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 $f \in \mathbf{Hom}(A,B)$, $g \in \mathbf{Hom}(B,A$. 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, $f \in \mathbf{Hom}(A,B)$. Suppose $g,h \in \mathbf{Hom}(B,A)$ 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, $f_1 = f_2$. □

Example 1.3.10.

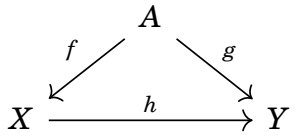
1. Let $\mathbf{Set}$ be defined by objects being sets and morphisms being functions. Indeed, letting $\circ$ be composition in the conventional way and letting $\mathbb{1}_X = \mathrm{id}_X : X \to X, x \mapsto x$, we see that this indeed defines a category.

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

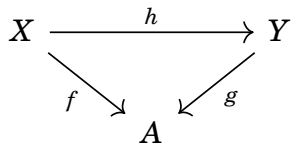
   $$(a,b)(b,c) := (a,c).$$

   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.

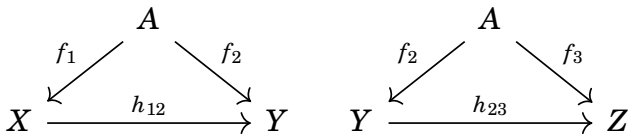
Definition 1.3.11. Let $\mathcal{C}$ be a category and A an object in $\mathcal{C}$. *The slice category of $\mathcal{C}$ under A* denoted $A/\mathcal{C}$ is the category whose objects are morphisms in $\mathbf{Hom}(\mathcal{C})$ with domain A and where a morphism from $f : A \to X$ and $g : A \to Y$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



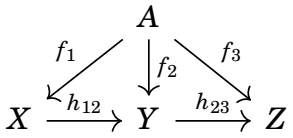
commutes. *The slice category of* $\mathcal{C}$ *over* A denoted $\mathcal{C}/A$ is $A/\mathcal{C}^{\mathrm{op}}$, i.e. objects are morphisms with codomain A and a morphism from $f : X \to A$ to $g : Y \to A$ is a morphism $h : X \to Y$ satisfying



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



to obtain the commutative diagram



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 $\text{Ob}(\mathcal{C}^{\text{op}}) := \text{Ob}(\mathcal{C})$ and where a morphism $f : A \to B$ in $\text{Hom}(\mathcal{C}^{\text{op}})$ is a morphism $f : B \to A$ in $\text{Hom}(\mathcal{C})$

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 $\text{Hom}(\mathcal{C}^{\text{op}})$. 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 $\text{Hom}(\mathcal{C}^{\text{op}})$ to be the identity morphism in $\text{Hom}(\mathcal{C})$, 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 $\text{Hom}(\mathcal{C}_1)$, then $\mathcal{F}(f) : \mathcal{F}(B) \to \mathcal{F}(A)$ is an isomorphism in $\text{Hom}(\mathcal{C}_2^{\text{op}})$*

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 $\text{Hom}(\mathcal{C})$,

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

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 $\text{Hom}(A, B)$ is a set for every object A, B in $\text{Ob}(\mathcal{C})$. It is *small* if $\text{Ob}(\mathcal{C})$ 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 $f \in \mathbf{Hom}(A,B)$ is a *monomorphism* if for every pair of morphisms $g_1, g_2 \in \mathbf{Hom}(C,A)$,

$$f g_1 = f g_2 \Rightarrow g_1 = g_2.$$

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

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

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

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

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 $\mathbf{Monoid}$ 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 $f, g \in \mathbf{Hom}^{\mathrm{Mon}}(M, N)$ and $x, y \in M$. Then

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

Furthermore we have

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

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

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

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 $\ker\ \rho = \rho^{-1}(\{e_N\})$ (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

$$(m_\alpha)\left((m'_\alpha)(m''_\alpha)\right) = (m_\alpha)(m'_\alpha m''_\alpha) = (m_\alpha(m'_\alpha m''_\alpha)) = ((m_\alpha m'_\alpha)m''_\alpha) = (m_\alpha m'_\alpha)(m''_\alpha) = \left((m_\alpha)(m'_\alpha)\right)(m''_\alpha)$$

