a hyperplane, i.e. for any $j \in \{1,\dots,n+1\}$, $[v_j] \notin H([v_1],\dots\widehat{[v_j]},\dots,[v_{n+1}]) := \left\{\left[\sum_{i\neq j} \lambda_i v_i\right] : (\lambda_i) \in K^n \setminus 0\right\}$ (the same is true for the $w_i's$). Then there is a projective change of coordinates $\varphi : \mathbb{P}^n \to \mathbb{P}^n$ such that $\varphi[v_i] = [w_j]$*

Proof. By Lemma 2.7.17, we have that $v_1,\dots,v_{n+1}$ resp. $w_1,\dots,w_{n+1}$ are linearly independent. Set $A = (v_{ij})^T$ and $B = (w_{ij})^T$, and pick $\varphi = BA^{-1}$. Then

$$\varphi([v_i]) = [BA^{-1}v_i] = [Be_i] = [w_i].$$

The result in other words amount to a linear change of basis. □

Lemma 5.1.53. *let $L = L([v],[w])$ and $\Lambda = L([u],[r])$ be distinct lines in $\mathbb{P}^2$. Then $L \cap \Lambda$ is a point.*

Proof. WLOG $[u] \notin L$, hence $[r] \in \mathbf{Span}([v],[w],[u])$, implying that there exist a unique(!) solution $(\lambda,\mu,\nu) \in K^3 \setminus 0$ to the equation

$$Ax = r,$$

where $A \in M_3(K)$ has columns v,w,u. We thus find that

$$L \cap \Lambda = \{[\lambda v + \mu w]\} = \{[\nu u + r]\}.$$

□

Lemma 5.1.54. *Let $H_1,\dots,H_m \subset \mathbb{P}^n$, hyperplanes with $m \leq n$. Then $\bigcap_1^m H_i \neq \emptyset$. Write $H_i = V(\sum_1^{n+1} a_{ij}x_j)$. If $(a_{i1},\dots,a_{i,n+1})$ are linearly independent, then $\bigcap_1^m H_i$ is an $n-m$-dimensional linear subvariety. In particular, the intersection of n such hyperplanes is a point.*

Proof. Since $(a_{ij}) : \mathbb{A}^{n+1} \to \mathbb{A}^m$ is a non-zero map it follows by rank-nullity that $\mathbf{dim}\ \bigcap_1^m H_i \geq 0$, hence $\bigcap_1^m H_i \neq \emptyset$. Under the further assumptions that (a_{ij}) have linearly independent, it follows that $\mathbf{dim}\ \bigcap_1^n H_i = \mathbf{dim}\ \mathbf{ker}((a_{ij}) - 1 = n+1-m-1 = n-m$ □

Lemma 5.1.55. *Let $H_1,\dots,H_{n+1}$ (defined by $\sum_1^{n+1} a_{ij}x_j$) and $H'_1,\dots,H'_{n+1}$ be hyperplanes in $\mathbb{P}^n$ that respectively do not all pass through a common point. Then there is a projective change of coordinates $\varphi : \mathbb{P}^n \to \mathbb{P}^n$ such that $\varphi(H_i) = H'_i$.*

Proof. The assumption that $H_1,\dots,H_{n+1}$ resp. $H'_1,\dots,H'_{n+1}$ do not pass through a common point means that $\bigcap_1^{n+1} H_i = \bigcap_1^{n+1} H'_i = \emptyset$, hence (a_{ij}) has null space $\mathbf{0}$, hence any selection of n rows in this matrix will be linearly independent. Then for each $i \in \{1,\dots,n+1\}$, $P_i := \bigcap_{j \neq i} H_i$ (cf. the above lemma) is a point not contained in $H(P_1,\dots,\widehat{P_i},\dots,P_{n+1}) = H_i$, which is also true for $P'_i := \bigcap_{j \neq i} H'_i$. We apply Lemma 5.1.52 to find a projective change of coordinates $\varphi : \mathbb{P}^n \to \mathbb{P}^n$ such that $\varphi(P_i) = P'_i$. By Lemma 5.1.50 it follows that

$$\varphi(H_i) = H(\varphi(P_1),\dots,\widehat{\varphi(P_i)},\dots,\varphi(P_{n+1})) = H(P'_1,\dots,P'_i,\dots,P'_{n+1}) = H'_i$$

□

Proposition 5.1.56. *There is a one-to-one correspondence*

$$\mathbb{P}^n \xleftrightarrow{\quad} \{H \subset \mathbb{P}^n : H \text{ is a hyperplane }\}$$

Proof. We map a point $[v_1,\dots,v_{n+1}]$ to $V\left(\sum_1^{n+1} v_i x_i\right)$ which is well-defined since $I = \lambda I$ for any ideal $I \subset R$, for any ring R, $\lambda \in R^*$. Given a hyperplane $H = V(f)$ where f is a form of degree 1, then $f = \sum_1^{n+1} a_i x_i$ for suitable $a_1,\dots,a_{n+1} \in K^{n+1}$ where at least one $a_i \neq 0$. Then $[a_1,\dots,a_{n+1}] \mapsto V(f)$. If $V\left(\sum_1^{n+1} v_i x_i\right) = V\left(\sum_1^{n+1} v'_i x_i\right)$, then $\sum_1^{n+1} v'_i x_i = \lambda \sum_1^{n+1} v_i x_i$ for some $\lambda \in K \setminus 0$, hence $\lambda(v'_1,\dots,v'_{n+1}) = (v_1,\dots,v_{n+1})$, meaning $[v] = [v']$ □

Remark 5.1.57. Denote the hyperplane corresponding to a point P by $P^\star$ and the point corresponding to a hyperplane H by $H^\star$.

Corollary 5.1.58. *Given a point P and a hyperplane H in $\mathbb{P}^n$, $P^{\star\star} = P$ and $H^{\star\star} = H$. Furthermore $P \in H \iff H^\star \in P^\star$*

Proof. The first statement is obvious since the functions are mutual inverses. For second statement suppose $[v_1,\dots,v_{n+1}] \in H = V\left(\sum_1^{n+1} w_i x_i\right)$. Then

$$\mathrm{ev}_{[w_1,\dots,w_{n+1}]}\left(\sum_1^{n+1} v_i x_i\right) = \sum_1^{n+1} w_i v_i = \mathrm{ev}_{[v_1,\dots,v_{n+1}]}\left(\sum_1^{n+1} w_i x_i\right) = 0 \Rightarrow H^\star = [w_1,\dots,w_{n+1}] \in P^\star.$$

The converse implication follows from the first statement of this corollary and the first implication. □

5.1.6 Affine and Projective Varieties

For this section we denote the identification of $\mathbb{A}^n$ with $U_{n+1} \subset \mathbb{P}^n$ by φ_{n+1}. Recall that this is given by $v \mapsto [v_1,\dots,v_n,1]$

Definition 5.1.59. Given an affine algebraic set $V = V(I) = V(f_1,\dots,f_m) \subset \mathbb{A}^n$ we define the *homogenization of* V to be the projective algebraic set

$$V^* := V(I^*) \subset \mathbb{P}^n.$$

Conversely given a projective algebraic set $V = V(I) = V(F_1,\dots,F_m) \subset \mathbb{P}^n$, $I \subset K[x_1,\dots,x_{n+1}]$, and $F_1,\dots,F_m \in K[x_1,\dots,x_{n+1}]$ are homogeneous, we define the *dehomogenization of* V to be the affine algebraic set

$$V_* := V(I_*) = V((F_1)_*,\dots,(F_m)_*)$$

Example 5.1.60. With the same setup as Example 2.9.114 where $R = K$ and consider $V := V(I)$. Note that $V = \{(a,a^2,a^3) : a \in K\}$. Let $f \in I(V)$, Then $f(a,a^2,a^3) = 0$ for any $a \in K$, hence $f(x,x^2,x^3) = 0$. In other words $(x,x^2,x^3) \in K[x,y,z]^3 = K[x][y,z]^3$ is a zero of f and hence $f \in \langle x - x, y - x^2, z - x^3\rangle = I$ by Proposition 2.9.38.
This illustrates that homogenization of a variety cannot be written as the vanishing set of the homogenization of the finite polynomials defining the affine variety.

Lemma 5.1.61. *Let $V \subset \mathbb{A}^n$ be algebraic. Then $I(V)^* \subset I(V^*)$. We also have that $\varphi_{n+1}(X_*) \subset X$ for $X = V(I) \subset \mathbb{P}^n$ algebraic.*

Proof. This follows from $V^* = V(I(V)^*)$. Let $[v,1] \in \varphi_{n+1}(X_*) = \varphi_{n+1}(V(I_*))$ and $f \in I(X)$. Then $f_* \in I(X)_* \supset I_*$, hence $f(v,1) = f_*(v)$ $\square$

Lemma 5.1.62. *Let $V = V(I), W = V(J) \subset \mathbb{A}^n$, $X, Y \subset \mathbb{P}^n$ be algebraic sets. We have the following results:*

1. *$\varphi_{n+1}(V) = V^* \cap U_{n+1}$ and $(V^*)_* = V$, $(X_*)^* \subset X$.*
2. *If $V \subset W$, then $V^* \subset W^*$. If $X \subset Y \subset$, then $X_* \subset Y_*$.*
3. *If V is a variety so is V^*.*
4. *$V^* = \overline{\varphi_{n+1}(V)}$, where $\overline{\bullet}$ is the Zariski-closure in $\mathbb{P}^n$.*
5. *If $\bigcup_1^k V_i$ is the decomposition of V into affine varieties, then $\bigcup_1^k V_i^*$ is the decomposition of V^* into projective varieties.*
6. *If $\emptyset \neq V \subsetneq \mathbb{A}^n$, then no component of V^* lies in H_∞ or contains it.*
7. *If an algebraic $X \subset \mathbb{P}^n$ and no component of X is in or contains H_∞, then $X_* \subsetneq \mathbb{A}^n$ and $(X_*)^* = X$.*

Proof. 1. Let $v \in V$, then $\varphi(V) \in U_{n+1}$. Let $f^* \in I^*$, $f \in I$. Then for $\lambda \in K \setminus 0$,

$$f^*(\lambda(v_1, \ldots, v_n, 1)) = \lambda^d (f^*)_*(v) = \lambda^d f(v) = 0 \Rightarrow f^*(\varphi(v)) = 0 \Rightarrow \varphi(v) \in V^* \cap U_{n+1}.$$

Conversely, if $[v_1, \ldots, v_n, 1] \in V^*$, then by the same computation $f(v) = f^*([v_1, \ldots, v_n, 1]) = 0 \Rightarrow v \in V \Rightarrow \varphi(v) \in \varphi(V)$.
The second equality follows from $(I^*)_* = I$. The second inclusion follows from $f = x_{n+1}^r (f_*)^*$ for some $r \geq 0$ for every $f \in I(V)$. 2. Since $I(V) \supset I(W)$, we have $I(V)^* \supset I(W)^*$ implying $V^* = V(I(V)^*) \subset V(I(W)^*) = W^*$. The second statement follows from $I(X) \supset I(Y) \Rightarrow I(X)_* \subset I(Y)_*$. 3. Since I is prime I^* is prime, hence $V^* = V(I^*)$ is a variety.
4. Let $U \subset \mathbb{P}^n$ be algebraic with $\varphi_{n+1}(V) \subset W$. Let $f \in I(W)$, then $f \in I(\varphi_{n+1}(V))$, hence f_* vanishes on every $[v, 1]$ where $v \in V$, implying $f_* \in I(V)$. Then $(f_*)^* \in I(V)^*$, hence $f = x_{n+1}^r (f_*)^* \in I(V)^*$ for some $r \geq 0$. We therefor see that $I(U) \subset I(V)^*$, hence $V = V(I(V)^*) \subset V(I(U)) = U$.
5. 4. and 2. implies $V^* = (\bigcup V_i)^* = \bigcup V_i^*$. 3. implies each V_i^*. We have that $(V_i^*)_* = V_i \not\subset V_j = (V_i^*)_*$ implying that $V_i^* \not\subset V_j^*$ by the second statement of 2.
6. Let C be a component of V. Then $\neq C \subsetneq \mathbb{A}^n$. Then by 1. $C^* \cap U_{n+1} = \varphi_{n+1}(C) \neq \emptyset$, hence $C^* \not\subset H_\infty$. Suppose $\emptyset \neq B \subset \mathbb{A}^n$ is algebraic such that $B^* \supset H_\infty$. Then $I(B)^* \subset I(B^*) \subset I(H_\infty) = \langle x_{n+1} \rangle$ by PNS. Pick $f \in I(B)$. Then $f^*(v, 0) = 0$ for every $v \in \mathbb{A}^n$. This correspondence to the highest degree terms of f being 0, implying $f = 0$, hence $I(B) = 0$, meaning $B = \mathbb{A}^n$. We thus have that $V \not\supset H_\infty$.
7. If $X_* = \mathbb{A}^n$, $X \supset (X_*)^* = \mathbb{P}^n \supset H_\infty$, hence trivially $X_* \subsetneq \mathbb{A}^n$. Let $Z = V(I)$ be a component of X. Then no component of Z lies in H_∞. By 1. it is sufficient to check that $Z \subset (Z_*)^*$, meaning it is sufficient to check that $I(Z) \supset I(Z_*)^*$, since in general $I(V)^* \subset I(V^*)$ implies that $V(I(V)^*) \supset V^*$. Let $f \in I(Z_*)$. Then, since PNS tells us that $I(Z_*) = \mathrm{rad}(I_*) = \mathrm{rad}(I)_* = I(Z)_*$, we get that for some $N \geq 0$, $f^N \in I(Z)_*$, hence for some $r \geq 0$, $x_{n+1}^r (f^n)^* \in I(V)$. Note that $I(Z)$ is prime and $x_{n+1} \notin I(Z)$ since $Z \not\subset H_\infty$, we get that $f^* \in I(Z)$, hence $I(Z_*)^* \subset I(Z)$. □

Remark 5.1.63. V^* is called the *projective closure of* V.

Lemma 5.1.64. *Let $V \subset W \subset \mathbb{P}^n$ be varieties, where $V = V(f)$ is a hypersurface (where f is a form). Then $V = W$ or $W = \mathbb{P}^n$*

Proof. The case $f \in K$ is trivial. So suppose $\deg f \geq 1$. Since V is a variety $I(V) = \langle f \rangle$ and f is prime and therefor irreducible. Then $\langle f \rangle \supset I(W)$, hence if $g \in I(W)$, then $g = qf$. Then $I(W) = \langle f \rangle J$ for some $J \subset K[x_1, \ldots, x_{n+1}]$. Then $W = V(f) \cup V(J)$, hence $W = V(f) = V$ or $W = V(J)$. Note that if J is non-trivial, then $V(J)$ and V are components of W, hence $J = 0$ or $J = K[\mathbf{x}]$. Only the first case is possible and in this case $W = \mathbb{P}^n$. □

Lemma 5.1.65. *Let $H_\infty \subset V \subset \mathbb{P}^n$ be a variety. Then $V = H_\infty$ or $V = \mathbb{P}^n$. If $\mathbb{P}^n_* = \mathbb{A}^n$ and if $(H_\infty)_* = \emptyset$.*

Proof. The first statement follows from the prior lemma. $\mathbb{P}^n_* = V^{\mathbb{A}}(0_*) = V^{\mathbb{A}}(0) = \mathbb{A}^n$ and $(H_\infty)_* = V^{\mathbb{A}}((x_{n+1})_*) = V^{\mathbb{A}}(1) = \emptyset$. □

Remark 5.1.66. There is a one-to-one correspondence between varieties in $\mathbb{P}^n$ that do not lie in H_∞ and non-empty varieties in $\mathbb{A}^n$, established via $V \mapsto V_*$ with mutual inverse $\emptyset \neq W \mapsto W^*$ (cf. statements 6. and 7. of the main lemma and the above lemma). The map

$$\Gamma(\mathbb{P}^n) = K[x_1, \ldots, x_{n+1}] \to K[x_1, \ldots, x_n] = \Gamma(\mathbb{A}^n), f \mapsto f_*,$$

is surjective with kernel $\langle x_{n+1} - 1\rangle$. Therefor $\Gamma(\mathbb{P}^n)/\langle x_{n+1} - 1\rangle \simeq \Gamma(\mathbb{A}^n)$. We thus establish that $K^h(V)/\langle x_{n+1} - 1\rangle K^h(V) \simeq Q(\Gamma^h(V)/\langle x_{n+1} - 1\rangle) \simeq K(V_*)$. Let $f \in K[x_1, \ldots, x_n]$. Then $x_{n+1}^d f^* + \langle x_{n+1} - 1\rangle = f + \langle x_{n+1} - 1\rangle$. We therefor get a surjective K-algebra map,

$$K(V) \to Q(\Gamma^h(V)/\langle x_{n+1} - 1\rangle), \frac{f}{g} \mapsto \frac{f + \langle x_n - 1\rangle}{g + \langle x_{n+1} - 1\rangle}.$$

It is clearly also surjective, since for a form $F \in K[x_1, \ldots x_{n+1}]$, $\mathrm{ev}_{x_{n+1} \mapsto 1}(F) = 0$ if and only if $F = 0$, hence $F \in \langle x_{n+1} - 1\rangle$ if and only if $F = 0$. Therefor $K(V) \simeq K^h(V)/K^h(V)\langle x_{n+1} - 1\rangle \simeq K(V_*)$. Suppose V is a variety not containing or contained in H_∞. Consider the map

$$K[x_1, \ldots, x_{n+1}] \to K[x_1, \ldots, x_n]/I(V_*), f \mapsto f_* + I(V_*),$$

Since $I(V)_* = I(V_*)$ we establish an identification of $\Gamma^h(V) \simeq \Gamma(V_*)$. This induces an isomorphism $K^h(V) \simeq K(V_*)$. We claim that $K^h(V) = K(V)$. We note that $x_{n+1} + I(V) \mapsto 1 + I(V_*)$, implying $x_{n+1} + I(V) = 1 + I(V)$ hence $\langle x_{n+1} - 1\rangle \subset I(V)$. If $f + I(V) \in \Gamma^h(V)$, then $f + I(V) = f_* + I(V) = x_{n+1}^r (f_*)^* + I(V)$ for any sufficiently large r. Then for $\frac{\alpha}{\beta} \in K^h(V)$, we may find representatives for α and β that are homogeneous of the same degree, meaning $\frac{\alpha}{\beta} \in K(V)$. This means $K(V) = K^h(V) \simeq K(V_*)$.
The above is just a way of saying that the map

$$K(V) \to K(V_*), \frac{a + I(V)}{b + I(V)} \mapsto \frac{a_* + I(V_*)}{b_* + I(V_*)}$$

is an isomorphism. Although not strictly necessary, I feel that this detour somewhat illuminating, simply due to the fact that along the way, we saw the exact relationship between $\Gamma^h(V)$ and $\Gamma(V_*)$, $K^h(V)$, $K(V)$ and $K(V_*)$ in any possible case.

Let $P = [v_1,\dots,v_n,1] \in V \cap U_{n+1}$. Then if $\frac{\alpha}{\beta} \in K(V)$ is defined at P, then $0 \neq \beta(P) = \beta_*(v_1,\dots,v_n)$, hence $\frac{\alpha_*}{\beta_*} \in K(V_*)$ is defined at $(v_1,\dots,v_n)$ as well. Therefor $\mathcal{O}_P(V) \simeq \mathcal{O}_{(v_1,\dots,v_n)}(V)$.

Example 5.1.67. 1. A quick note on the algebraic subsets of $\mathbb{P}^1$. Consider such a set $V = V(F_1,\dots,F_m)$ where F_i are non-zero forms. We may write $F_i = \prod_1^{d_i} L_{ij}$ for linear forms L_{ij}. Then

$$X = \bigcup_{(i_1,\dots,i_m)} V(L_{1i_1},\dots,L_{mi_m}).$$

Note that the vanishing set of a linear form in $\mathbb{P}^1$ is a point. For each $(i_1,\dots,i_m)$, if $L_{1i_1},\dots,L_{mi_m}$ are distinct (up to multiplication by a unit), then $V(L_{1i_1},\dots,L_{mi_m}) = \emptyset$. If they are equal then $V(L_{1i_1},\dots,L_{mi_m})$ is a point. The algebraic subsets of $\mathbb{P}^1$ are therefor $\mathbb{P}^1, \emptyset$ or finite union of points.

2. By the above, we conclude that the varieties of $\mathbb{P}^1$ are therefor $\mathbb{P}^1, \emptyset$ and hyperplanes, which in $\mathbb{P}^1$ are the same as points

3. Suppose $\emptyset \neq V \subset \mathbb{P}^2$ is a variety that is not contained or does not contain H_∞. Then $\emptyset \neq V_* \subsetneq \mathbb{A}^n$ and $(V_*)^* = V$. Then V_* is a variety since V_* has the same number of components as V. This means $V_* = V(f)$ where $f \in K[x,y]$ for $\deg f \geq 1$ or $V_* = \{x-a, y-b\}$ for $a,b \in K$, hence $V = V(f^*)$ or $V = V(x - az, y - bz)$. If V contains H_∞, $V = H_\infty$ or $V = \mathbb{P}^2$. Suppose $V = V(I) \subsetneq H_\infty$. Note that $V(I)$ is in bijection with $W = \{[v_1,v_2] \in \mathbb{P}^1 : \forall f \in I, f(v_1,v_2,0)\}$, which is a finite set for otherwise $V(I) = H_\infty$. Then in particular $\#V = 1$, hence V is a point $V(bx - ay, z)$ where $[a,b] \in \mathbb{P}^1$. In conclusion the varieties in $\mathbb{P}^2$ are $\mathbb{P}^2 = (\mathbb{A}^2)^*, \emptyset$, $V(f^*)$, where $f \in K[x,y]$ is irreducible, $V(I(P)^*)$ where P is a point in $\mathbb{A}^2$, $V = H_\infty$, or $V(bx - ay, z) \subsetneq H_\infty$ for $[a,b] \in \mathbb{P}^1$.

Example 5.1.68. 1. Let's classify the lines through $P = [0,1,0] \in \mathbb{P}^2$. Note that $\{P\} = V(x,z)$. Let $l = ax + by + cz$. Suppose $L := V(l)$ passes through P, i.e. that $l(0,1,0) = 0$. Then $b = 0$, hence $l = ax + cz$, where $[a,c] \in \mathbb{P}^1$. If $a = 0$, then $L = L_\infty$. If $a \neq 0$, then L is a *vertical line*, i.e. for a suitable $\lambda \in K$ $L = L_\lambda := V(x - \lambda z) = \{[\lambda, t, 1] : t \in K\} \cup \{P\}$. In other words, L_λ is the vertical line $x = 1$ in the affine xy-plane with an added point P on the line at infinity.

2. Consider an arbitrary collection of points $P_1,\ldots,P_m \in \mathbb{P}^n$. Then since there is a one-to-one correspondence between $\mathbb{P}^n$ and the hyperplanes in $\mathbb{P}^n$ via the map $P \mapsto P^\star$. Then pick a point P not in $P_i^\star$ for any i. We may do this since the intersection of proper subsets (of $\mathbb{P}^n$) itself is a proper subset (of $\mathbb{P}^n$). Then $P_i = P_i^{\star\star} \notin P^\star$. Here we use Corollary5.1.58 a few times together with the fact that $\mathbb{P}^n$ is an infinite set for $n \geq 1$. It should be noted that one can deduce the fact from more simple principles.

5.2 Multiprojective Algebraic Sets & Multiprojective Varieties

5.2.1 Basic Definitions

Definition 5.2.1. Let $S \subset K[x_{ij} : i \in \{1,\ldots,m\}, j \in \{1,\ldots,n_i+1\}, n_i \geq 1]$. We define the *multiprojective vanishing set* S to be

$$V^{\mathrm{mul}\mathbb{P}}(S) := V(S) := \left\{([v_1],\ldots,[v_m]) \in \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m} : f([v_1],\ldots,[v_m]) = 0 \text{ for all } f \in S\right\}.$$

A set $X \subset \mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_m}$ is called a *(multiprojective) algebraic set* if it the multiprojective vanishing set of some subset of $K[\mathbf{x}]$. X is a *(multiprojective) hypersurface* if $X = V(F)$ for some m-homogeneous $F \in K[\mathbf{x}]$. If $\deg F = 1$ it is called a *(multiprojective) hyperplane*.

Remark 5.2.2. By Lemma 2.9.123 we may assume that any multiprojective algebraic set may be written as the multiprojective vanishing set of some m-homogeneous ideal, i.e. as the multiprojective vanishing set of a finite set of m-forms. The system of multiprojective algebraic sets form the *Zariski topology* on $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ with $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m} = V^{\mathrm{mul}\mathbb{P}}(0)$ and $\emptyset = V^{\mathrm{mul}\mathbb{P}}(K[\mathbf{x}])$. Vanishing sets interact with $\subset$ as expected.

Definition 5.2.3. Let $X \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$. We define *the ideal of* X to be

$$I^{\mathrm{mul}\mathbb{P}}(X) := I(X) := \left\{f \in K[\mathbf{x_1},\ldots,\mathbf{x_m}] : f([v_1],\ldots,[v_m]) = 0 \text{ for every } ([v_1],\ldots,[v_m]) \in X\right\}.$$

Remark 5.2.4. $I^{\mathrm{mul}\mathbb{P}}(X)$ is an m-homogeneous ideal. The interactions of $I^{\mathrm{mul}\mathbb{P}}(\bullet)$ with $\subset$ and $V^{\mathrm{mul}\mathbb{P}}(\bullet)$ are as expected.

Definition 5.2.5. An algebraic set $X \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ is called *reducible* if there a proper algebraic subsets of X, Y,Z say, such that $X = Y \cup Z$. If an algebraic set $V \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ is not reducible, it is called *irreducible* or is said to be a *multiprojective variety.*

Remark 5.2.6. As expected any multiprojective set has a unique decomposition into multiprojective varieties. An algebraic set V is a variety if and only if $I^{\mathrm{mul}\mathbb{P}}(V)$ is prime.

Definition 5.2.7. For a variety $V \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ we define the cone of V to be

$$C(V) := \left\{(v_1, \dots, v_m) \in \mathbb{A}^{n_1+1} \times \cdots \times \mathbb{A}^{n_m+1} : ([v_1], \dots, [v_m]) \in V \text{ or } v_i = 0\right\}$$

Remark 5.2.8. Lemmas 5.1.20 and 5.1.21 are easily generalizable to the multiprojective case.

There is a natural way of "embedding" a projective algebraic set $V \subset \mathbb{P}^{n_i}$ in $\mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ by considering

$$\mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_{i-1}} \times V \times \mathbb{P}^{n_{i+1}} \times \cdots \mathbb{P}^{n_m}.$$

We do find that if $V = V^{\mathbb{P}}(I)$, then $\mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_{i-1}} \times V \times \mathbb{P}^{n_{i+1}} \times \cdots \mathbb{P}^{n_m} = V^{\mathrm{mul}\mathbb{P}}(I)$. In general, given $V_i = V(I_i) \subset \mathbb{P}^{n_i}$ algebraic sets, we have that

$$V^{\mathrm{mul}\mathbb{P}}(\langle \bigcup I_i \rangle) = V_1 \times \cdots \times V_m \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}.$$

Lemma 5.2.9. *Let I be an m-homogeneous ideal in $K[\mathbf{x_1}, \dots, \mathbf{x_m}]$. The following are equivalent.*

1. $V^{\mathrm{mul}\mathbb{P}}(I) = \emptyset$.

2. $V^{\mathbb{A}}(I) \subset \{0\}$.

3. $\mathrm{rad}(I) = I^{\mathbb{A}}(V^{\mathbb{A}}(I)) \supset \langle \mathbf{x_1}, \dots \mathbf{x_m} \rangle$.

4. $\langle \mathbf{x_1}, \dots \mathbf{x_m} \rangle^D \supset I$ *for a suitable large* $D \geq 0$.

Proof. The proof is the same as the singly projective version using proper generalizations of the results involved □

Theorem 5.2.10. *Let $I \subset K[\mathbf{x_1}, \dots, \mathbf{x_m}]$ be an m-homogeneous ideal. Then*

1. $V^{\mathrm{mul}\mathbb{P}}(I) = \emptyset$ *if and only if* $\langle \mathbf{x_1}, \dots, \mathbf{x_m} \rangle^D \subset I$ *for some* $D \geq 0$.

2. *If* $V^{\mathrm{mul}\mathbb{P}}(I) \neq \emptyset$, *then* $I^{\mathrm{mul}\mathbb{P}}(V^{\mathrm{mul}\mathbb{P}}(I)) = \mathrm{rad}(I)$.

Proof. 1. follows from the prior lemma. 2. follows from the generalizations of the results about cones. □

Definition 5.2.11. We define the *multiprojective coordinate ring of a variety* V to be the integral domain.

$$\Gamma^m(V) := K[\mathbf{x_1},\dots,\mathbf{x_m}]/I^{\mathrm{mul}\mathbb{P}}(V)$$

We set $K^b(V) := Q(\Gamma^m(V))$. We define *the function field of* V to be

$$K(V) := \left\{ z \in K^b(V) : z = \frac{f}{g} \text{ for suitable } m\text{-forms } f, g \in \Gamma^m(V) \right\}.$$

We say for an element $z \in K(V)$ and a point $P \in \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ that z *is defined at* P if there are m-forms of the same m-degree, $\alpha, \beta \in \Gamma^m(V)$, such that $z = \frac{\alpha}{\beta}$ and P is not a zero of β. We define *the local ring of* V *at* P to be

$$\mathcal{O}_P(V) = \{z \in K(V) : z \text{ defined at } P\}$$

Remark 5.2.12. Suppose $z = \frac{\alpha}{\beta} \in K(V)$. Then for a $P = ([v_1],\dots,[v_m]) \in \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ with $\beta(v_1,\dots,v_n) \neq 0$. We have for each $\lambda_1,\dots,\lambda m \in K \setminus 0$ that

$$\frac{\alpha(\lambda_1 v_1,\dots,\lambda_m v_m)}{\beta(\lambda_1 v_1,\dots,\lambda_m v_m)} = \frac{\prod_1^m \lambda_i^{d_i}}{\prod_1^m \lambda_i^{d_i}} \frac{\alpha(v_1,\dots,v_m)}{\beta(v_1,\dots,v_m)} = \frac{\alpha(v_1,\dots,v_m)}{\beta(v_1,\dots,v_m)}.$$

Hence we get a well-defined evaluation at P of multiprojective rational functions defined at P, given by $\mathcal{O}_P(V) \ni \frac{\alpha}{\beta} \mapsto \frac{\alpha(v_1,\dots,v_m)}{\beta(v_1,\dots,v_m)}$.

Definition 5.2.13. For $(i_1,\dots,i_m)$, define

$$\varphi_{i_1,\dots,i_m} : \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m} \to U_{i_1} \times \cdots \times U_{i_m} \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$$

$$(v_1,\dots,v_m) \mapsto ([v_{11},\dots,v_{1,i_1-1},1,v_{1,i_1+1},\dots,v_{1n_1}],\dots,[v_{m1},\dots,v_{m,i_m-1},1,v_{m,i_m+1},\dots,v_{mn_m}])$$

Remark 5.2.14. In this way we identify $U_{i_1} \times \cdots \times U_{i_m}$ with $\mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$.

Definition 5.2.15. The *multiprojective closure* of an affine algebraic set $V = V(I) \subset \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$ is the set

$$V^* := V(I^*)$$

where for an $f \in K[x_{ij} : i \in \{1,\dots,m\}, j \in \{1,\dots,n_i\}]$, f^* denotes polynomial obtained from the following recursive process: set $f_0 = f$ and $f_{i+1} = f_i^*$ where $f_i \in K[\mathbf{x_1},\dots,\widehat{\mathbf{x_i}},\dots,\mathbf{x_m}, x_{n_1+1},\dots,x_{n_{i-1}+1}][\mathbf{x_i}]$ for $i \in \{1,\dots,m\}$. We therefor define $I^* := \langle \{f^* : f \in I\} \rangle$.

For a polynomial $f \in K[x_{ij} : i \in \{1,\dots,m\}, j \in \{1,\dots,n_i+1\}]$ we define f_* to be the image of f under the evaluation map taking x_{i,n_i+1} to 1 and x_{ij} to x_{ij} for every other pair of indices.

Lemma 5.2.16. *Let $V = V(I), W = V(J) \subset \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$, $X, Y \subset \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ be algebraic sets. We have the following results:*

1. *$\varphi_{n_1+1,\dots,n_m+1}(V) = V^* \cap U_{n_1+1} \times \cdots \times U_{n_m+1}$ and $(V^*)_* = V$, $(X_*)^* \subset X$.*

2. *If $V \subset W$, then $V^* \subset W^*$. If $X \subset Y \subset$, then $X_* \subset Y_*$.*

3. *If V is a variety so is V^*.*

4. *$V^* = \overline{\varphi_{n_1+1,\dots,n_m+1}(V)}$, where $\overline{\bullet}$ is the Zariski-closure in $\mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_m}$.*

5. *If $\bigcup_1^k V_i$ is the decomposition of V into affine varieties, then $\bigcup_1^k V_i^*$ is the decomposition of V^* into projective varieties.*

6. *If $\emptyset \neq V \subsetneq \mathbb{A}^n$, then no component is contained in or contains $\mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_{i-1}} \times H_\infty^i \times \mathbb{P}^{n_{i+1}} \times \cdots \times \mathbb{P}^{n_m}$ for each i where $H_\infty^i := V^{\mathbb{P}}(x_{i,n_i+1})$.*

7. *If $X \subset \mathbb{P}^n$ is an algebraic set lies not in or is contained in any $\mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_{i-1}} \times H_\infty^i \times \mathbb{P}^{n_{i+1}} \times \cdots \times \mathbb{P}^{n_m}$ for each i, then $X_* \subsetneq \mathbb{A}^n$ and $(X_*)^* = X$.*