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 $\alpha_1, \ldots, \alpha_r \in A$ and $\beta_1, \ldots, \beta_p \in A$, $m_\alpha = e_\alpha$ for every $\alpha \in A \setminus \{\alpha_1, \ldots, \alpha_r\}$ and $m'_\alpha = e_\alpha$ for every $\alpha \in A \setminus \{\beta_1, \ldots, \beta_p\}$. Then $m_\alpha m'_\alpha = e_\alpha$ for every $\alpha \in A \setminus \{\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_p\}$ 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 $i \in \{n+1, \ldots, n+m\}$. 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 $G = \langle g_1, \ldots, g_n \rangle$ for some $g_1, \ldots, g_n \in G$. 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 $g_1, g_2 \in G$,

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

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

$$s((h_\alpha) + (h'_\alpha)) = s((h_\alpha + h'_\alpha)) = \sum_{\alpha\in A} \left(h_{\alpha_i} + h'_{\alpha_i}\right) = \sum_{1}^{n} \left(h_{\alpha_i} + h'_{\alpha_i}\right) = \sum_{1}^{n} h_{\alpha_i} + \sum_{1}^{n} h'_{\alpha_i}$$
$$= \sum_{\alpha\in A} h_\alpha + \sum_{\alpha\in A} h'_\alpha = s((h_\alpha)) + s((h'_\alpha)).$$

□

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 $i\in\{1,\dots,n\}$. 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 $\alpha\in A\setminus\{\alpha_1,\dots,\alpha_n\}$ 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 $k \in \ker\ \rho$. Then

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

hence $g(\ker\ \rho)g^{-1} \subset \ker\ \rho$. Conversely $k \in g(\ker\ \rho)g^{-1}$ since $k = g\left(g^{-1}kg\right)g^{-1}$ and $g^{-1}kg \in \ker\ \rho$ 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 $(g_1H, g_2H) = (g_1'H, g_2'H)$, then trivially

$$(g_1g_2)H = (g_1H)(g_2H) = (g_1'H)(g_2'H) = (g_1'g_2')H.$$

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

   $$r_1 * (r_2 * r_3) = (r_3 r2) r_1 = r_3(r_2 r_1) = (r_1 * r_2) * r_3$$

   and

   $$r_1 * 1 = 1 r_1 = r_1 = r_1 1 = 1 * r_1$$

   and lastly

   $$r_1 * (r_2 + r_3) = (r_2 + r_3) r_1 = r_2 r_1 + r_3 r_1 = r_1 * r_2 + r_1 * r_3,$$

   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

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

□

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 $a \mapsto [(a,0)]$. □

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 ${}_R M := M$.
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 ${}_R M_S := M$.

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

$$m(r_1 * r_2) = (r_2 r_1)m = r_2(r_1 m) = r_2(m r_1) = (m r_1) r_2.$$

"$\Leftarrow$": 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 ${}_R M, {}_R N$ 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. $(r_1+r_2)(m_\alpha)=((r_1+r_2)m_\alpha)=(r_1m_\alpha+r_2m_\alpha)=(r_1m_\alpha)+(r_2m_\alpha)=r_1(m_\alpha)+r_2(m_\alpha)$.
2. $r_1((m_\alpha)+(m'_\alpha))=r_1(m_\alpha+m'_\alpha)=(r_1m_\alpha+r_1m'_\alpha)=(r_1m_\alpha)+(r_1m'_\alpha)=r_1(m_\alpha)+r_1(m'_\alpha)$.
3. $1(m_\alpha)=(1m_\alpha)=(m_\alpha)$.
4. $(r_1r_2)(m_\alpha)=((r_1r_2)m_\alpha)=(r_1(r_2m_\alpha))=r_1(r_2m_\alpha)=r_1(r_2(m_\alpha))$.

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

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

□

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 $\beta\in A$,

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

Furthermore, we have that

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

□

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

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

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 $\alpha_1,\dots,\alpha_n \in A$. Let $j \in \{1,\dots,n\}$. Then

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

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 $m_1,\dots,m_n \in M$ and every finite sequence $r_1,\dots,r_n \in R$,

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

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 $x_1,\dots,x_n \in X\setminus\{y\}$ and $r_1,\dots,r_{n+1} \in R$. 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

$$r_1r_2 - r_1'r_2' = r_1r_2 - r_1r_2' + r_1r_2' - r_1'r_2' = r_1(r_2 - r_2') + (r_1 - r_1')r_2' \in I,$$

hence $r_1r_2 + I = r_1'r_2' + I$ 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. "⇒": 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.
"⇐": 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. □

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

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

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

$$r(a_1 a_2) = (ra_1)a_2$$

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

$$A \underset{i_1}{\overset{p_1}{\rightleftarrows}} A \oplus B \underset{i_2}{\overset{p_2}{\rightleftarrows}} B$$

$i_1 p_1 = \mathbb{1}_A$, $p_2 i_2 = \mathbb{1}_B$ 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

$$\begin{array}{ccc} K & \xrightarrow{0} & B \\ \downarrow{\scriptstyle k} & \nearrow{\scriptstyle f} & \\ A & & \end{array}$$

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 $\mathrm{im}\ \rho = \ker \rho'$. 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.⇒ 2.": by exactness $\ker\rho = \operatorname{im} 0 = 0$ and $\operatorname{im}\ \rho' = \ker\ 0 = M''$. "2.⇒ 1.": $\operatorname{im} 0 = 0 = \ker\ \rho$ and $\ker\ 0 = M'' = \operatorname{im}\ \rho'$ □

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

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

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 $\ker\ \rho = \{e\}$ 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 $k \in \ker\ \rho$. Then $\rho(k) = e = \rho(e)$, implying $k = e$, hence $k \in \{e\}$. □

Corollary 2.6.10. *Let $\sigma : R \to S$ be a ring homomorphism. Then $\ker\ \sigma = 0$ 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

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

□

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,

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

□

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. □

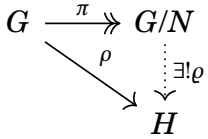
Corollary 2.6.14. *Let R be a ring, M, N left/right R-modules, $\rho : M \to N$ a left/right module homomorphism and $L \subset M$, $K \subset N$ be submodules. Then the $\vartheta : M/L \to N/K$, $m+L \mapsto \rho(m)+K$ is a well-defined module homomorphism if and only if $\rho(L) \subset K$*

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 $N \subset \ker \rho$. 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*



commutes.

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

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

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 $M/\ker\ \rho \simeq \rho(M)$ via the left/right R-module homomorphism $\varrho : M/\ker\ \rho \to \rho(M)$, $m + \ker\ \rho \mapsto \rho(m)$*

Proof. Theorem 2.6.18 and Corollary 2.6.17 ensures that $\varrho : M/\ker\ \rho \to \rho(M)$ 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 $A/\ker\ \sigma \simeq \sigma(A)$ via the R-algebra homomorphism $\vartheta : A/\ker\ \sigma \to \sigma(A), a + \ker\ \sigma \mapsto \sigma(a)$.*

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 $N' \subset \ker\ \pi$. Clearly $N'/N \subset \ker\ \varpi$. Let $gN' \in \ker\ \varpi$. Then $gN = 0$, hence $g \in N$. Thus $gN' \in N/N'$. Then $\ker\ \varpi = N/N'$, hence by the 1st isomorphism theorem

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

□

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. □

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 □

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.

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

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

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

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

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:

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

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 $W := \text{im}\ \rho_2 = \ker\ \rho_3$. 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. □

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

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

□

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

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

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

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

□

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

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

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

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

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 $m_\lambda = \mathrm{id}$, hence it is continuous.

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

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

is well-defined, since

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

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 $\iota : K^n \to \pi^{-1}(U_i) \setminus 0, v \mapsto (v_1, \ldots, v_{i-1}, 1, v_{i+1}, \ldots, v_{n+1})$ with $\pi|_{\pi^{-1}(U_i)} : \pi^{-1}(U_i) \to U_i$, which are continuous maps. Let $U \subset K^n$ be open. Let $Q := \{v \in K^{n+1} : v_i \in K, (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{n+1}) \in KU\} \simeq_{S_{n+1}} K \times KU$. 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 $\pi|_{S^n(\mathbb{R})} : S^n(\mathbb{R}) \to \mathbb{P}^n(\mathbb{R})$, $S^n(K) = \{v \in \mathbb{R}^{n+1} : \|v\| = 1\}$. 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 $[v_1], \ldots, [v_m] \in \mathbb{P}^n$ we define