Proof. 1. If $f \in I^a st$, then $f(\varphi_{n_1+1,\dots,n_m+1}(v_1,\dots,v_m)) = f([v_1,1],\dots,[v_m,1]) = f_*(v_1,\dots,v_m) = 0$ for every $\varphi(v_1,\dots,v_m) \in \varphi_{n_1+1,\dots,n_m+1}(V)$. Conversely if $v \in V^* \cap U_{n_1+1} \times \cdots \times U_{n_m+1}$. Then for a $f \in I$, $f(v_1,\dots,v_m) = f^*([v_1,1],\dots,[v_m,1]) = 0$, hence $v = \varphi(v_1,\dots,v_m)$. It is fairly obvious that $(I^*)_*$, hence the second equality follows. Similarly $f = \left[\prod_1^m x_{n_i+1}^{r_i}\right](f_*)^*$, hence we get the inclusion of $(X_*)^*$ in X.
2. - 5. follow from the exact same arguments using the proper generalizations.
6. Let C be a component of V. Then $\emptyset \neq C \subsetneq \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$. Then $C^* \cap U_{n_1+1} \times \cdots \times U_{n_m+1} = \varphi_{n_1+1,\dots,n_m+1}(C) \neq \emptyset$. Suppose $\emptyset \neq B \subset \mathbb{A}^n$ is algebraic such that $B^* \supset \mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_{i-1}} \times H_\infty^i \times \mathbb{P}^{n_{i+1}} \times \cdots \times \mathbb{P}^{n_m}$ for some i. Then $I(B)^* \subset I(B^*) \subset \langle x_{n_i+1} \rangle$ by the generalization of PNS. Let $f \in I(B) \subset K[\mathbf{y_j} : j \neq i][\mathbf{y_i}]$. Then $f^*([v_i,0]) = 0$ for every $v_i \in \mathbb{A}^{n_i}$. Suppose d_i was the $\mathbf{y_i}$-degree of f. Writing $f = f_{d_i} + \sum_{j \leq d_i} f_j$, we see that

$$f^* = f_{d_i} + \sum_{j \leq d_i} x_{n_i+1}^{d_i-j} f_j.$$

The vanishing condition on f^*, hence corresponds to $f_{d_i} = 0$, implying $I(B) = 0$, meaning $B = \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$. We therefor conclude that $V \not\supset \mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_{i-1}} \times H_\infty^i \times \mathbb{P}^{n_{i+1}} \times \cdots \times \mathbb{P}^{n_m}$. 7. The case $X_* = \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$ is trivially not possible under the assumptions. We thus have that $X_* \subsetneq \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$. Let $Z = V^{\mathrm{mul}\mathbb{P}}(a)$ be a component of X. Then Z lies not in $\mathbb{P}^{n_1} \times \cdots \mathbb{P}^{n_{i-1}} \times H_\infty^i \times \mathbb{P}^{n_{i+1}} \times \cdots \times \mathbb{P}^{n_m}$ for any i. It is sufficient to prove that $Z \subset (Z_*)^*$ and therefor sufficient to check that $I(Z) \supset I(Z_*)^*$

in $K[\mathbf{x_1},\dots,\widehat{\mathbf{x_i}},\dots,\mathbf{x_m}][\mathbf{x_i}]$. For an $f\in I(Z_*)$ we utilize the generalization of PNS to get that for suitably large $N\geq 0$, and for suitable r, $x_{n_i+1}^r(f^N)^*\in I(Z)$, using the fact that $x_{n_i+1}\notin I(Z)$ and that Z is irreducible we get that $f^*\in I(Z)$. □

Lemma 5.2.17. *Let $V\subset W\subset \mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_m}$ be multiprojective varieties such that $V=V^{\mathrm{mul}\mathbb{P}}(f)$ for some m-form f. Then $V=W$ or $W=\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_m}$. It follows that if $\mathbb{P}^{n_1}\times\cdots\mathbb{P}^{n_{i-1}}\times H_\infty^i\times\mathbb{P}^{n_{i+1}}\times\cdots\times\mathbb{P}^{n_m}\subset W\subset\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_m}$, then $W=\mathbb{P}^{n_1}\times\cdots\mathbb{P}^{n_{i-1}}\times H_\infty^i\times\mathbb{P}^{n_{i+1}}\times\cdots\times\mathbb{P}^{n_m}$ or $W=\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_m}$*

Proof. The proof of the first statement is identical to the projective case, the second statement immediately follows from the first due to $\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_{i-1}}\times H_\infty^i\times\mathbb{P}^{n_{i+1}}\times\cdots\times\mathbb{P}^{n_m}=V(x_{i,n_i+1})$. □

5.2.2 Algebraic Geometry in Multispaces

We extend the theory to so-called *multispaces* which are defined to be $\mathbb{A}^n\times\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_m}$ for $n,n_1,\dots,n_m\geq 1$ with the convention that $\mathbb{A}^0$ is a point.

Definition 5.2.18. We say that a point $P=(v,[v_1],\dots,[v_m])\in\mathbb{A}^n\times\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_m}$ is a *zero* of a polynomial $f\in K[x_1,\dots,x_n,\mathbf{y_1},\dots,\mathbf{y_m}]$. If for every $\lambda_1,\dots,\lambda_m\in K\setminus 0$

$$f(v,\lambda_1 v_1,\dots,\lambda_m v_m)=0$$

and we write $f(P)=0$.

Remark 5.2.19. If $f\in K[x_1,\dots,x_n][\mathbf{y_1},\dots,\mathbf{y_m}]$ is an m-form, then if $(v,v_1,\dots,v_m)\in\mathbb{A}^n\times\mathbb{A}^{n_1+1}\times\cdots\mathbb{A}^{n_m+1}$ is a zero of f so is $(v,[v_1],\dots,[v_m])\in\mathbb{A}^n\times\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_m}$.

Definition 5.2.20. Let $S\subset K[x_1,\dots,x_n,\mathbf{y_1},\dots,\mathbf{y_m}]$. We define the *vanishing set of S* in multispace to be

$$V(S):=\left\{P\in\mathbb{A}^n\times\mathbb{P}^{n_1}\times\cdots\times\mathbb{P}^{n_m}:\forall f\in S, f(P)=0\right\}.$$

A subset X in a multispace is said to be *algebraic* if it is the vanishing set of such an S

Remark 5.2.21. Again it is readily verifiable $V(S)=V(\langle S\rangle)$. By Lemma 2.9.123 we may therefor assume that any algebraic set is the vanishing set of some ideal $I\subset K[\mathbf{x},\mathbf{y_1},\dots,\mathbf{y_m}]$ such that $I\subset K[\mathbf{x}][\mathbf{y_1},\dots,\mathbf{y_m}]$ is m-homogeneous. Hence we may assume that an algebraic set arises as the vanishing set of $f_1,\dots,f_l\in K[\mathbf{x},\mathbf{y_1},\dots,\mathbf{y_m}]$ such that when seen as elements of $K[\mathbf{x}][\mathbf{y_1},\dots,\mathbf{y_m}]$ these are m-forms. The algebraic sets define closed sets in a *Zariski topology* on multispaces and interacts with $\subset$ as expected. This therefor lays the ground for a natural generalization of the affine and multiprojective (therefor also projective) theory we have already developed.

To completely reconcile the above claim of generalization we also introduce the ideal of subsets of multispace.

Definition 5.2.22. Let $X \subset \mathbb{A}^n \times \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$. The *ideal of* X is the ideal

$$I(X) = \{f \in K[\mathbf{x}, \mathbf{y_1}, \ldots, \mathbf{y_m}] : \forall P \in X, f(P) = 0\}.$$

Remark 5.2.23. If X is algebraic, then $I(X) \subset K[\mathbf{x}][\mathbf{y_1}, \ldots, \mathbf{y_m}]$ is m-homogeneous. $I(\bullet)$ interacts with $V(\bullet)$ and $\subset$ in the way one expects.

Definition 5.2.24. An algebraic set in multispace is *reducible* if it is the union of two proper algebraic subsets. An algebraic that is not reducible is called *irreducible* and is a *variety.*

Remark 5.2.25. An algebraic subset of multispace is a variety if and only if the ideal of the subset is prime. There is a unique decomposition of an algebraic set in multispace into a union of varieties that are not subsets of each other.

We could define the cone of a algebraic set in multispace and write generalizations of proper lemmas to prove. Things to check!.

Theorem 5.2.26. *Let $I \subset K[\mathbf{x}][\mathbf{y_1}, \ldots, \mathbf{y_m}]$ be m-homogeneous. Then $I(V(I)) = \mathrm{rad}(I)$.*

Definition 5.2.27. For $(i_1, \ldots, i_m)$ we define

$$\begin{aligned} \varphi_{i_1, \ldots, i_m} : \mathbb{A}^n \times \mathbb{A}^{n_1} \times \cdots \mathbb{A}^{n_m} &\to \mathbb{A}^n \times U_{i_1} \times \cdots \times U_{i_m} \\ (v, v_1, \ldots, v_m) &\mapsto (v, \varphi_{i_1, \ldots, i_m}(v_1, \ldots, v_m)) \end{aligned}$$

We homogenize an element f in $K[\mathbf{x}, \mathbf{y_1}, \ldots, \mathbf{y_m}]$ by homogenizing f with respect to each set of the y-variables which we denote f^*. We homogenize an ideal in the obvious way and take the projective closure in the obvious way.

Remark 5.2.28. A generalization of Proposition 5.2.16 goes through. The addition of an affine coordinates changes nothing in the approach of the proof. The conditions in 6. and 7. of this result should be given in terms of $\mathbb{A}^n \times \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_{i-1}} \times H^i_\infty \times \mathbb{P}^{n_{i+1}} \times \cdots \times \mathbb{P}^{n_m}$. What we obtain from this result is, given a variety $V \subset \mathbb{A}^n \times \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$, V^* defines a variety in $\mathbb{A}^n \times \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ which is neither contained or contains

$V(x_{n_i+1})$ for any i. If $V = \mathbb{A}^n \times \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$ then $V^* = \mathbb{A}^n \times \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$. Conversely given a variety $Z \subsetneq \mathbb{A}^n \times \mathbb{P}^{n_1} \times \mathbb{P}^{n_m}$ satisfying the conditions of 7. we get that X_* is a non-empty variety, not equal to $\mathbb{A}^n \times \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$, and if $X = \mathbb{A}^n \times \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$, then $X_* = \mathbb{A}^n \times \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$. The maps $\bullet^*$ and $\bullet_*$ are mutual inverses under the restrictions of 6. and 7. Hence we get a one-to-one correspondence between non-empty varieties in $\mathbb{A}^n \times \mathbb{A}^{n_1} \times \cdots \times \mathbb{A}^{n_m}$ and varieties in $\mathbb{A}^n \times \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ intersecting $\mathbb{A}^n \times U_{n_1+1} \times \cdots \times U_{n_m+1}$.

Definition 5.2.29. Let V be a variety in multispace. We define the *coordinate ring of* V to be

$$\Gamma(V) := K[\mathbf{x}, \mathbf{y_1}, \ldots, \mathbf{y_m}]/I(V).$$

Define the *ring of ration functions on* V to be

$$K(V) := \left\{ z = \frac{\alpha}{\beta} \in Q(\Gamma(V)) : \alpha, \beta \text{ are } m\text{-forms of the same degree} \right\}.$$

In the above we view $I(V) \subset K[\mathbf{x}][\mathbf{y_1}, \ldots, \mathbf{y_m}]$, hence we identify $Q(\Gamma(V))$ with $Q(K[\mathbf{x}][\mathbf{y_1}, \ldots, \mathbf{y_m}]/I(V))$.

The coordinate ring is an integral domain an $K(V)$ is a field. These are generalizations of the affine/projective/multiprojective cases. Given a point $P = (v, [v_1], \ldots, [v_m]) \in V \subset \mathbb{A}^n \times \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$ where V is a variety and a rational function $\frac{\alpha}{\beta} \in K(V)$, where P is not a zero of β. Then for every $\lambda_1, \ldots, \lambda_m \in K \setminus 0$

$$\frac{\alpha(v, \lambda_1 v_1, \ldots, \lambda_m v_m)}{\beta(v, \lambda_1 v_1, \ldots, \lambda_m v_m)} = \frac{\alpha(v, v_1, \ldots, v_m)}{\beta(v, v_1, \ldots, v_m)}$$

A rational function on V is defined at P if it has a representation where the denominator does not vanish on P. The local ring of such functions is denoted $\mathcal{O}_P(V)$. The pole set of a rational function on V, ϕ; i.e. the set of points in V on which ϕ is not defined is denoted $\mathcal{P}(\phi)$.

Remark 5.2.30. (cf. Remark 5.1.66) Consider $V = \mathbb{A}^n \times \mathbb{P}^{n_1} \times \cdots \times \mathbb{P}^{n_m}$. Then

$$\Gamma(V)/\langle y_{n_1+1} - 1, \ldots, y_{n_m+1} - 1\rangle \to \Gamma(V_*), f + \langle y_{n_1+1} - 1, \ldots, y_{n_m+1} - 1\rangle \mapsto f_*.$$

One sees that

$$K(V) \simeq Q(\Gamma(V)/\langle y_{n_1+1} - 1, \ldots, y_{n_m+1} - 1\rangle) \simeq K(V_*)$$

defines an isomorphism. When V does not contain or is not contained in $V_{y_{n_i+1}}$ for some i, then

$$\Gamma(V) \to \Gamma(V_*), f + I(V) \mapsto f_* + I(V_*),$$

defines an isomorphism. From this we get that $x_{n+1}+I(V)=1+I(V)$, and hence that

$$K(V)=Q(\Gamma(V))\simeq Q(\Gamma(V_*))=K(V_*).$$

In any case it is clear to see that if $\alpha(P)\neq 0$ for some $P=(v,[v_1,1],\dots,[v_m,1])\in V\subset \mathbb{A}^n\times U_{n_1+1}\times\cdots\times U_{n_m+1}$, then $0\neq\alpha(P)=\alpha_*(v,v_1,\dots,v_m)$, hence $\mathcal{O}_P(V)\simeq\mathcal{O}_{(v,v_1,\dots,v_m)}(V)$.

Remark 5.2.31. The pole set of a rational function $\phi\in K(V)$ is an algebraic set. Indeed, $J_\phi:=\{g\in K[\mathbf{x}][\mathbf{y_1},\dots,\mathbf{y_m}]:(g+I(V))\phi\in\Gamma(V)\}$ is a m-homogeneous such that $\mathcal{P}(\phi)=V(J_\phi)$.

Definition 5.2.32. We define the *Segre embedding* of $\mathbb{P}^{n_1}\times\mathbb{P}^{n_2}$ to be the map

$$S_{n_1n_2}:=S:\mathbb{P}^{n_1}\times\mathbb{P}^{n_2}\to\mathbb{P}^{n_1+n_2+n_1n_2}$$
$$([v_1],[v_2])\mapsto[v_{11}v_{21},\dots,v_{11}v_{2,n_2+1},\dots,v_{1,n_1+1}v_{21},\dots,v_{1,n_1+1}v_{2,n_2+1}]$$

Remark 5.2.33. This map is well-defined. First of all $\#\{(i,j):1\le i\le n_1+1,1\le j\le n_2+1\}=(n_1+1)(n_2+1)=n_1n_2+n_1+n_2+1$ and for some pair (i,j), $v_{1i}\neq 0\neq v_{2j}$, hence the image of S is within the assigned co-domain. Secondly, if $(v_1,v_2)=(\lambda_1v_1,\lambda_2v_2)$ for some $\lambda_1,\lambda_2\in K\setminus 0$, then for each $\lambda\in K\setminus 0$,

$$(\lambda\lambda_1\lambda_2v_{1i}v_{2j})=\lambda\lambda_1\lambda_2(v_{1i}v_{2j}),$$

hence S is well-defined. Note also that if $v_{1i}v_{2i}=\lambda v_{1i}v_{2j}$ for a pair (i,j) with $v_{1i},v_{2j}\neq 0$, then $v_{1i}=\lambda v_{1i}$ and $v_{2j}=\lambda v_{2j}$ which implies the map is injective.

Lemma 5.2.34. *Let $n_1,n_2\ge 1$, and $W=V(f_1,\dots,f_m)\subset\mathbb{P}^{n_1+n_2+n_1n_2}$ be a variety. Define*

$$g_i:=f_i(x_{11}x_{21},\dots,x_{11}x_{2,n_2+1},\dots,x_{1,n_1+1}x_{21},\dots,x_{1,n_1+1}x_{2,n_2+1}),$$

which are polynomials in $K[\mathbf{x_1},\mathbf{x_2}]$. Then

$$S^{-1}(W)=V(g_1,g_2,\dots,g_m).$$

Proof. Indeed, substituting v into g_i is the same as substituting $S(v)$ into f_i, hence

$$v\in S^{-1}(W)\iff g_i(v)=f_i(S(v))=0\iff v\in V(g_1,\dots,g_m).$$

□

Lemma 5.2.35. *Set $V:=V(\{y_{ij}y_{kl}-y_{il}y_{kj}:i,k\in\{1,\dots,n_1+1\},j,l\in\{1,\dots,n_2+1\}\})$. Then $S(\mathbb{P}^{n_1}\times\mathbb{P}^{n_2})=V$, $S(U_i\times U_j)=V\cap U_{ij}$ and V is a variety.*

Proof. Let $S(v) \in S(\mathbb{P}^{n_1} \times \mathbb{P}^{n_2})$. Then

$$\mathrm{ev}_{S(v)}(y_{ij}y_{kl} - y_{il}y_{kj}) = v_{1i}v_{2j}v_{1k}v_{2l} - v_{1i}v_{2l}v_{1k}v_{2j} = 0.$$

Conversely, if $v \in V$, then $v_{ij}v_{kl} = v_{il}v_{kj}$. For some pair k,l, $v_{kl} \neq 0$. Then for any $(i,j) \neq (k,l)$,

$$v_{ij} = (v_{il}v_{kj})/v_{kl}$$

. Define $w_{1i} = v_{il}/v_{kl}$ and $w_{2j} = v_{kj}/v_{kl}$. Define $w_1 = (w_{11},\dots,w_{1,n_1+1})$ and $w_2 = (w_{21},\dots,w_{2,n_2+1})$. Consider $\overline{S} : \mathbb{A}^{n_1} \setminus 0 \times \mathbb{A}^{n_2} \setminus 0 \to \mathbb{A}^{n_1+n_2+n_1n_2} \setminus 0, (u_1,u_2) \mapsto (u_{1i}u_{2j})$. Then

$$\overline{S}(w_1,w_2)_{ij} = w_{1i}w_{2j} = v_{il}v_{kj}/v_{kl} = v_{ij} \Rightarrow S([w_1],[w_2]) = [\overline{S}(w_1,w_2)] = [v].$$

In particular if $([v_1],[v_2]) \in U_i \times U_j$, then

$$\overline{S}(v_1,v_2)_{ij} = v_{1i}v_{2j} \neq 0 \Rightarrow S([v_1],[v_2]) = [\overline{S}(v_1,v_2)] \in U_{ij}.$$

Suppose $V = V_1 \cup V_2$ is a decomposition of V. Then

$$S^{-1}(V_1) \cup S^{-1}(V_2) = S^{-1}(V_1 \cup V_2) = S^{-1}(V) = S^{-1}(S(\mathbb{P}^{n_1} \times \mathbb{P}^{n_2})) = \mathbb{P}^{n_1} \times \mathbb{P}^{n_2}$$

is a decomposition of $\mathbb{P}^{n_1} \times \mathbb{P}^{n_2}$, meaning $S^{-1}(V_i) = \mathbb{P}^{n_1} \times \mathbb{P}^{n_2}$ for some i, hence $V_i = V$ for some i. □

5.3 Projective Plane Curves

5.3.1 Definitions and Basic Results

Definition 5.3.1. We say that two non-constant forms $F,G \in K[x,y,z]$ are equivalent if there is a $\lambda \in K \setminus 0$ such that $G = \lambda F$. In this case we write $F \sim G$. An element of $\{F \in V_K(d,3) : d \geq 1\}/\sim$ is called a *projective plane curve*. The *degree* of a curve is just the degree of a representative of curve.