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

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

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

and conversely

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

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

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

Proof. "$\Rightarrow$": Suppose there is a $[v_i] \in \mathbf{Span}([v_1], \ldots, \widehat{[v_i]}, \ldots, [v_m])$. Then for some $\lambda_1, \ldots, \lambda_m \in K^m$ 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 $v_1, \ldots, v_m$ are not linearly independent. "$\Leftarrow$": Suppose there are $(\lambda_1, \ldots, \lambda_m) \in K^m \setminus 0$ such that $\sum_1^m \lambda_i v_i = 0$. Then for some $j \in \{1, \ldots, m\}$, $\lambda_j \neq 0$, hence

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

□

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 $rr_i = r_i r = 1$. Then

$$r(r_1 - r_2) = rr_1 - rr_2 = 1 - 1 = 0 \Rightarrow r_1 - r_2 = r_1 r(r_1 - r_2) = 0 \Rightarrow r_1 = r_2.$$

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

$$(rr')\left(r'^{-1}r^{-1}\right) = r\left(r'\left(r'^{-1}r^{-1}\right)\right) = r\left(\left(r'r'^{-1}\right)r^{-1}\right) = r\left(er^{-1}\right) = rr^{-1} = 1,$$

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. "⇒": 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.
"⇐": 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. "⇒": 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 $b \in I$. □

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. □

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. □

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 $p \mid x$. □

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

$$\mathbf{rad}(I) = \{r \in R : r^n \in I \text{ for some } n > 0\}.$$

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.

$$\lambda_{i,j} \in R \quad (i \in \{1,\dots,m\}, j \in \{1,\dots,2n\})$$

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. $\operatorname{lcm}(a, 0) =$ 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 $p_1,\dots,p_n \in R$ unique and $v_1,\dots,v_n \geq 1$ such that

$$r = \prod_1^n p_i^{v_i}.$$

Remark 2.8.45. By uniqueness we more precisely mean that given $q_1,\dots,q_m \in R$ another sequence of distinct irreducible elements and $w_1,\dots,w_m \geq 1$ such that

$$r = \prod_1^m q_i^{w_i},$$

then $m = n$ and there is some bijection $\omega : \{1,\dots,n\} \to \{1,\dots,n\}$ (i.e. a permutation $\omega \in \mathscr{S}_n$) and units $a_1,\dots,a_n \in R$ such that

$$p_i = a_i q_{\tau(i)} \text{ and } v_i = w_{\tau(i)},$$

for each $i \in \{1,\dots,n\}$.

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 irreduciblity of p. We can then find irreducible $q_1,\dots,q_n,p_1,\dots,p_m,p_{m+1},\dots,p_l \in R$ and $v_1,\dots,v_m,w_1,\dots,w_m,w_{m+1},\dots,v_l \geq 1$ 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 $i \in \{1,\dots,l\}$ and a unit $s \in R$ such that $w_i = 1$ and $p = sp_i$. If $i \in \{1,\dots,m\}$ then

$$a = \prod_1^m p_j^{w_j} = s^{-1} p \prod_{j \in \{1,\dots,m\} \setminus \{i\}} p_j^{w_j} \Rightarrow p \mid a.$$

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 $v_i = w_{\tau(i)}$. □

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

$$\langle a \rangle \subset \langle a_1 \rangle \subset \dots$$

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

$$\frac{r_1}{x_1} + \left(\frac{r_2}{x_2} + \frac{r_3}{x_3}\right) = \frac{r_1}{x_1} + \frac{r_2 x_3 + r_3 x_2}{x_2 x_3} = \frac{r_1 x_2 x_3 + r_2 x_3 x_1 + r_3 x_2 x_1}{x_1 x_2 x_3} = \frac{(r_1 x_2 + r_2 x_1) x_3 + r_3 x_2 x_1}{x_1 x_2 x_3}$$
$$= \frac{r_1 x_2 + r_2 x_1}{x_1 x_2} + \frac{r_3}{x_3} = \left(\frac{r_1}{x_1} + \frac{r_2}{x_2}\right) + \frac{r_3}{x_3}.$$

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 $1:=\frac{1}{1}$. 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 im $R\hookleftarrow X^{-1}R$ is a subring of $X^{-1}R$ (or that $R\hookleftarrow X^{-1}R$ is a ring homomorphism), hence 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

$$\widehat{X}:=\{r\in R:\exists r'\in R, r'r\in X\}.$$

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 $r \in X$. □

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*

$$\begin{aligned} \iota : R &\hookrightarrow X^{-1}R \\ r &\mapsto \frac{r}{1} \end{aligned}$$

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

$$sr = sr\frac{1}{y}y = r\frac{s}{y}y = y \in X \Rightarrow r \in \widehat{X}.$$

□

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. □

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 $\sigma \in \mathbf{Hom}((R,X),(S,Y))$ and $\tau \in \mathbf{Hom}((S,Y),(T,Z))$, then $\sigma(X) \subset Y$ hence

$$\tau \circ \sigma(X) = \tau(\sigma(X)) \subset \tau(Y) \subset Z \subset T \setminus 0.$$

Clearly $\mathbb{1}_{(R,X)} := \mathrm{id}_R \in \mathbf{Hom}((R,X),(R,X))$. □

Proposition 2.8.74. *Call the category described in the above lemma $\mathscr{C}$. The assignment of a pair $((R,X),\sigma)$ in $(Ob(\mathscr{C}),\mathbf{Hom}(\mathscr{C}))$ to $(X^{-1}R,\overline{\sigma})$ in the category of integral domains defines a covariant functor.*

Proof. Let $\tau \in \mathbf{Hom}((S,Y),(T,Z))$, $\sigma \in \mathbf{Hom}((R,X),(S,Y))$ 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. "⇒": 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 $u = u'$.
"⇐": 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 $r = \min\{n \geq 0 : t^n \in I\}$. 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 $\nu_p\left(\frac{a}{n}\right) = \mathbf{max}(\{n \geq 0 : p^n \mid \frac{a}{n}\})$. 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 $\frac{a}{b} = \frac{s}{t}p^{\nu_p(a)-\nu_p(b)} = \frac{s}{t}p^{\mathbf{ord}(\frac{a}{b})}$, 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* $z = \frac{a}{b}t^{\mathbf{ord}(z)} \in K \setminus R$*, 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 $a, b \in K$, $\nu(ab) = \nu(a) + \nu(b)$,

3. for every $a, b \in K$, $\nu(a+b) \geq \min(\nu(a), \nu(b))$.

Definition 2.8.87. Let K be a field with an order function ν. We define the ring induced by ν to be the set

$$R_\nu(K) := R_\nu := \nu^{-1}(\mathbb{Z}_{\geq 0} \cup \{\infty\}) = \{r \in K : \nu(r) \geq 0\}.$$

Define the ideal induced by ν to be the set

$$\mathfrak{m}_\nu := \nu^{-1}(\mathbb{Z}_{\geq 1} \cup \{\infty\}) = \{r \in R_\nu : \nu(r) > 0\} \subset R_\nu$$

Lemma 2.8.88. *Let K be a field with an order function ν. We collect the following facts about R_ν and $\mathfrak{m}_\nu$:*

(i) R_ν is subring of K and hence is an integral domain.

(ii) For every $u \in R_\nu$,

$$u \in R_\nu^* \iff \nu(u) = 0.$$

(iii) $\mathfrak{m}_\nu$ is the unique maximal ideal of R_ν, hence R_ν is local.

(iv) $R_\nu = K$ if and only if ν is trivial. If ν is non-trivial, R_ν is not a field.

Proof. (i) Property 2. ensures that R_ν is closed under multiplication while property 3. ensures that R_ν is closed under addition. $0 \in R_\nu$ by property 1. Note that $\nu(1) = \nu(1 \cdot 1) = \nu(1) + \nu(1)$, hence $\nu(1) = 0$, hence $1 \in R_\nu$. Let $u \in R_\nu^*$. Then $0 = \nu(1) = \nu(uu^{-1}) = \nu(u) + \nu(u^{-1})$, hence $\nu(u^{-1}) = -\nu(u)$ and since $\nu(u) \geq 0$, it follows that $\nu(u^{-1}) = 0$. We thus in particular find that $\nu(-1) = 0$, hence for any $r \in R_\nu$ we have that $\nu(-r) = \nu(-1 \cdot r) = \nu(-1) + \nu(r) = \nu(r)$. It thus follows that $-r \in R_\nu$. We thus get that R_ν is a subring of K.
(ii) "$\Rightarrow$": This was already proven in the proof of (i).
"$\Leftarrow$": Let $u \in R_\nu$ such that $\nu(u) = 0$. For some $v \in K \setminus 0$, $uv = vu = 1$. Then $0 = \nu(1) = \nu(uv) = \nu(u) + \nu(v) = \nu(v)$, hence $v \in R_\nu$, which means u is a unit in R_ν.
(iii) $\mathfrak{m}_\nu = R_\nu \setminus R_\nu^*$, hence it is sufficient to prove that $\mathfrak{m}_\nu$ is an ideal by Proposition 2.8.59. $\mathfrak{m}_\nu$ is closed under addition by property 3. Let $r \in R_\nu$, $x \in \mathfrak{m}_\nu$. Then $\nu(rx) = \nu(r) + \nu(x) \geq 0 + 1 = 1$, hence $rx \in \mathfrak{m}_\nu$. (iv) "$\Rightarrow$": Suppose ν is not trivial. Then there is some $x \in K \setminus 0$ such that $\nu(x) \geq 1$, then $\nu(x^{-1}) \leq 1$, hence $x^{-1} \notin R_\nu$.
"$\Leftarrow$": Suppose ν is trivial. Then for $x \in K$