Remark 5.3.2. Every non-trivial algebraic set is characterized by such curves. Consider the factors of F, $F_1,\dots,F_m$, Then $I(V(F)) = F_1 \cdots F_m$, we thus (as in the affine case) lose information about multiplicities. We call these factors *components*. For an irreducible f, $V(f)$ is irreducible. There is a one-to-one correspondence between projective varieties in $\mathbb{P}^2$ and irreducible projective plane curves f. We define $\Gamma(F) := \Gamma^h(V(F))$, $K(F) := K(V(F))$, for a point $P \in F$, $\mathcal{O}_P(F) := \mathcal{O}_P(V(F)$, when F is irreducible. If $P = [v,w,1] \in U_3$ then $V(F)$ does not contain, and is not contained in L_∞, hence $\mathcal{O}_P(F) \simeq \mathcal{O}_{(v,w)}(F_*)$, since if $\frac{\beta}{\alpha} \in K(F)$, then P is a zero of α if and only if $\alpha_*(v,w) = 0$. We thus find that the isomorphism, $\Gamma(F) \simeq \Gamma(F_*)$ extends to an isomorphism $\mathcal{O}_P(F) \simeq \mathcal{O}_{(v,w)}(F_*)$. This result is obviously also true for U_1, U_2 with $F_{*,1}, F_{*,2}$. The upshot of this is that the results of 4.4.2 carries over to projective curves. Thus upon, for an arbitrary curve F, $P \in U_i$, $i = 1,2,3$ we define $m_P(F) := m_{(v_k,v_l)}(F_{*,i})$, $k,l \neq i$ which will only depend the local ring of each component of F (with multiplicity) by Corollary 2.9.110. This definition is independent of choice of U_i. Indeed for a component C passing through $P = [v_1, v_2, v_3] \in U_i \cap U_j$. Then $L_{\infty,k} = \{[w_1, w_2, w_3] \in \mathbb{P}^3 : w_k = 0\}$ is not contained in and does not contain C. Then $\mathcal{O}_{(v_k,v_h)}(C_{*,i}) \simeq \mathcal{O}_P(C) \simeq \mathcal{O}_{(v_l,v_m)}(C_{*,j})$ for $k,h \neq i$, $l,m \neq j$. From this point we therefor let $\bullet_*$ denote dehomogenization with respect to the appropriate variable. The multiplicity is also invariant under a projective change of coordinates. Let $\varphi = (l_1, l_2, l_3)$ be a projective change of coordinates. Then

$$\mathcal{O}_{\varphi^{-1}(P)}(C(\varphi_1, \varphi_2, \varphi_3)) \simeq \mathcal{O}_P(C).$$

We say that a point on F is *simple* if $m_P(F) = 1$ and *multiple* if $m_P(F) > 1$.

Lemma 5.3.3. *Let F be a projective plane curve and $P \in F$. Then P is a multiple point of F if and only if $F_x(P) = F_y(P) = F_z(P) = 0$*

Proof. WLOG $P = [v_1, v_2, 1] \in U_3$, since multiplying with z commutes with $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$,

$$(F_x)_* = (F_*)_x \text{ and } (F_y)_* = (F_*)_y.$$

Using Euler's formula we get (in any characteristic)

$$dF = xF_x + yF_y + zF_z.$$

"$\Rightarrow$": If P is a multiple point of F. Then $F_x(P) = (F_x)_*(v_1, v_2) = (F_*)_x(v_1, v_2) = 0$ and $F_y(P) = (F_y)_*(v_1, v_2) = (F_*)_y(v_1, v_2) = 0$. Then

$$0 = dF(P) = v_1 F_x(P) + v_2 F_y(P) + F_z(P) = F_z(P).$$

(Note that morally in the above, we take evaluation in P to evaluation in an arbitrary representative of P).
"$\Leftarrow$": We get that $(F_*)_x(v_1, v_2) = F_x(P) = 0$ and $(F_*)_y(v_1, v_2) = F_y(P) = 0$. It follows that $m_P(F) = m_{(v_1,v_2)}(F_*) > 1$. □

Definition 5.3.4. Let F be a projective curve and $P_1, \dots, P_n \in \mathbb{P}^2$ be distinct points. Let L be a line that does pass through these points (cf. Example 5.1.68 2.). We define $F_* = \frac{F}{L^{\deg F}} \in K(\mathbb{P}^2)$.

Remark 5.3.5. 1. Suppose Λ is another line that does not pass through $P_1,\ldots,P_n$. Then

$$\frac{F}{\Lambda^{\deg\ F}} = \left(\frac{L}{\Lambda}\right)^d \frac{F}{L^d},$$

and since $L/\Lambda \in K(\mathbb{P}^2)^*$, F_* is unique up to multiplication by a unit.

2. After an appropriate projective change of coordinates, φ say. We may arrange that $\varphi(L) = L_\infty = z$ is line passing through no $Q_i = \varphi(P_i)$. Hence $Q_1,\ldots,Q_n \in U_3$ are distinct point that L_∞ does not pass through and $G = F \circ \varphi$ is curve of degree $\deg\ F$. Therefor, under the identification $K(\mathbb{P}^2) \simeq K(\mathbb{A}^2), \alpha/\beta \mapsto \alpha_*/\beta_*$, $G_* = G/z^{\deg\ F} \mapsto G(x,y,1)/1^{\deg\ F}$, hence in this case $\bullet_*$ is in fact just dehomogenization under the aforementioned identification.

3. If F is an irreducible curve, then F_* is irreducible by Corollary 2.9.110. Therefor if P is simple, then $\mathcal{O}_P(F)$ is a DVR. Then we have an order function $\mathbf{ord}_P^F$ on $Q(\mathcal{O}_P(F)) = K(F)$.

Definition 5.3.6. Let F be an irreducible projective curve with $P \in F$ a simple point. Consider a form $G \in K[x,y,z]$. We then define $\mathbf{ord}_P^F(G) := \mathbf{ord}_P^F(G_*)$.

Remark 5.3.7. Let H be any form of degree $\deg\ G$ such that $H(P) \neq 0$. Then since $\frac{1}{H+I(F)}$ is unit in $\mathcal{O}_P(F)$, $\mathbf{ord}_P^F(G/H) = \mathbf{ord}_P^F(G)$.

Definition 5.3.8. Let F, G be projective plane curves and $P \in \mathbb{P}^2$. We define the *intersection number of F and G at P* to be

$$I(P, F \cap G) := \dim\ \mathcal{O}_P(\mathbb{P}^2)/\langle F_*, G_* \rangle$$

Remark 5.3.9. One should have in mind that we dehomogenize with respect to a non-zero coordinate of P and that the definition is independent of which one we choose (if there is more than one). This has the properties of the affine version (cf. 4.4.3), in 3. however φ should be a projective change of coordinates and in 7. h should be a form of degree $\deg\ G - \deg\ F$ (simply to ensure that $g + hf$ is a form).

Definition 5.3.10. A projective line L is a *tangent* of a curve projective plane F if $I(P, L \cap F) > m_P(F)$. A multiple point $P \in F$ is called *ordinary* if F has $m_P(f)$ distinct tangents at P.

Remark 5.3.11. Let the Q be appropriate affine coordinates that are identified with P under some φ_i. We see that $I(P, L \cap F) > m_P(F)$ if and only if $I(Q, L_* \cap F_*) > m_Q(F_*)$ if and only if L_* is tangent to F_* at Q (cf. Corollary 4.4.57).

Proposition 5.3.12. *Let F be a curve at $P \in U_i$ and Q be the identification with P in $\mathbb{A}^2$. Then the tangents of F at P are the tangents of F_* at Q.*

Proof. Let $F_* = \prod_1^n l_i^{r_i} + \dots$, where $l_1, \dots, l_m$ are the tangents of F.

$$I(P, l_i \cap F) = I(Q, l_i \cap F_*) > m_Q(F_*) = m_P(F),$$

hence l_i is a tangent of F. Suppose conversely that L is a tangent of F. Then

$$I(Q, L_* \cap F_*) = I(P, L \cap F) > m_P(F) = m_Q(F_*).$$

Note that L_* is a 1-degree polynomial, for otherwise $I(P, L \cap F) \in \{0, \infty\}$. Therefor L_* is a tangent for F_* at Q, hence L_* is a form, meaning $L_* = L$. □

Example 5.3.13.

1. Consider the curve $F = xy^4 + yz^4 + xz^4$ over K with characteristic 0. Consider the system of polynomial equations

$$\begin{cases} F = 0 \\ F_x = y^4 + z^4 = 0 \\ F_y = 4xy^3 + z^4 = 0 \\ F_z = 4yz^3 + 4xz^3 = 0 \end{cases}$$

Suppose $z = 1$, then y needs to be a solution to $y^4 + 1$, which we denote by s. Then x has to be a common zero of $4s^3x + 1$ and $4s + 4x$, hence there is no solution in U_3. If $y = 1$, we find that $z = s$ and again x has to be a common zero of $4x + 1$ and $4s^3 + 4xs^3$, which is not possible in characteristic 0. Suppose $x = 1$. Then $4yz^3 + 4z^3 = 0$, hence $y = -1$ or $z = 0$. In the first case $z^4 - 4 = 0$, hence $z = \pm 2$, but then $y^2 + z^2 = 5 \neq 0$, hence no such solution can exist. If $z = 0$, then $4y^3 = 0$, hence $y = 0$. Hence in U_1 the only solution is $[1,0,0]$. Note that $F(1, y, z) = y^4 + z^4 + yz^4$. Then the lowest degree form of $F(1, x, y)$ at $(0,0)$ is

$$y^4 + z^4 = (y^2 + s^2z^2)(y^2 - s^2z^2) = \underbrace{(y + s^3z)}_{l_1}\underbrace{(y - s^3z)}_{l_2}\underbrace{(y + sz)}_{l_3}\underbrace{(y - sz)},$$

hence by the prior proposition these are the tangents of F at $[1,0,0]$. Consider also $F(x - s, y + s, 1) = (x - s)(y + s)^4 + y + s + x - s = (x - s)(y^4 + 4sy^3 + 6s^2y^2$.

2. Set $F = x^2y^3 + x^2z^3 + y^2z^3$. The multiple points of F are the solution to the

system of polynomial equations

$$\begin{cases} F = 0 \\ F_x = 2xy^3 + 2xz^3 = 0 \\ F_y = 3x^2y^2 + 2yz^3 = 0 \\ F_z = 3x^2z^2 + 3y^2z^2 = 0 \end{cases}$$

We first aim to find a solution in U_3. We identify that $(0,0)$ is a valid solution. Note that $x = 0$ if and only if $y = 0$. We investigate whether there can exist others. Suppose $x \neq 0$. Then $2xy^3 + 2x = 0 \iff y^3 = -1$ call any solution to this equation α. Then x must satisfy $3\alpha^2x^2 + 2\alpha = 0$, hence $x^2 = \frac{-2}{3\alpha}$. Call this value β. Note however that

$$\alpha(3\beta^2 + 3\alpha^2) = -2 - 3 = -5 \neq 0.$$

So the only solution is $[0,0,1]$. In U_2, $[0,1,0]$ is a solution. again $x = 0$ if and only if $z = 0$. Suppose $x \neq 0$. Then $2x+2xz^3 = 0$, hence we must have that $z = \alpha$. Then we must have that $3x^2\alpha^2+3\alpha^2 = 0 \iff x^2 = -1$. Call any solution of this equation i. Then $-3\alpha^2 + 2\alpha = 0 \iff \alpha(-3\alpha + 2) = 0$, hence $[0,1,0]$ is the only solution. In U_1 we get that $[1,0,0]$ is a solution with $y = 0 \iff z = 0$. We must have that $y^3 = -z^3$ and $0 = 3z^2 + 3y^2z^2 = 3z^2(y^2 + 1)$, hence $y = i$ for a solution with $y \neq 0$ to exist. But then $z^3 = -i^3$, hence $0 = -3 - 2ii^3 = -3 - 2 = -5 \neq 0$, which implies $[1,0,0]$ is the only possible solution.

Note that $F(x,y,1) = x^2y^3 + x^2 + y^2$, hence the tangents of F at $[0,0,1]$ are $(x + iy)(x - iy)$. Note secondly that $F(x,1,z) = x^2 + x^2z^3 + z^3$, hence F has x as a double tangent at $[0,1,0]$. Note lastly $F(1,y,z) = y^3 + z^3 + y^2z^3$, hence the tangents of F at $[1,0,0]$ are $x + z, 2x + (1 + i\sqrt{3})z, 2x + (1 - i\sqrt{3})z$.

3. Consider $F = y^2z - x(x-z)(x-\lambda z) = y^2z - (x^2 - xz)(x - \lambda z) = y^2z - x^3 + (z + \lambda z)x^2 - \lambda xz^2$ where $\lambda \in K$. To find multiple points of F in $\mathbb{P}^2$ we solve the system of polynomial equations

$$\begin{cases} F = 0 \\ F_x = 3x^2 - 2(1+\lambda)zx + \lambda z^2 = 0 \\ F_y = 2yz = 0 \\ F_z = y^2 + (1+\lambda)x^2 + 2\lambda xz = 0 \end{cases}$$

Note that either $y = 0$ or $z = 0$. If $z = 0$, then $3x^2 = 0$, hence $x = 0$, but then $y^2 = 0$. Assume $z = 1$, then $y = 0$ and x must be a common zero of

$3x^2 - 2(1+\lambda)x + \lambda$ and $F(x,0,1) = x(x-1)(x-\lambda)$. $[0,0,1]$ is a multiple point if and only if $\lambda = 0$, since $F_x(0,0,1) = \lambda$. If $x = 1$

$$F_x(1,0,1) = 1 - \lambda,$$

then $[1,0,1]$ is a multiple point if and only if $\lambda = 1$. Similarly, if $x = \lambda$,

$$F(\lambda,0,1) = 3\lambda^2 - 2\lambda - 2\lambda^2 + \lambda = \lambda^2 - \lambda = \lambda(\lambda - 1),$$

hence $[\lambda,0,1]$ is a multiple point if and only $\lambda = 0$ or $\lambda = 1$. In conclusion, when $\lambda = 0$, $[0,0,1]$ is a multiple point, in which case $F(x,y,1) = y^2 + x^2 - x^3$, hence the tangents at this point are $x + iy$ and $x - iy$. When $\lambda = 1$, $[1,0,1]$ is a multiple point and $F(x+1,y,1) = y^2 - x^2(x+1) = y^2 + x^2 - x^3$, hence the tangent at this point are $x + iy$ and $x - iy$ as well. For every other value of λ, F is non-singular.

4. Consider $F = x + y + z$. Then $F_x = F_y = F_z = 1$, hence F has no multiple points

5. Consider $F = x^n + y^n + z^n$, $n > 1$, Then $F_x = nx^{n-1}$, $F_y = ny^{n-1}$ and $F_z = nz^{n-1}$, we see that the only solution to $F_x = F_y = F_z = 0$ is $x = y = z = 0$, hence F has no multiple points.

Example 5.3.14. 1. Let $F = y^2 z - x(x-2z)(x+z)$ and $G = y^2 + x^2 - 2xz$. Note that

$$\begin{cases} F_* = y^2 - x^3 + x^2 + 2x \\ G_* = y^2 + x^2 - 2x \end{cases}$$

We check whether there are common zeros of F and G in U_3. (a,b) is a common zero of $F(x,y,1)$ and $G(x,y,1)$ if and only if a is a zero of $x(x-2)(x+1) - x(x-2) = x(x-2)(x+1-1) = x^2(x-2)$, hence $a = 0$ or $a = 2$ and $b = 0$. The common zeros of F and G in U_3 therefor are $P_1 = [0,0,1]$ and $P_2 = [2,0,1]$.

$$\begin{aligned} I(P_1, F \cap G) &= I((0,0), y^2 - x^3 + x^2 + 2x \cap y^2 + x^2 - 2x) = I((0,0), y^2 - x^3 + x^2 + 2x \cap x^3 - 4x) \\ &= I((0,0), y^2 - x^3 + x^2 + 2x, x) + I((0,0), y^2 - x^3 + x^2 + 2x \cap x^3 - 4) \\ &= I((0,0), y^2, x) = 2. \end{aligned}$$

Secondly, note that

$$\begin{cases} F_*(x+2,y) = y^2 - x(x+2)(x+3) \\ G_*(x+2,y) = y^2 + x(x+2) \end{cases}$$

$$\begin{aligned} I(P_2, F\cap G) &= I((0,2), F_*\cap G_*) = I((0,0), F_*(x,y+2,1)\cap G(x,y+2,1)) \\ &= I((0,0), y^2 - x(x+2)(x+3)\cap y^2 + x(x+2)) = I((0,0), y^2\cap y^2+x^2+2x) \\ &= 2I((0,0), y\cap y^2+x^2+2x) = 2. \end{aligned}$$

Note that $F(x,y,0) = x^3$ and $G(x,y,0) = y^2 + x^2$. It thus follows that F and G have no other common zeros.