$$\nu(x) = \begin{cases} 0 & \text{if } x \neq 0 \\ \infty & \text{otherwise} \end{cases},$$

in any case $v(x) \geq 0$, hence $x \in R_v$.

□

Theorem 2.8.89. *Let K be a field and v a non-trivial order function on K. Then $R = \{z \in K : v(z) \geq 0\}$ is a DVR with maximal ideal $\mathfrak{m} := \{z \in R : v(z) > 0\}$ such that $Q(R_v) = K$. Conversely, if R is a DVR with quotient field K, then $\mathbf{ord} : K \to \mathbb{Z} \cup \{\infty\}$ 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_v is a PID. Let $I \subset R_v$ be a non-zero ideal. Let $s \in I$ be given such that $v(s) = \min(v(I))$. Let $x \in I$. For some $q \in K$, $x = qs$. Since

$$v(x) = v(q) + v(s) \Rightarrow v(q) = v(x) - v(s) \geq 0,$$

it follows that $s \mid x$ in R_v, hence $I = \langle s \rangle$. Let $k \in K$. Then either $k \in R_v$ or $k^{-1} \in R_v$. In the first trivially $k \in Q(R_v)$. In the second case $k = (k^{-1})^{-1} = \frac{1}{k^{-1}} \in Q(R_v)$.
For the second statement, we have for $a \in R$ that $\mathbf{ord}(a) = \infty$ 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 $l := \mathbf{ord}(b)$, $m := \mathbf{ord}(a)$.First, we get that

$$ab = uvt^{m+l} \Rightarrow \mathbf{ord}(ab) = m + l = \mathbf{ord}(a) + \mathbf{ord}(b),$$

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 $d := \mathbf{ord}(q) \geq 0$. This implies that $a + b = \alpha t^{d+\min(l,m)}$, hence

$$\mathbf{ord}(a + b) = d + \min(m,l) \geq \min(m,l) = \min(\mathbf{ord}(a), \mathbf{ord}(b)),$$

proving property 3. □

Remark 2.8.90. One notes from the above proof that the uniformizing parameter for R_v is $t \in R_v$ where $v(t) = \min(v(\mathfrak{m}_v))$.

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 $\mathbf{ord}(a_i) < \mathbf{ord}(a_j)$ for every $j \neq i$, then $\mathbf{ord}(\sum_1^n a_j) = a_i$ 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 $m_1 := \mathbf{ord}(a_1) < \mathbf{ord}(a_2) := m_2$. 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 $\mathbf{ord}(a_1) = \mathbf{ord}(a_1 + a_2)$.
WLOG $i = n + 1$. Note that $\mathbf{ord}(a_n + a_{n+1}) = \mathbf{ord}(a_{n+1})$, hence setting $a'_n = a_n + a_{n+1}$, we get $\mathbf{ord}(a'_n) = \mathbf{ord}(a_{n+1}) < a_j$ for every $j \in \{1, \dots, n-1\}$. By induction it follows that

$$\mathbf{ord}(\sum_1^{n+1} a_j) = \mathbf{ord}(\sum_1^{n-1} a_j + (a_n + a_{n+1})) = \mathbf{ord}(\sum_1^{n-1} a_i + a'_n) = \mathbf{ord}(a'_n) = \mathbf{ord}(a_{n+1}).$$