Definition 5.3.15. Two curves F and G are *projectively equivalent* if there is a projective change of coordinates φ such that $\varphi^{-1}(F) = G$

Remark 5.3.16. Any statement that only depends on local rings of curves will thus be true for any curve in an equivalence class under this equivalence relation. This is useful since we may prove a general statement on projective curves by reducing it to a statement to nicer equivalent curves

Lemma 5.3.17. *Let P be a simple point of a projective curve F. Then the tangent of F at P is given by*

$$F_x(P)x + F_y(P)y + F_z(P)z$$

Proof. WLOG $P = [0,0,1] \in U_3$. Then then tangent of F at P is equal to the tangent of F_* at $(0,0)$ which is given by $L = (F_x)_*(0,0)x + (F_y)_*(0,0)y$. Set $d := \deg F$. Using Euler's formula

$$0 = dF(P) = F_z(P),$$

hence we find that

$$L = F_x(P)x + F_y(P)y + F_z(P)z.$$

$\square$

Lemma 5.3.18. *Let P be a point on a projective curve F. Then*

$$m_P(F_x) \geq m_P(F) - 1.$$

Proof. WLOG $P = [0,0,1]$. Set $m := m_P(F) = m_0(F_*)$. We consider $F_* = F(x,y,1) = \sum_m^d f_i$. Then

$$(F_*)_x = \sum_m^d (f_i)_x.$$

If $(f_i)_x = 0$ for every i, then trivially $\infty = m_P(F_x) = m_0((F_*)_x) > m_0(F_*) = m_P(F)$. So let $n \geq m$ be the smallest index such that $(f_n)_x \neq 0$. Then $(f_n)_x$ is a form degree $n-1$, hence

$$m_P(F_x) = m_0((F_*)_x) = n - 1 \geq m_P(F) - 1$$

$\square$

Proposition 5.3.19. *Any two curves F and G with no common components, have only finitely many points of intersection*

Proof. If F and G have no common zeroes, then $\gcd(F_{*,i}, G_{*,i}) = 1$, $i = 1,2,3$ by Corollary 2.9.110. Hence

$$\#V(F,G) \le \sum_1^3 \#(V(F,G) \cap U_i) < \infty.$$

It can also be seen from the fact that

$$\sum_P I(P, F \cap G) \le \sum_1^3 \sum_P I(P, F_{*,i} \cap G_{*_i}) < \infty,$$

by Corollary 4.4.56. ☐

Proposition 5.3.20. *Let F be an irreducible projective curve. F has only finitely many multiple points.*

Proof. By proposition 4.4.22, F has only finitely many multiple points in each U_i (note that this proof works in any characteristic).
(**Note to self:** In Fulton this is marked as an obligatory exercise while the proposition is not. The (I think) intended proof uses the fact that WLOG $F_x \neq 0$ (since F is non-constant), implying $(F_x)_* \neq 0$ hence $\deg\,(F_x)_* < \deg\, F_*$, implying $F_* \nmid (F_x)_*$, hence $\gcd(F_*, (F_x)_*) = 1$, hence $F(P) = F_x(P) = F_y(P) = F_z(P) = 0$ for only finitely many P in each U_i. This ONLY works in characteristic 0. To prove the characteristic $p > 0$ case, we would apply Lemma 4.4.21 to F in each U_i, before using the $\operatorname{char} K = 0$-argument. This is just to say, I haven't strayed from the books approach in going to the simple two line proof.) ☐

Example 5.3.21.

1. Let C be an irreducible *conic*; i.e. a projective curve of degree 2. In this example we aim to show that up to projective equivalence $C = yz - x^2$. Suppose $P = [0,0,1]$ is a simple point of C with y as tangent line of C at P. Write $C = ax^2 + by^2 + cz^2 + dxy + exz + fyz$. Then $C_* = ax^2 + by^2 + c + dxy + ex + fy$. Then $c = e = 0$ and $f \neq 0$, hence $C = ax^2 + by^2 + dxy + fyz$. If $a = 0$, C is reducible. If $a \neq 0$, it follows from Eisenstein's criterion using y that $C \in K[z,y][x]$ is irreducible. WLOG $C = fyz - (ax^2 + by^2 + dxy)$ with $a, f \neq 0$. Set φ equal to the projective change of coordinates induced by $\operatorname{diag}(a^{-1}, f^{-1}, 1)$. Then $C^\varphi = yz - x^2 - \alpha y^2 - \beta xy$. Consider then the projective change of coordinates ϕ induced by
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \beta & \alpha & 1 \end{pmatrix}$$

Then

$$\begin{aligned}(C^{\varphi})^{\phi} &= y(\beta x + \alpha y + z) - x^2 - \alpha y^2 - \beta xy \\ &= \beta xy + \alpha y^2 + yz - x^2 - \alpha y^2 - \beta xy = yz - x^2.\end{aligned}$$

Given an arbitrary irreducible conic we apply a projective change of coordinates such that it has P as a simple point and y as a tangent line at P to see that it will be projectively equivalent to $F := yz - x^2$. Note that

$$\begin{cases} F_x = 2x \\ F_y = z \\ F_z = y \end{cases}$$

hence any irreducible conic will be non-singular when $\operatorname{char} K \neq 2$ and the only common zero of the partial derivatives when $\operatorname{char} K = 2$, is $[1,0,0]$ which is not a zero of F.

2. Consider an irreducible cubic C with a cusp $P := [0,0,1]$ and tangent y at P. We aim to show that up to projective equivalence an irreducible cubic with a cusp is $y^2z - x^3$. Write

$$\begin{aligned}C = a_{300}x^3 + a_{030}y^3 + a_{003}z^3 + a_{210}x^2y + a_{201}x^2z \\ + a_{120}xy^2 + a_{021}y^2z + a_{102}xz^2 + a_{012}yz^2 + a_{111}xyz\end{aligned}$$

Then

$$\begin{aligned}C_* = a_{300}x^3 + a_{030}y^3 + a_{003} + a_{210}x^2y + a_{201}x^2 \\ + a_{120}xy^2 + a_{021}y^2 + a_{102}x + a_{012}y + a_{111}xy.\end{aligned}$$

Then since P is a double point and y is the only tangent of C, we necessarily have that $a_{003} = a_{102} = a_{012} = a_{111} = a_{201} = 0$ and $a_{021} \neq 0$ hence

$$C = ay^2z - bx^3 - cy^3 - dx^2y - exy^2,$$

where $a \neq 0$. If $b = 0$, then $y \mid C$, hence $b \neq 0$. Take φ to be the projective change of coordinates induced by

$$\begin{pmatrix} b^{-1} & 0 & 0 \\ 0 & a^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then $C^{\varphi} = y^2z - x^3 - \alpha y^3 - \beta x^2 y - \gamma x y^2$ Let ϕ be the projective change of coordinates induced by

$$\begin{pmatrix} 1 & -\beta/3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

We thus get that

$$\begin{aligned}
(C^{\varphi})^{\phi} &= y^2z - (x-\beta/3y)^3 - \alpha y^3 - \beta(x-\beta/3y)^2 y - \gamma(x-\beta/3y)y^2 \\
&= y^2z - x^3 + \beta x^2 y - \beta^2/3xy^2 + \beta^3/27y^3 - \alpha y^3 - \beta x^2 y - \beta^2/9y^3 + 2\beta/3xy^2 - \gamma xy^2 + \gamma\beta/3y^3 \\
&= y^2z - x^3 - (\underbrace{\beta^2/3 - 2\beta/3 + \gamma}_{A})xy^2 - (\underbrace{-\beta^3/27 + \alpha^3 + \beta^2/9 - \gamma\beta/3}_{B})y^3
\end{aligned}$$

Let ψ be the projective change of coordinates induced by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ A & B & 1 \end{pmatrix}$$

Then

$$((C^{\varphi})^{\phi})^{\psi} = y^2(Ax+By+z) - x^3 - Axy^2 - By^3 = y^2z - x^3 =: F.$$

Then

$$\begin{cases} F_x = 2x \\ F_y = 2yz \\ F_z = y^2 \end{cases}$$

Then we see that $[0,0,1]$ is a singularity. It is clear that it's the only singularity for $\mathbf{char}\, K \neq 2$. In characteristic 2 another candidate would be $[\alpha,0,\beta]$, $(\alpha,\beta) \neq 0$, but for such a point to be a zero of F, $\beta = 0$. Therefor any irreducible cubic with a cusp has only one multiple point, namely the cusp.

3. Up to projective equivalence there is only one irreducible cubic with a node. Consider an irreducible cubic C with a node (a double point at which there are distinct tangents of the curve) at $P := [0,0,1]$ and tangents x and y Write

$$\begin{aligned}
C = {} & a_{300}x^3 + a_{030}y^3 + a_{003}z^3 + a_{210}x^2y + a_{201}x^2z \\
& + a_{120}xy^2 + a_{021}y^2z + a_{102}xz^2 + a_{012}yz^2 + a_{111}xyz
\end{aligned}$$

Then

$$\begin{aligned}
C_* = {} & a_{300}x^3 + a_{030}y^3 + a_{003} + a_{210}x^2y + a_{201}x^2 \\
& + a_{120}xy^2 + a_{021}y^2 + a_{102}x + a_{012}y + a_{111}xy.
\end{aligned}$$

Then $a_{003} = a_{201} = a_{021} = a_{102} = a_{012} = 0$ and $a_{111} \neq 1$. Hence

$$C = axyz - bx^3 - cy^3 - dx^2y - exy^2$$

with $a \neq 0$. If b or c are 0, then $x \mid C$ respectively $y \mid C$, hence $b, c \neq 0$. Changing coordinates by φ induced by $\mathbf{diag}(b^{-1}, c^{-1}, a^{-1})$, we get

$$C^{\varphi} = xyz - x^3 - y^3 - \alpha x^2 y - \beta xy^2.$$

Let ϕ be the projective change of coordinates induced by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & \beta & 1 \end{pmatrix}$$

Then we find that

$$(C^{\varphi})^{\phi} = xy(z + \alpha x + \beta y) - x^3 - y^3 - \alpha x^2 y - \beta xy^2 = xyz - x^3 - y^3 =: F$$

See that

$$\begin{cases} F_x = yz - 3x^2 \\ F_y = xz - 3y^2 \\ F_z = xy \end{cases}$$

In any characteristic it is true that $x = 0$ or $y = 0$, plugging this into F, we see that both x and y have to be 0, hence any irreducible cubic with a node has only the node as a multiple point.

Lemma 5.3.22. *Let F be a curve of degree $d > 0$ passing through $Q := [0,0,1]$. Then*

$$\sum_{P} I(P, F \cap x) = d \ \text{ or } \sum_{P} I(P, F \cap y) = 0$$

Proof. Since $F(0,0,1) = 0$, $F_* = ax^d + by^d + \ldots$ where $a \neq 0$ or $b \neq 0$, then $x \nmid F_*$ or $y \nmid F_*$. Note that F does not pass through $[\alpha, \beta, 0]$. Suppose $x \nmid F_*$. It follows from Proposition 4.4.56 that

$$\sum_{P} I(P, F \cap x) = \sum_{P \in \mathbb{A}^2} I(P, F_* \cap x) = \dim\ K[x,y]/\langle F_*, x\rangle = \dim\ K[y]/\langle F_*(0,y)\rangle = d,$$

where the last equality follows from Lemma 2.9.46. $\square$

Remark 5.3.23. Given an arbitrary curve F of degree d and line L not contained in F that intersect at a point Q, we may take a projective change of coordinates taking Q to $[0,0,1]$,L to x and F to some degree n curve G. Then

$$\sum_{P} I(P, F \cap L) = \sum_{P} I(P, G, x) = d.$$

Proposition 5.3.24. *An irreducible cubic is either non-singular or has exactly one double point.*

Proof. Let F be an irreducible curve of degree n with two multiple points $P_1, P_2 \in F$. Let L be the line through these two points. Note that since F is irreducible, $L \not\subset F$. Then by prior lemma

$$n = \sum_P I(P, F \cap L) = I(P_1, F \cap L) + I(P_2, F \cap L) + \dots \geq m_{P_1}(F) + m_{P_2}(F) \geq 4,$$

hence any irreducible curve of degree < 4 has at most one multiple point. Could it be a triple point? No, since any cubic with a triple is projectively equivalent to $x^3 + y^3 = (x+y)(x^2 - xy + y^2)$. Note that if an irreducible cubic admits a double point, the tangents at that point are either distinct or not, hence up to projective equivalence the curve is either $y^2z - x^3$ (the irreducible cubic with a cusp) or $xyz - x^3 - y^3$ (the irreducible cubic with a node).
This hints at another proof: An irreducible cubic cannot have $m_P(F) > 3$ for any point. We have also seen that it cannot have $m_P(F) = 3$ at any point. So if it has a multiple point it is either a node or a cusp in which case the curve is either projectively equivalent to $y^2z - x^3$ resp. $xyz - x^3 - y^3$, meaning such a curve has only one double point. □

Lemma 5.3.25. *Let $P, P_1, \dots, P_n \in \mathbb{P}^2$ be distinct points. Then there are infinitely many lines passing through P not passing through P_1.*

Proof. the set of lines through P is in one-to-one correspondence with points distinct from P. Indeed, set $P = [v]$ and consider $Q = [w], Q' = [w']$, $L(Q,P) = L(Q',P)$. Then for suitable $\alpha, \beta, \gamma \in K$, we have that $w = \alpha v + \beta w' = \alpha v + \gamma w$, implying $\alpha = 0$ and $w = \beta w'$ with $\beta \neq 0$. This means that $\mathbb{P}^2 \setminus \{P_1, \dots, P_n, P\}$ is in bijection with the lines passing through P and not passing through $P_1, \dots, P_n$, implying that there are infinitely many such lines. □

Remark 5.3.26. Let $P \in F$ be a simple point on a curve. Let $P_1, \dots, P_n \in \mathbb{P}^2$ be distinct points distinct from P. Then there are infinitely many lines not passing through $P_1, \dots, P_n$ that intersect F transversely at P.

Lemma 5.3.27. *Let C be an irreducible projective plane curve. Consider $P_1, \dots, P_n \in C$ all distinct and simple with integers $m_1, \dots, m_n$. Then there is a $z \in K(C)$ with $\mathrm{ord}^C_{P_i}(z) = m_i$ for each i.*

Proof. By the prior lemma and remark we can for each i pick a line L_i passing through P_i and not passing through P_j for $j \neq i$ that intersects C transversally at P_i. We may also pick a line L_0 that passes through no P_i. Pick

$$z := \prod_1^n L_i^{m_i} L_0^{-\sum_{j\neq i} m_j}.$$

Indeed notice first that by Proposition 4.4.54

$$\text{ord}_{Q_i}^{C_*}\left(\prod_1^n (L_i)_*^{m_i}\right) = I\left(P_i, C \cap \prod_1^n L_i^{m_i}\right) = m_i,$$

where Q_i is the appropriate affine coordinate corresponding to P_i. Since the element $\prod_1^n (L_i)_*^{m_i} \in \mathcal{O}_{Q_i}(C_*)$ is identified with $z \in \mathcal{O}_{P_i}(C)$, we find that $\text{ord}_{P_i}(z) = \text{ord}_{Q_i}(\prod_1^n (L_i)_*^{m_i}) = m_i$. □

Lemma 5.3.28. *Let $P \in F$ be a point on an irreducible curve. Suppose $I(P, F \cap z) = 1$ and $P \neq [1,0,0]$. Then $F_x(P) \neq 0$.*

Proof. We prove the contrapositive: If G is a curve that intersect with z at a point P and $G_x(P) = 0$, then $P = [\alpha, \beta, 0]$ and by Euler's formula

$$0 = (\deg\ G)G(P) = \beta G_y(P).$$

Then $P = [1,0,0]$ or $G_y(P) = 0$, which implies $G_z(P) = 0$, meaning $I(P, F \cap z) > 1$. □

5.3.2 Linear Systems Of Curves

The projective plane curves of $V(d,3)$, $(d > 1)$ has $\{\mathbf{x}^v : |v| = d\}$ as a basis over K. Set $N = \binom{d+2}{2} = \frac{(d+1)(d+2)}{2}$, we then get a commutative diagram

$$\begin{tikzcd}
V(d,3) \arrow[r, "\sim"] \arrow[d, two heads] & \mathbb{A}^N \arrow[d, two heads] \\
V(d,3) \setminus 0 \arrow[r, "\sim"] \arrow[d, two heads] & \mathbb{A}^N \setminus 0 \arrow[d, two heads] \\
(V(d,3) \setminus 0)/\sim \arrow[r, "\sim"] & \mathbb{P}^{N-1}
\end{tikzcd}$$

which is a natural motivation to study the behavior of curves of fixed degree via the identification of these with points in $\mathbb{P}^{N-1} = \mathbb{P}^{\frac{d(d+3)}{2}}$. We denote this space by $\mathcal{L}_d$.

Definition 5.3.29. Let d be some fixed positive integer. Let S be a subset of $\mathcal{L}_d$. We then have a natural identification of S with a subset of $\mathbb{P}^{\frac{d(d+3)}{2}}$. If $S \subset \mathbb{P}^{\frac{d(d+3)}{2}}$ is a linear subvariety, we say that it is *linear system of plane curves.*

Remark 5.3.30. Throughout this subsection we fix a positive integer d; and make the choice to denote $(d+1)(d+2)/2$ by N and the standard basis elements of $V(d,3)$ be $M_1,\ldots,M_d$.

Proposition 5.3.31. *Let d be some fixed positive integer and Fix a point $P \in \mathbb{P}^2$. The set S of curves passing through P is a linear system of plane curves (in particular a hyperplane).*

Proof. Let $F = \sum_1^N a_i M_i$, $[a_1,\ldots,a_N] \in \mathbb{P}^{N-1}$ be degree d curve. Then $F(P) = 0$ if and only if $\sum_1^d a_i M_i(P) = 0$. There is at least one monomial that does not vanish on P. This means that $\sum_1^N M_i(P)y_i \in K[y_1,\ldots,y_N] \setminus 0$ and that $S = V(\sum_1^N M_i(P)y_i)$, meaning S is a hyperplane. □

Corollary 5.3.32. *Let $P_1,\ldots,P_n \in \mathbb{P}^2$ be an arbitrary set of points. If $n \leq N-1$ The set of curves passing through these points is a linear system of plane curve.*

Corollary 5.3.33. *Let $P_1,\ldots,P_{N-1} \in \mathbb{P}^2$ be distinct points. There is exactly one curve of degree d passing through each of these points.*

Proof. This follows from Lemma 5.1.54. □

Lemma 5.3.34. *Let $\varphi =$ be a projective change of coordinates on $\mathbb{P}^2$. Then the map $\mathcal{L}_d \to \mathcal{L}_d, F \mapsto F^\varphi$ induces a projective change of coordinates on $\mathbb{P}^{N-1}$.*

Proof. We define $\phi : \mathbb{P}^{N-1} \simeq \mathcal{L}_d \xrightarrow{\varphi} \mathcal{L}_d \simeq \mathbb{P}^{N-1}$. Let $F\mathcal{L}_d$. Note that since φ is invertible F^φ is a degree d curve, since otherwise $F = (F^\varphi)^{\varphi^{-1}} = 0$, hence $\phi([v_1,\ldots,v_N])$ is an element of $\mathbb{P}^{N-1}$. For each M_i we get that

$$M_i \circ \varphi = \sum_1^d b_{ij} M_j,$$

hence

$$F^\varphi = \sum_1^d a_i M_i \circ \varphi = \sum_1^d a_i \sum_1^d b_{ij} M_j = \sum_{j=1}^d \left[\sum_{i=1}^d b_{ij} a_i \right] M_j.$$

Therefor

$$\phi([v_1,\ldots,v_N]) = [\sum_1^d b_{i1} v_i, \ldots, \sum_1^d b_{iN} v_i]$$

hence ϕ is induced by the linear transformation on $\mathbb{A}^N$, $(b_{ij})^T$. This is invertible since it has mutual inverse induced by $F \mapsto F^{\varphi^{-1}}$, and is thus a projective change of coordinates. □

Definition 5.3.35. Let $P_1,\ldots,P_n \in \mathbb{P}^2$ be distinct points and $1 \le r_i \le d+1$ for each i. We define

$$V(d;r_iP_1,\ldots,r_nP_n) := \left\{F \in \mathcal{L}_d : m_{P_i}(F) \ge r_i, 1 \le i \le n\right\}$$

Lemma 5.3.36. *Let $P \in \mathbb{P}^2$ and $1 \le r \le d+1$. $V(d;rP) \subset \mathbb{P}^{N-1}$ is a linear subvariety of dimension*

$$N-1-\frac{r(r+1)}{2}.$$

In other words **codim** $V(d;rP) = \frac{r(r+1)}{2}$

Proof. By the prior lemma we may assume that $P=[0,0,1]$. Note that $F \in \mathcal{L}_d$ has multiplicity greater than r if and only if the coefficient of F at $x^iy^jz^k$ is 0 for every (i,j) with $i+j<r$, which is readily seen by writing $F = \sum_0^d F_i z^{d-i}$ for forms F_i. There are $D := \#\{(i,j) : i+j<r\} = \binom{r-1+2}{r} = \binom{r+1}{r} = \frac{r(r+1)}{2}$ It follows that $V(d;rP)$ can be identified with

$$\left\{[v_1,\ldots,v_{N-D},0\ldots,0] \in \mathbb{P}^{N-1}\right\} = V(y_{N-D+1},\ldots,y_N)$$

It thus follows that **dim** $V(d;rP) = N-1-D$. □

Lemma 5.3.37. *Let $P_1,\ldots,P_n \in \mathbb{P}^2$ be distinct points and let $r_1,\ldots,r_n$ be positive integers $\le d+1$. Then $V(d;r_1P_1,\ldots,r_nP_n)$ is linear subvariety of dimension*

$$\ge N-1-\sum_1^n \frac{r_i(r_i+1)}{2}.$$

Proof. Note that

$$V(d;r_1P_1,\ldots,r_nP_n) = \bigcap_1^n V(d;r_iP_i),$$

hence by Lemma 2.7.8 and the prior lemma, it follows that

$$\begin{aligned}\textbf{dim } V(d;r_1P_1,\ldots,r_nP_n) &= N-1-\textbf{codim } V(d;r_1P_1,\ldots,r_nP_n) \ge N-1-\sum_1^n \textbf{codim } V(d;r_iP_i)\\ &= N-1-\sum_1^n \frac{r_i(r_i+1)}{2}\end{aligned}$$

□

Theorem 5.3.38. *Let $P_1,\ldots,P_n \in \mathbb{P}^2$ be distinct points and let $r_1,\ldots,r_n$ be positive integers. Suppose $d \ge \left[\sum_1^n r_i\right]-1$. Then*

$$\textbf{dim } V(d;r_1P_1,\ldots,r_nP_n) = N-1-\sum_1^n r_i.$$

Proof. We prove statement by induction in value $m := \left[\sum_1^n r_i\right] - 1$. In the case $m = 0$ we get that $r_i = 1$ for some i and $r_j = 0$ for $j \neq i$. Then

$$\begin{aligned} V(d;r_1P_1,\ldots,r_nP_n) = V(d;r_iP_i) \Rightarrow \dim\ V(d;r_1P_1,\ldots,r_nP_n) &= N-1-\frac{r_i(r_i+1)}{2} \\ &= N-q-\sum_1^n \frac{r_j(r_j+1)}{2}. \end{aligned}$$

So assume $m > 1$, $d > 1$.

Case 1: Suppose first that $r_i = 1$ for each i. Then $V(d;P_1,\ldots,P_n)$ is the intersection of n hyperplanes. By induction $\dim\ V(d;P_1,\ldots,P_{n-1}) = N-n$. Choose lines through L_i through P_i not passing through P_j for $j \neq i$ for $i \in \{1,\ldots,n-1\}$ and L_0 a line not passing through $P_1,\ldots,P_{n-1}$. Then $F = L_0^{d-n+1}\prod_1^{n-1} L_i \in V(d;P_1,\ldots,P_{n-1})$ but not in $V(d;P_1,\ldots,P_n)$, hence the hyperplanes generating $V(d;P_1,\ldots,P_n)$ are linearly independent (since $V(d;P_1,\ldots,P_n) \subsetneq V(d;P_1,\ldots,P_{n-1})$, meaning $\dim\ V(d;P_1,\ldots,P_n) = N-n-1 = N-1-\sum_1^n \frac{r_i(r_i+1)}{2}$.

Case 2: Suppose some $r_i > 1$; WLOG $r := r_1 > 1$ and $P := P_1 = [0,0,1]$. Set

$$V_0 := V(d;(r-1)P, r_2P_2,\ldots,r_nP_n),$$

and define

$$V_i := \left\{F \in V_0 : F_* = \sum_{j=i}^{r-1} a_j x^j y^{r-1-j} + \text{ higher order terms}\right\}, \quad (i = 1,\ldots,r).$$

Note that $V_r = V(d;rP,\ldots,r_nP_n)$ and in particular that

$$V_0 \supset V_1 \supset \cdots \supset V_{r-1} \supset V_r = V(d;rP,\ldots,r_nP_n).$$

If we can prove that non of these inclusions hold with equality, then by the induction hypothesis

$$N-1-\frac{(r_1-1)(r_1)}{2} - \sum_2^n \frac{r_i(r_i+1)}{2} = \dim\ V_0 > \dim\ V_1 > \ldots > \dim\ V_{r-1} > \dim V_r,$$

hence

$$\begin{aligned} \dim\ V \leq N-1-\sum_2^n \frac{r_i(r_i+1)}{2} - \frac{(r_1-1)r_1}{2} + r_1 &= N-1-\sum_2^n \frac{r_i(r_i+1)}{2} - \frac{(r_1-1)r_1+2r_1}{2} \\ = N-1-\sum_1^n \frac{r_i(r_i+1)}{2} \leq \dim\ V, \end{aligned}$$

where the last bound is due to Lemma 5.3.37. We would then conclude that $\dim\ V = N-1-\sum_1^n \frac{r_i(r_i+1)}{2}$.

We proceed to prove the sufficient claim: Set

$$W_0 := V(d-1;(r-2)P, r_2P_2,\ldots,r_nP_n)$$

and

$$W_i := \left\{ F \in W_0 : F_* = \sum_{j=i}^{r-2} a_j x^j y^{r-2-j} + \text{ higher order terms} \right\} \quad (i = 1, \ldots, r-1),$$

by induction

$$W_0 \supsetneq W_1 \supsetneq \cdots \supsetneq W_{r-2} \supsetneq W_{r-1} = V(d-1; (r-1)P, r_2P_2, \ldots, r_nP_n).$$

Fix an $i \in \{0, \ldots, r-2\}$ and pick an $F \in W_i \setminus W_{i+1}$. Then $F_* = \sum_{j=i}^{r-2} a_j x^j y^{r-2-j} + \ldots$ with $a_i \neq 0$, hence

$$(yF)_* = yF_* = \sum_{j=i}^{r-2} a_j x^j y^{r-1-j} + \ldots,$$

implying $yF \in V_i \setminus V_{i+1}$. Pick $F \in W_{r-2} \setminus W_{r-1}$. Then

$$F_* = a_{r-2} x^{r-2} + \cdots,$$

where $a_{r-2} \neq 0$. Then

$$(xF)_* = a_{r-2} x^{r-1} + \cdots,$$

hence $xF \in V_{r-1} \setminus V_r$ establishing the desired claim. $\square$

Example 5.3.39. 1. Let $P_1, \ldots, P_4 \in \mathbb{P}^2$ be distinct and set $V := V(2; P_1, \ldots, P_4)$. Suppose all 4 points lie on a line. WLOG $P_1 = [1,0,0]$,$P_2 = [0,1,0]$,$P_3 = [\alpha, \beta, 0]$ and $P_4 = [\gamma, \beta, 0]$ with $\alpha, \beta, \gamma, \delta \neq 0$ Let

$$F = ax^2 + by^2 + cxy + dz^2 + exz + fyz \in V(2; P_1, \ldots, P_4)$$

Then we need have that

$$\begin{cases} 0 = F(P_1) = a \\ 0 = F(P_2) = b \\ 0 = F(P_3) = a\alpha^2 + b\beta^2 + c\alpha\beta \\ 0 = F(P_4) = a\gamma^2 + b\delta^2 + c\gamma\delta \end{cases}$$

Hence $a = b = c = 0$, hence $V \simeq \mathrm{Span}([e_4],[e_5],[e_6])$ hence $\dim\ V = 2$. If one of $P_1, \ldots, P_4$ does not sit on a common line with the other points, we may WLOG assume that $P_1 = [1,0,0]$, $P_2 = [0,1,0]$, $P_3 = [0,0,1]$ and $P_4 = [\alpha, \beta, 0]$ with $\alpha, \beta \neq 0$. Let $F = ax^2 + by^2 + cz^2 + dxy + exz + fyz \in V(2; P_1, \ldots, P_4)$. Then

$$\begin{cases} 0 = F(P_1) = a \\ 0 = F(P_2) = b \\ 0 = F(P_3) = c0 = F(P_4) = \alpha^2 a + \beta^2 b + \alpha\beta d \end{cases}$$

Then we see that $a = b = c = d = 0$, hence $V \simeq \mathrm{Span}([e_4],[e_5])$, hence $\dim\ V = 1$.

2. Consider the points $[1,0,0],[0,1,0],[0,0,1],[1,1,1],[1,2,3] \in \mathbb{P}^2$. Denote them P_i. A curve

$$F = ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

is in $V(2;P_1,\dots,P_5)$ if and only if

$$\begin{cases} 0 = F(P_1) = a \\ 0 = F(P_2) = b \\ 0 = F(P_3) = c \\ 0 = F(P_4) = a+b+c+d+e+f \\ 0 = F(P_5) = a+4b+9c+2d+3e+6f \end{cases}$$

hence F is in $V(2;P_1,\dots,P_5)$ if and only if (a,b,c,d,e,f) is in the null-space of

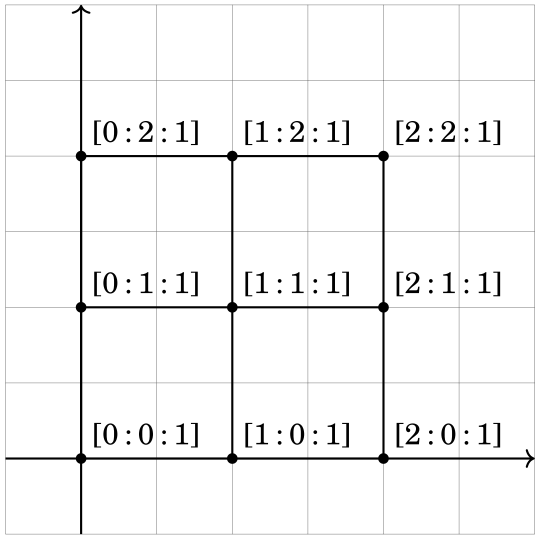
$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 4 & 9 & 2 & 3 & 6 \end{pmatrix}$$

We thus see that $a = b = c = 0$ and $d = 3f$ and $e = -4f$, hence there is only one curve passing through these five points, namely $3xy - 4xz + yz$. Note that by Eisenstein in $K[x,y][z]$ using x this is irreducible. By Example 5.3.21 1. there (up to projective equivalence) is only one such curve, which is non-singular.

3. Consider the nine points

$$[0,0,1],[0,1,1],[1,0,1],[1,1,1],[0,2,1],[2,0,1],[1,2,1],[2,1,1],[2,2,1] \in \mathbb{P}^2,$$

which we denote $P_1,\dots,P_9$. By a result seen in this subsection, we know only that $\dim V(3;P_1,\dots,P_9) \geq 0$, so it may be that there are only one. We aim to show that there are in fact infinitely many such. Below is a sketch of the positions of the points in U_3 (in $\mathbb{P}^2(\mathbb{R})$):



We see that any of the points lie on either $x, x-z$ or $x-2z$ and on either $y, y-z$ or $y-2z$. It follows that $f := x(x-z)(x-2z)$ and $g := y(y-z)(y-2z)$ are two distinct cubics passing through each of the 9 points. It follows that $L(f,g) \subset V(3;P_1,\ldots,P_9)$, hence $\#V(3;P_1,\ldots,P_9) \geq \#L(f,g) = \infty$.

5.3.3 Bézout's Theorem

Theorem 5.3.40. *(Bezout's Theorem)*
Let F and G be projective plane curves with no common components. Then

$$\sum_P I(P, F \cap G) = (\deg\ F)(\deg\ G)$$

Proof. By Proposition 5.3.19 F and G have only finitely many points of intersection. WLOG non of these lie on L_∞. Then

$$\sum_P I(P, F \cap G) = \sum_P I(P, F_* \cap G_*) = \dim\ K[x,y]/\langle F_*, G_* \rangle.$$

We define

$$\Gamma_* := K[x,y]/\langle F_*, G_* \rangle, \qquad \Gamma := K[x,y,z]/\langle F, G \rangle, \qquad R = K[x,y,z].$$

We let Γ_d denote the space of forms of degree d, and $R_d := V(d,3)$ for $d \geq 1$. Set $n := \deg\ F$ and $m := \deg\ G$. It is sufficient to prove that $\dim\ \Gamma_* = \dim\ \Gamma_d$ and that $\dim\ \Gamma_d = mn$ for some (large enough) d.
Claim 1: $\dim\ \Gamma_d = mn$ for every $d \geq m+n$. Consider the sequence

$$0 \longrightarrow R \xrightarrow{\tau} R \times R \xrightarrow{\sigma} R \xrightarrow{\pi} \Gamma \longrightarrow 0$$

where $\tau : C \mapsto (GC, -FC)$, $\sigma : (A,B) \mapsto AF + BG$ and π is the canonical surjection. It is clear that $\sigma\tau = 0 = \pi\sigma$ hence the sequence is exact. Note that τ restricts to a linear map from R_{d-m-n} to $R_{d-m} \times R_{d-n}$, that σ restricts to a linear map from $R_{d-m} \times R_{d-n}$ to R_d and that π restricts to linear map from R_d to Γ_d. We thus get by exactness (due to Lemma 2.7.4) that

$$\begin{aligned}
&\dim\ \Gamma_d = \dim\ R_d - \dim\ R_{d-m} \times R_{d-n} + \dim\ R_{d-m-n} \\
&= \frac{(d+1)(d+2) + (-d+m-1)(d-m+2) + (-d+n-1)(d-n+1) + (d-m-n+1)(d-m-n+2)}{2} \\
&= mn.
\end{aligned}$$

Let $\overline{\bullet}$ denote $\bullet + \langle F, G \rangle$.

Claim: The map $\alpha : \Gamma \to \Gamma, \overline{H} \mapsto \overline{zH}$ is injective. Suppose $zH = AF + BG$. Note that since $z \cap F = \emptyset, z \cap G = \emptyset$, one sees that $F_0 := F(x,y,0)$, $G_0 := G(x,y,0)$ are non-zero forms in $K[x,y]$ that are coprime.

Indeed, suppose $J, I \in K[x_1, \ldots, x_{n+1}]$ such that $J(x_1, \ldots, x_n, 0) \neq 0 \neq I(x_1, \ldots, x_n, 0)$ have a non-trivial common factor D, then $\emptyset \neq H \subset J(x_1, \ldots, x_n, 0), H(x_1, \ldots, x_n, 0)$, meaning $x_{n+1} \subset J(x_1, \ldots, x_{n+1}), H(x_1, \ldots, x_{n+1})$.

Set $A_0 := A(x,y,0)$ and $B_0 := B(x,y,0)$. Then

$$A_0 F_0 = -B_0 G_0.$$

Hence for some $C \in K[x,y]$, $A_0 = -G_0 C$ and $B_0 = F_0 C$. Set $A_1 := A + CG$ and $B_1 := B - CF$. Note that

$$A_1(x,y,0) = 0, \quad B_1(x,y,0) = 0,$$

hence $A_1 = zA'$ and $B_1 = zB'$ for some $A', B' \in K[x,y,z]$, hence

$$zH = AF + BG + CFG - CFG = A_1 F + B_1 F = z(A'F + B'G) \Rightarrow H = A'F + B'G \in \langle F, G \rangle$$

proving that $\mathbf{ker}\ \alpha = 0$.

Combining results: Let $d \geq m+n$ be given. Pick a basis $\{\overline{A_1}, \ldots, \overline{A_{nm}}\} \subset \Gamma_d$ for suitable $A_i \in R_d$. Consider the restriction of α to $\Gamma_d \to \Gamma_{d+1}, \overline{H} \mapsto \overline{zH}$, since $\mathbf{ker}\ \alpha|_{\Gamma_d} = 0$. It follows by rank-nullity that the restriction of α is an isomorphism. Therefor $\{\overline{zA_i} : 1 \leq i \leq mn\}$ constitutes a basis of Γ_{d+1}, and by induction $\{\overline{z^r A_i} : 1 \leq i \leq mn\}$ form a basis for Γ_{d+r} for every $r \geq 0$. We claim that setting $a_i := (A_i)_* + \langle F_*, G_* \rangle$, $\{a_1, \ldots, a_{mn}\} \subset \Gamma_*$ constitutes a basis. Let $h = H + \langle F_*, G_* \rangle \in \Gamma_*$, then $\overline{z^N H^*} \in \Gamma_{d+r}$ for some sufficiently large $N \geq 0$ and $r \geq 0$, hence

$$z^N H^* = \sum_1^{mn} \lambda z^r A_i + BF + CG$$

for some $\lambda \in K$, $B,C \in K[x,y,z]$. One then sees that

$$H = (z^N H^*)_* = \sum_1^{nm} \lambda_i (A_i)_* + B_* F_* + C_* G_* \Rightarrow h = \sum_1^{nm} \lambda_i a_i.$$

Suppose, $\sum_1^{nm} \lambda_i a_i = 0$. Then $\sum_1^{nm} \lambda (A_i)_* = BF_* + CG_*$ for some $B,C \in K[x,y]$. We then for suitably large $r,s,t \geq 0$ that

$$z^s B^* G + z^t C^* = z^r (BF_* + CG_*)^* = z^r \left(\sum_1^{mn} \lambda_i A_i \right)^* = \sum_1^{mn} \lambda_i z^{r_i} A_i \sum R_{d+r_i},$$

implying that $0 = \sum_1^{mn} \overline{\lambda_i z^{r_i} A_i} \in \sum \Gamma_{d+r_i}$. Since $\Gamma_{d+l} \cap \Gamma_{d+k} = 0$ (this is seen readily), it follows that $\{z^{r_i} A_i\}$ is a basis of $\sum \Gamma_{d+r_i}$, hence in particular $\{z^{r_i} A_i\}$ are algebraically independent, meaning $\lambda_i = 0$. It follows that $\dim\ \Gamma_* = \dim\ \Gamma_d = mn$. □

Corollary 5.3.41. *Let F and G be curves with no common components. Then*

$$\sum_P m_P(F) m_P(G) \leq (\deg\ F)(\deg\ G).$$

Proof. This follows from property 5 of intersection numbers in conjunction with Bezout's theorem. □

Corollary 5.3.42. *If F and G (still have no common components) meets in $(\deg\ F)(\deg\ G)$ distinct points, then these points are simple.*

Proof. Under this extra assumption, $m_P(F) m_P(G) \geq 1$ for each point of intersection, meaning $\sum_P m_P(F) m_P(G) \geq (\deg\ F)(\deg\ G)$, hence $\sum_P m_P(F) m_P(G) = (\deg\ F)(\deg\ G)$. Then $m_P(G) m_P(F) = 1$ at each point of intersection, for otherwise we would have a strict inequality. We thus conclude that $m_P(F) = m_P(G) = 1$ at every point of intersection. □

Corollary 5.3.43. *If F and G have exactly $(\deg\ F)(\deg\ G)$ common points, then they have no common components.*

Proof. This just follows from Bezout and the fact that $\sum_P I(P, F \cap G) < \infty$ if and only if F and G intersect properly at every point. □

Proposition 5.3.44. *Every non-singular projective curve F is irreducible.*

Proof. Suppose F is reducible with two components G,H. G and H have at least one point in common, since if they have a component in common, then they intersect at infinitely many points and if they have no components in common, then by Bezout $\sum_P I(P, G \cap H) = (\deg\ G)(\deg\ H) > 1$, meaning they have at least one point in common. Let $P \in G \cap H \subset F$. Then $m_P(F) = m_P(GH) = m_P(G) + m_P(H) \geq 2$, hence F is singular. □

Remark 5.3.45. It is not the case that the above is true in the affine case. Consider for example $f = y(y+1)$. The zeroes of this curve are $(\alpha, 0)$ and $(\alpha, -1)$ where $\alpha \in K$. Since $f(x+\alpha, y+0) = f(x,y) = y^2 + y$ and $f(x+\alpha, y-1) = y(y-1) = y^2 - y$, hence f is non-singular.

5.3.4 Bounds on the Number of Multiple Points of a Curve

Proposition 5.3.46. *Let F be an irreducible projective plane curve of degree d with $F_x \neq 0$. Then*

$$\sum_P m_P(F)(m_P(F)-1) \leq d(d-1),$$

hence F has at most $\frac{d(d-1)}{2}$ multiple points.

Proof. Since $F_x \neq 0$, we have that $\deg F_x = d-1$, hence the first bound follows from Corollary 5.3.41 and Lemma 5.3.18,

$$\sum_P m_P(F)(m_P(F)-1) \leq \sum_P m_P(F) m_P(F_x) \leq d(d-1)$$

. Note that the number of multiple points is given by $\sum_P (m_P(F)-1)$. It then follows that

$$2\sum_P (m_P(F)-1) \leq \sum_{Q \text{ multiple}} \sum_P m_Q(F)(m_P(F)-1) \leq \sum_P m_P(F)(m_P(F)-1) \leq d(d-1).$$

It thus follows that

$$\sum_P (m_P(F)-1) \leq \frac{d(d-1)}{2}.$$

Another way to derive the bound:

$$\sum_P (m_P(F)-1) \leq \sum_P \frac{m_P(F)(m_P(F)-1)}{2} \leq \frac{d(d-1)}{2}.$$

$\square$

In Example 5.3.21 we saw that the optimal bound on the number of multiple points for an irreducible conic is 0 and for an irreducible cubic 1, indicating that there is a better bound in the general case.

Theorem 5.3.47. *Let F be an irreducible curve of degree $d \geq 1$. Then*

$$\sum_P (m_P(F)-1) \leq \frac{(d-1)(d-2)}{2}.$$

Proof. Set

$$r := \frac{(d-1)(d-1+3)}{2} - \sum_P \frac{m_P(F)(m_P(F)-1)}{2} \geq \frac{d(d-1)}{2} - \sum_P \frac{m_P(F)(m_P(F)-1)}{2} \geq 0.$$

Pick $Q_1, \ldots, Q_r \in F$ simple. Let $P_1, \ldots, P_l \in F$ be the multiple points. Then

$$V := V(d-1; Q_1, \ldots, Q_r, (m_{P_1}(F)-1)P_1, \ldots, (m_{P_l}(F)-1)P_l)$$

is a linear subvariety of dimension greater then $\frac{(d-1)(d-1+3}{2} - r - \sum_P \frac{m_P(F)(m_P(F)-1)}{2} = 0$. It follows that can pick a curve $G \in V$. Note that if $P \neq Q_i$ is simple, then $m_P(G) \geq 0 = m_P(F) - 1$. so $m_P(G) \geq m_P(F) - 1$ for each $P \in F$ Since $\deg\ G = d - 1 < d = \deg\ F$ and F is irreducible, $\gcd(G, F) = 1$, implying

$$\begin{aligned} r + \sum_{P \neq Q_i} m_P(F)(m_P(F)-1) &\leq \sum_1^r m_{Q_i}(G) m_{Q_i}(F) + \sum_{P \neq Q_i} m_P(G) m_P(F) \\ &= \sum_P m_P(G) m_P(F) \leq d(d-1). \end{aligned}$$

Here we use Corollary 5.3.41 for the last upper bound. Inserting the value of r into the left-hand side, we see that

$$\begin{aligned} \frac{(d-1)(d+2)}{2} + \sum_P \frac{m_P(F)(m_P(F)-1)}{2} \leq \frac{2d(d-1)}{2} \Rightarrow \sum_P \frac{m_P(F)(m_P(F)-1)}{2} &\leq \\ \frac{(d-1)(2d-d-2)}{2} &= \frac{(d-1)(d-2)}{2}. \end{aligned}$$

We therefor immediately get that $\sum_P (m_P(F) - 1) \leq \frac{(d-1)(d-2)}{2}$ □

Remark 5.3.48. This bound is sharp in the cases $d = 1, 2, 3$.

Proposition 5.3.49. *Let F be a projective plane curve of degree $d \geq 1$ with c components, each of which is not multiple. Then*

$$\sum_P \frac{m_P(F)(m_P(F)-1)}{2} \leq \frac{(d-1)(d-2)}{2} + c - 1 \leq \frac{d(d-1)}{2}.$$

Proof. write $F = F_1 F_2$ where F_1 is a component and F_2 is a product of $c-1$ remaining simple components with $c > 1$. Set $d_1 := \deg\ F_1$ and $d_2 := \deg\ F_2$. By induction and Bezout, it follows that

$$\begin{aligned}
\sum_P \frac{m_P(F)(m_P(F)-1)}{2} &= \sum_P \frac{m_P(F_1)(m_P(F_1)+m_P(F_2)-1)}{2} + \sum_P \frac{m_P(F_2)(m_P(F_1)+m_P(F_2)-1)}{2} \\
&= \sum_1^2 \sum_P \frac{m_P(F_i)(m_P(F_i)-1)}{2} + \sum_P m_P(F_1)m_P(F_2) \\
&\le \frac{(d_1-1)(d_1-2)+(d_2-1)(d_2-2)+2d_1d_2}{2} + c-1-1 \\
&= \frac{d_1^2+2-d_1-2d_1+d_2^2+2-d_2-2d_2+2d_1d_2}{2} + c-1-1 \\
&= \frac{(d_1+d_2)^2-(d_1+d_2)-2(d_1+d_2)+2}{2} + c-1 \\
&= \frac{(d-1)(d-2)}{2} + c-1.
\end{aligned}$$

Note that in the last step we use $d_1+d_2=d$. □

Proposition 5.3.50. *Assume* char $K=0$. *Let* F *be an irreducible curve of degree* $d \ge 1$. *Let* $P \in \mathbb{P}^2$ *and set* $r = m_P(F)$. *For all but finitely many lines* L *through* P, L *intersects* F *in* $d-r$ *distinct points.*

Proof. WLOG $P=[0,0,1]$. The lines through P are

$$L_\lambda := V(x-\lambda y) = \{[\lambda,1,t] : t \in K\} \cup \{P\} \quad (\lambda \in K)$$

(cf. Example 5.1.68 1.) together with $L=V(y)$. It is therefor sufficient to prove the statement for the L_λ. Write $F=\sum_r^d H_i z^{d-i}$ with $H_i \in K[x,y]$ a form degree i and $H_r \neq 0$. Consider for each $\lambda \in K$ the polynomial

$$G_\lambda := F(\lambda,1,T) \in K[T]$$

whose roots are in one-to-one correspondence with $L_\lambda \cap F$. It is therefor sufficient to prove that G_λ has $n-r$ distinct roots for all but finitely many $\lambda \in K$. Suppose $\lambda \in K$ is given such that $H_r(\lambda,1) \neq 0$ (making G_λ a $d-r$-degree polynomial) and $F \cap F_z \cap L_\lambda = \{P\}$, or equivalently that the common roots of G_λ and $(G_\lambda)_T = F_z(\lambda,1,T)$ is the empty set. Then by the contrapositive of result the $d-r$ roots of G_λ are all distinct. Clearly $H_r(\lambda,1) \neq 0$ for all but finitely many λ. Since $H_i(0,0)=0$ for all $i \ge 1$, it follows that $P \in F_z$. Note that since F is irreducible so is $F(x,1,z)$ by Corollary 2.9.110, hence $F(x,1,z)$ and $F_z(x,1,z)$ are co-prime. It then follows that $F(x,1,z) \cap F_z(x,1,z)$ is finite. In particular there are only finitely $\lambda \in K$ such that $\emptyset \neq F(\lambda,1,T) \cap F_z(\lambda,1,T) = G_\lambda \cap (G_\lambda)_T$. It follows that G_λ has $d-r$ distinct roots for all but finitely many λ. □

Proposition 5.3.51. *The above proposition extends to curves F with no multiple components.*

Proof. Write $F = \prod_1^n F_i$. Set $r_i := m_P(F_i)$ and $d_i := \deg\ F_i$. There are infinitely many lines passing through P that do not pass through $\bigcap_1^n F_i$. Denote this set $\mathcal{L}$. Then for each i, all but finitely many lines in $\mathcal{L}$, intersect F_i in $d_i - r_i$ points. Then L intersect each F_i in $d_i - r_i$ points for all but finitely many i. Let such an L be given. Then

$$L \cap F = L \cap \bigcup_1^n F_i = \bigsqcup_1^n L \cap F_i,$$

hence

$$\#(L \cap F) = \#\left(\bigsqcup_1^n L \cap F_i\right) = \sum_1^n d_i - r_i = d - \sum_1^n m_P(F_i) = d - m_P(F) = d - r.$$

□

Example 5.3.52. Suppose $\mathbf{char}\ K = p > 0$. Set $F := x^{p+1} - y^p z$ and $P := [0,1,0]$. $L_\infty \cap F = \{P\}$. Let $\lambda \in K$. Then $L_\lambda \cap F = \{t : \lambda^{p+1} - t^p = 0\} \cup \{P\} = \left\{\left[\lambda, 1, \lambda^{\frac{1}{p}}\lambda\right]\right\} \cup \{P\}$ More explicit description?. Note that

$$\begin{cases} F_x = x^p \\ F_y = 0 \\ F_z = y^p \end{cases}$$

5.3.5 Max Noether's Fundamental Theorem

Definition 5.3.53. A *zero-cycle* on $\mathbb{P}^2$ is an element of the free abelian group generated by $\mathbb{P}^2$, i.e. $\mathbb{Z}[\mathbb{P}^2]$.

Definition 5.3.54. We define the *degree* of a zero-cycle $s = \sum_P n_P P \in \mathbb{Z}[\mathbb{P}^2]$ is the quantity

$$\deg\ s := \sum_P n_P.$$

For a $t = \sum_P m_P \in \mathbb{Z}[\mathbb{P}^2]$, we write $s \geq t$ if $n_P \geq m_P$ for each $P \in \mathbb{P}^2$.

Definition 5.3.55. Let F and G be curves with no common components. We define the *intersection cycle of F and G* to be

$$F \bullet G := \sum_P I(P, F \cap G) P \in \mathbb{Z}[\mathbb{P}^2]$$

Remark 5.3.56. By Bezout $\deg\ F \bullet G = (\deg\ F)(\deg\ G)$

The following properties of intersection cycles are trivial consequences of properties of intersection numbers:

Lemma 5.3.57. *Let F, G, H be curves and A a form of degree* $\deg\ G - \deg\ F$. *Then*

1. $F \bullet G = G \bullet F$.

2. $F \bullet GH = F \bullet G + F \bullet H$.

3. $F \bullet G + AF = F \bullet G$.

Definition 5.3.58. Consider projective plane curves F, G, H where F and G have not common components. Let $P \in \mathbb{P}^2$. We say that *Noether's condition (with respect to F, G and H) is satisfied at P* if $H_* \in \langle F_*, G_* \rangle \subset \mathcal{O}_P(\mathbb{P}^2)$.

Theorem 5.3.59. *(Max Noether's Fundamental Theorem/MNFT)*
Let F, G, H be projective plane curves where F and G have no common components. Then there are curves A and B with $\deg\ A = \deg\ H - \deg\ F$ *and* $\deg\ B = \deg\ H - \deg\ G$ *such that*

$$H = AF + BG$$

if and only if Noether's conditions are satisfied at every $P \in F \cap G$.

Proof. "$\Rightarrow$": Is trivial
"$\Leftarrow$": WLOG $F \cap G \cap z = \emptyset$. Since

$$K[x,y]/(K[x,y]F_* + K[x,y]G_*) \simeq \prod_{P=[v,1]\in F\cap G} \mathcal{O}_v/(\mathcal{O}F_* + \mathcal{O}_v G_*)$$

and $H_* \in \mathcal{O}F_* + \mathcal{O}_v G_*$ for each $P = [v,1] \in F \cap G$, we get that $H_* \in K[x,y]F_* + K[x,y]G_*$, hence for suitable $a, b \in K[x,y]$, $H_* = aF_* + bG_*$. For a suitably large $N \geq 0$,

$$z^r H = z^N (H_*)^* = z^N (aF_* + bG_*)^* = a^* z^s (F_*)^* + b^* z^t (G_*)^* = a^* z^{s'} F + b^* z^{t'} F,$$

hence $z^r H = AF + BG$ for suitable forms $A, B \in K[x,y,z]$. By the proof of Bezout, $K[x,y,z]/\langle F, G \rangle \to K[x,y,z]/\langle F, G \rangle, \Lambda \mapsto z^r \Lambda$ is injective, hence

$$H = A'F + B'G$$

for some $A', B' \in K[x,y,z]$, hence writing $A' = A'_{\deg\ H - \deg\ F} + \ldots$ and $B' = B'_{\deg\ H - \deg\ G} + \ldots$, as a result of cancellations

$$H = A'_{\deg\ H - \deg\ F} F + B'_{\deg\ H - \deg\ G} G.$$

$\square$

Proposition 5.3.60. *Let F,G,H be projective plane curves and $P \in F \cap G$. Noether's condition at P are satisfied if:*

1. *F and G intersect transversally at P and $P \in H$.*

2. *P is simple on F and $I(P,H \cap F) \geq I(P,G \cap F)$.*

3. *F and G have distinct tangents at P and $m_P(H) \geq m_P(F) + m_P(G) - 1$.*

Proof. 1. Note that 1. implies 2. since then $I(P,H \cap F) \geq 1 = I(P,F \cap G)$, so it is sufficient to prove 2.
2. We get that $\mathbf{ord}_P^F(H) = I(P,H \cap F) \geq I(P,G \cap G) = \mathbf{ord}_P^F(F)$, hence $\mathcal{O}_P(F) \ni \overline{H_*} = uL^k$ and $\mathcal{O}_P(F) \ni \overline{G_*} = uL^h$ with $h \leq k$, hence $\overline{H_*} \in \langle \overline{G_*} \rangle \in \mathcal{O}_P(F)$. Then using $\mathcal{O}_P(F)/\langle \overline{G_*} \rangle \simeq \mathcal{O}_P(\mathbb{P}^2)/\langle F_*,G_* \rangle$, we find that $H_* + \langle F_*,G_* \rangle = 0$.
3. WLOG $P = [0,0,1]$. Note that then $m_P(H_*) \geq m_P(F_*) + m_P(G_*) - 1$. This implies that $H_* \in \langle x,y \rangle^{m_P(F_*)+m_P(G_*)-1}$, hence by Lemma 4.4.48, $0 = H_* + \langle F_*,G_* \rangle \in \mathcal{O}_{(0,0)}(\mathbb{A}^2)/\langle F_*,G_* \rangle \simeq \mathcal{O}_P(\mathbb{P}^2)/\langle F_*,G_* \rangle$ □

Corollary 5.3.61. *Let F and G be projective plane curves with no common components. Then there is a curve B where $B \bullet F = H \bullet F - G \bullet F$ if one of following two conditions are satisfied:*

1. *F and G intersect in $(\deg F)(\deg G)$ points and H passes through each of these points.*

2. *All points of $F \cap G$ are simple points of F and $H \bullet F \geq G \bullet F$.*

Proof. If 1. is satisfied, then F and G intersect transversally at every point of intersection, hence by 1. of the prior proposition Noether's condition is satisfied at every point of intersection.
If 2. is satisfied then $I(P,H \cap F) \geq I(P,G \cap F)$, hence 2. of the prior propositions shows that Noether's condition is satisfied at every $P \in F \cap G$.
In either case, this means $H = AF + BG$ for suitable forms A,B by Max Noether's Fundamental Theorem. □

Proposition 5.3.62. *Let F,G,H be plane curves where $\gcd(F,G) = 1$, $P \in F \cap G$.*

1. *When P is simple, then Noether's condition at P is satisfied at P if and only if $I(P,F \cap H) \geq I(P,F \cap G)$.*

2. *When F and G meet transversally at P, then Noether's condition at P is satisfied at P if and only if $P \in H$.*

Proof. WLOG $P = [0,0,1]$. 1"$\Rightarrow$ ": Write $H_* = \frac{\alpha}{\beta}F_* + \frac{\lambda}{\mu}G_*$ for some $\frac{\alpha}{\beta}, \frac{\lambda}{\mu} \in \mathcal{O}_{0,0}(\mathbb{A}^2)$. Then

$$\zeta H_* = \alpha F_* + \lambda G_*, \quad (\zeta := \beta\mu).$$

Then

$$\begin{aligned} I(P, H \cap G) &= I((0,0), H_* \cap G_*) = I((0,0), \zeta H_* \cap G_*) = I((0,0), \alpha F_* \cap G_*) \\ &= I((0,0), \alpha \cap G_*) + I((0,0), F_* \cap G_*) \geq I((0,0), F_* \cap G_*) = I(P, F \cap G) \end{aligned}$$

" $\Leftarrow$": Follows from Proposition 5.3.60.
2.Follows from 1. □

Remark 5.3.63. The above shows that in Proposition 5.3.60 in case 1. resp. case 2. if we presuppose the condition on F and G, then the condition on H is equivalent to Noether's condition being satisfied for F, G and P. The next example shows that the same augmentation can not be made in case 3.

Example 5.3.64. Consider $F = x^2 + x + y$, $G = x^2$ and $H = x + y$ and $P = [0,0,1]$. Note that F and G have distinct tangents, namely $x + y$ resp. x. One sees that $H_* = H = F - G = F_* - G_*$, so Noether's condition is satisfied at P wrt. F and G. Note that $1 = m_P(H) < 2 = m_P(F) + m_P(G) - 1$. So the condition on H in case 3. in Proposition 5.3.60 is not equivalent to Noether's condition on P for curves F, G, H with $\gcd(F,G) = 1$, $P \in F \cap G$ and F and G having distinct tangents.

Proposition 5.3.65. *Let F be an irreducible projective plane curve. Suppose $z \in K(F)$ is given such that $z \in \mathcal{O}_P(F)$ for every $P \in F$. Then $z \in K$.*

Proof. write $z = \frac{H + \langle F \rangle}{G + \langle F \rangle}$, for some equidegree forms $H, G \in K[\mathbf{x}]$, where $G(P) \neq 0$ for every $P \in F$. Then $F \cap G = \emptyset$ hence Noether's condition is vacuously satisfied for every $P \in F \cap G$. Then by MNFT there are forms $A, B \in K[\mathbf{x}]$ such that $\deg AF = \deg H$, $\deg BG = \deg H = \deg G$. Then $B \in K$ and

$$z = \frac{H + I(F)}{G + I(F)} = \frac{AF + BG + I(F)}{G + I(F)} = \frac{BG + I(F)}{G + I(F)} = B + I(F) \in K.$$

□

5.3.6 Applications of Noether's Theorem

Proposition 5.3.66. *Let C, C' be cubics such that $C \bullet C' = \sum_1^9 P_i$. Let Q be a conic such that $Q \bullet C = \sum_1^6 P_i$. Assume $P_1, \ldots, P_6$ are simple on C. Then P_7, P_8 and P_9 lie on a line.*

Proof. Since $P_1,\ldots,P_6$ are distinct, $I(P_i,C'\cap C)\geq 1 = I(P_i,Q\cap C)$ for $i=1,\ldots,6$. For $i=7,8,9$ (7 8 9?! I guess that's why 6 is afraid of 7), $I(P_i,C'\cap C)\geq 1\geq I(P_i,Q\cap C)$, since $I(P_i,Q\cap C)=1$ if $P_i\in Q\cap C$ and 0 otherwise. It follows that $C'\bullet C\geq Q\bullet C$. By Corollary 5.3.61 there is an L such that $L\bullet C = C'\bullet C - Q\bullet C = P_7+P_8+P_9$. Since $\deg\ C=3$ and $(\deg\ L)(\deg\ C)=\deg\ L\bullet C=3$, we get that $\deg\ L=1$, hence L is a line passing through P_7,P_8,P_9. □

Definition 5.3.67. Given $2n$ points in the projective plane $P_1,\ldots,P_{2n}$, we form a $2n$-gon by connecting these points via $2n$ lines $L_i := L(P_i,P_{i+1})$ for $i\in\{1,\ldots,2n-1\}$ and $L_{2n} := L(P_{2n},P_1)$. For the first n lines, we define the *opposite side of* L_i, denoted $\mathsf{op}(L_i)$ to be L_{i+n}. For the remaining lines, the opposite to L_i is L_{i-n}.

Remark 5.3.68. Thus we get bijection of any collection of n line segments, LS satisfying that the opposite segment of any $L\in LS$ is not LS, to the opposite line segments to those in LS, via op.

Corollary 5.3.69. *(Pascal's Theorem) Let Q be an irreducible conic (an ellipse, parabola or hyperbola). Pick 6 distinct points on Q, in some ordering $P_1,\ldots,P_6$ and form the hexagon containing these points. Then the points $P_i := L_i\cap \mathsf{op}(L_i)$, $i:=1,2,3$ lie on a line. More generally given a set $LS := \{L_{i_1},L_{i_2},L_{i_3}\}$, satisfying the conditions of the prior remark, the points $Q_j := L_{i_j}\cap\mathsf{op}(L_{i_j})$ lie on a line.*

Proof. Indeed, consider the cubics $C := L_{i_1}L_{i_2}L_{i_3}$ and $C' := \mathsf{op}(L_{i_1})\mathsf{op}(L_{i_2})\mathsf{op}(L_{i_3})$ and apply the prior proposition. □

Corollary 5.3.70. *(Pappus' Theorem) Let L_1,L_2 be two lines, $P_1,P_2,P_3\in L_1$ and $Q_1,Q_2,Q_3\in L_2$ such that $P_i,Q_j\notin L_1\cap L_2$. Set $L_{ij} := L(P_i,Q_j)$. For each i,j,k with $\{i,j,k\}=\{1,2,3\}$, set $R_k := L_{ij}\bullet L_{ji}$. Then R_1,R_2,R_3 lie on a line.*

Proof. Indeed, Consider $Q := L_1L_2$, $C := L_{12}L_{13}L_{23}$ and $C' := L_{21}L_{31}L_{32}$. By construction non of the L_{ij}'s are components of Q. We see that $C\bullet C' = \sum_1^3 P_i+Q_i+R_i$ and $Q\bullet C = \sum_1^3 P_i+Q_i$ and also note that we chose P_i,Q_i such that they are simple. We are thus in position where we can apply the proposition to see that R_1,R_2,R_3 are on a line. □

Proposition 5.3.71. *Let C,C',C'' be cubics with C irreducible, and $Q\in C$. Suppose $C'\bullet C=\sum_1^9 P_i$, where the P_i are simple but not necessarily distinct points on C and that $C''\bullet C = Q+\sum_1^8 P_i$, then $Q=P_9$.*

Proof. Take any line L passing through P_9. Note that $L \bullet C = P_9 + R + S$ for some $R,S \in C$. Then $LC'' \bullet C = C' \bullet C + Q + R + S$. By Noether's Theorem there is a line L' passing through Q,R,S. But then $L = L'$, hence $P_9 + R + S = L \bullet C = L' \bullet C = Q + R + S$, meaning $P_9 = Q$. □

Definition 5.3.72. Let C be a non-singular cubic. Take any two points P,Q on C. Take the unique line L such that $L \bullet C = P + Q + R_{P,Q}$ for some unique $R_{P,Q} \in C$. When $P \neq Q$, $L = L(P,Q)$ and when $P = Q$, L is the tangent at P. Define a binary operation on C

$$\begin{aligned} \boxplus : C \times C &\to C \\ (P,Q) &\mapsto R_{P,Q} \end{aligned}$$

Remark 5.3.73. With this operation C becomes a commutative magma I.e. $\boxplus$ is a commutative operation.

Definition 5.3.74. Let C be a non-singular cubic. Choose any point O on C. Define a binary operation

$$\begin{aligned} \oplus_O := \oplus : C \times C &\to C \\ (P,Q) &\mapsto O \boxplus (P \boxplus Q) \end{aligned}$$

Proposition 5.3.75. *Let C be a non-singular cubic. With $\oplus$, C becomes an additive group with O being the identity.*

Proof. We first show that the operation is associative. Let $P,Q,R \in C$. We pick unique lines L_1,L_2,L_3,M_1,M_2,M_3 such that

$$\begin{aligned} L_1 \bullet C &= P + Q + P \boxplus Q \\ L_2 \bullet C &= P \oplus Q + R + (P \oplus Q) \boxplus R \\ L_3 \bullet C &= O + Q \boxplus R + \underbrace{O \boxplus (Q \boxplus S)}_{Q \oplus R} \\ M_1 \bullet C &= O + P \boxplus Q + \underbrace{O \boxplus (P \boxplus Q)}_{P \oplus Q} \\ M_2 \bullet C &= Q + R + Q \boxplus R \\ M_3 \bullet C &= P + Q \oplus R + P \boxplus (Q \oplus R) \end{aligned}$$

Set $C' := L_1 L_2 L_3$ and $C'' := M_1 M_2 M_3$. Then

$$\begin{aligned} C' \bullet C &= O + P + Q + R + P \boxplus Q + Q \boxplus R + P \oplus Q + Q \oplus S + (P \oplus Q) \boxplus R, \\ C'' \bullet C &= O + P + Q + R + P \boxplus Q + Q \boxplus R + P \oplus Q + Q \oplus S + P \boxplus (Q \oplus R). \end{aligned}$$

Applying the prior proposition, it follows that $(P \oplus Q) \boxplus R = P \boxplus (Q \oplus R)$. Then $O + (P \oplus Q) \boxplus R (P \oplus Q) \oplus R = O + P \boxplus (Q \oplus R) + P \oplus (Q \oplus R)$, hence $(P \oplus Q) \oplus R = P \oplus (Q \oplus R)$. Moreover, the line through O and $O \oplus P$ is the line through O and P, hence $O \oplus P = P$. Define $-P := P \boxplus (O \boxplus O)$. Then the third intersection of C with the line through P and $-P$ is $O \boxplus O$. Moreover the third point of intersection of C with the line through O and $O \boxplus O$ is O, hence $P \oplus -P = O \boxplus (P \boxplus -P) = O \boxplus (O \boxplus O) = O$. Commutativity, follows from commutativity of $\boxplus$. ☐

Definition 5.3.76. Let C be a cubic with no multiple components. Define

$$C^o := \{P \in C : P \text{ simple}\}.$$

Remark 5.3.77. When C is irreducible, we may define $\boxplus$ as in the non-singular case, since $L(P,Q)$ is not a component of C in this case there is a unique point $R_{P,Q}$ such that $L(P,Q) \bullet C = P + Q + R_{P,Q}$. We see that $I(R_{P,Q}, L(P,Q) \cap C) = 1$, hence $R_{P,Q}$ is simple, meaning $\boxplus$ is indeed well-defined. Pick a point $O \in C^o$. Upon for $P,Q \in C^o$ defining, $P \oplus Q := O \boxplus (P \boxplus Q)$, the same computations as in the nonsingular case shows that $(C^o, \oplus)$ is an additive group.
If C has a non-trivial component Then I don't know.

Proposition 5.3.78. *Consider an irreducible cubic C. Let $O, O' \in C^\circ$ be given. Set $Q := O \boxplus O'$. Then*

$$\begin{aligned} \alpha : (C^\circ, \oplus_O) &\to (C^\circ, \oplus_{O'}) \\ P &\mapsto Q \boxplus P \end{aligned}$$

defines a group isomorphism.

Proof. ☐

Proposition 5.3.79. *Let P_1, P_2, P_3 be distinct points on an irreducible conic Q. Let L_1, L_3, L_5 be the tangents at each of the respective points. Let $L_2 := L(P_1, P_2)$, $L_4 := L(P_2, P_3)$ and $L_6 := L(P_3, P_1)$. Then $Q_i := L_i \cap \mathrm{op}(L_i) = L_i \cap L_{i+3}$ for $i = 1,2,3$ are collinear*

Proof. Set $C = L_1 L_3 L_5$, $C' := L_2 L_4 L_6$. For each i note that there are exactly two lines in C' intersecting P_i. Since neither of these points are tangents to Q, they intersect the tangent to P_i transversally. It follows that $I(P_i, C \cap C') = 2$, hence

$$C \bullet C' = \sum_1^3 2P_i + Q_i.$$

Additionally

$$Q \bullet C' = \sum_1^3 2P_i,$$

it follows from a Corollary 5.3.61 2. that there is a line L for which $L \bullet C' = \sum_1^3 Q_i$, hence Q_1, Q_2, Q_3 are collinear. □

Proposition 5.3.80. *Let $P_1, \ldots, P_5$ be distinct points on an irreducible conic Q. Let L_1 be the tangent at P_1 and connect the points with 5 lines L_i as before. Defining Q_1, Q_2, Q_3 as before, we get that these are collinear.*

Proof. Define C and C' as before. Then

$$C \bullet C' = 2P_1 + \sum_2^5 P_i + Q_1 + Q_2 + Q_3$$

and

$$C \bullet Q = 2P_1 + \sum_2^5 P_i.$$

We get a line through Q_1, Q_2, Q_3 in the usual way. □

Remark 5.3.81. The above result gives a way to construct a tangent to a point on an irreducible conic using only a straightedge: Call the point P. Draw 4 other distinct points also distinct from P. Call these points $P_1, \ldots, P_4$. Find the intersection of the line $L(P_4, P)$ and $L(P_1, P_2)$. Call that point Q_1. Do the same with $L(P_1, P)$ and $L(P_3, P_4)$. Call that point Q_2. Find the intersection of $L(P_2, P_3)$ and $L(Q_1, Q_2)$. Call that point Q_3. Then $L(P, Q_3)$ is the tangent of the conic at P.

Proposition 5.3.82. *Consider the hexagon formed by points $P_1, \ldots, P_6$. Suppose the intersections of the sides with their opposite side lie on a line. Then $P_1, \ldots, P_6$ lie on a conic.*

Proof. □

Proposition 5.3.83. *Let C be an irreducible cubic and L a line such that $L \bullet C = P_1 + P_2 + P_3$ for P_1, P_2, P_3 distinct. Let L_i be the tangent to C at P_i. Then $L_i \bullet C = 2P_i + Q_i$ for some Q_i. Then Q_1, Q_2, Q_3 are collinear.*

Proof. □

Proposition 5.3.84. *On a cubic, a line through two flexes on the cubic, passes through a third flex on the cubic.*

Proof. □