Since $\mathbf{ord}(\sum_1^n a_j) < \mathbf{ord}(a_j) \leq \infty$, 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 $\mu \in K \setminus 0$ $\mathbf{ord}(z - \mu) = \mathbf{ord}(\mu) = 0$, 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 $\lambda_0, \dots, \lambda_n \in K$, $z_n \in R$ 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 $\lambda_0, \dots, \lambda_n, \mu_1, \dots, \mu_n \in K$ 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

$$\sigma:K\to\mathfrak{m}^n/\mathfrak{m}^{n+1}$$
$$\lambda\mapsto\lambda t^n+\mathfrak{m}^{n+1}$$

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 $\mathbf{dim}\ \mathfrak{m}^n/\mathfrak{m}^{n+1}=\mathbf{dim}\ K=1$. □

Lemma 2.8.95. *We keep the setup from the prior results. For each* $n\geq 0$, $\mathbf{dim}\ R/\mathfrak{m}^n=n$. *It follows that* $\mathbf{ord}(z)=\mathbf{dim}\ R/\langle z\rangle=\mathbf{dim}\ R/\mathfrak{m}^{\mathbf{ord}(z)}$ *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

$$\mathbf{dim}\ R/\mathfrak{m}^{n+1}=\mathbf{dim}\ \mathfrak{m}^n/\mathfrak{m}^{n+1}+\mathbf{dim}\ R/\mathfrak{m}^n=n+1,$$

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 $\mathbf{ord}(z)=\mathbf{dim}\ R/\mathfrak{m}^{\mathbf{ord}(z)}$. □

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 $x_1,\dots,x_n$ 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 $x_1,\dots,x_n$, i.e. $R[x_1,\dots,x_n]$. 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 $(v_1,\dots,v_n) \neq (w_1,\dots,w_n)$. 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 $s_1,\dots,s_n \in Z(S)$ 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 $s_1,\dots,s_n \in Z(S)$ is algebraically independent over R if and only if $\{s_1^{v_1}\cdots s_n^{v_n}\}_{(v_1,\dots,v_n)\in\mathbb{N}^n}$ 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 $R[x_1,\dots,x_n]$ as the set of functions

$$\mathrm{Pol}(R^n,R) := \left\{ f : R^n \to R : \begin{array}{c} f(x_1,\dots,x_n)=\sum_{v\in\mathbb{N}^n} a_v x_1^{v_1}\cdots x_n^{v_n} \text{ for some} \\ \text{finite sequence } \{a_v\}_{v\in\mathbb{N}^n}\subset R \text{ for all } (x_1,\dots,x_n)\in R^n \end{array} \right\},$$

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

$$|\bullet| : \mathbb{N}^n \to \mathbb{N}$$
$$v = (v_1, \ldots, v_n) \mapsto \sum_1^n v_i,$$

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 $v = (v_1, \ldots, v_n)$ and $w = (w_1, \ldots, w_n)$ in $\mathbb{N}^n$,

$$v + w := (v_1 + w_1, \ldots, w_n + v_n).$$

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 $v = (v_1, \ldots, v_n)$ and $w = (w_1, \ldots, w_n)$. Then

$$|v + w| = |(v_1 + w_1, \ldots, v_n + w_n)| = \sum_1^n (v_i + w_i) = \sum_1^n v_i + \sum_1^n w_i = |v| + |w|.$$

□

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 $R[\![x_1,\ldots,x_n]\!]$, 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 $i\in\{1,\ldots,n\}$ 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 $x_1,\ldots,x_n$. 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 $y_1,\ldots,y_n$ 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. □

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). □

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 $\sigma\in\mathrm{Hom}(R,S)$, $\tau\in\mathrm{Hom}(S,T)$. Then $\overline{\tau\circ\sigma}=\overline{\tau}\circ\overline{\sigma}$. We also have that $\overline{id_R}=id_{R[x_1,\ldots,x_n]}$. In other words $(R,\sigma)\mapsto(R[x_1,\ldots,x_n],\overline{\sigma})$ 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 $R[x_1,\ldots,x_n]\overset{\overline{\sigma}}{\simeq}S[x_1,\ldots,x_n]$.*

Proof. An immediate consequence of functoriality. □

Example 2.9.20. It is in general not true that if $R[x_1,\ldots,x_n]\simeq S[x_1,\ldots,x_n]$, 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

$$\mathrm{ev}:S^n\times R[x_1,\ldots,x_n]$$
$$((s_1,\ldots,s_n),f)=\left((s_1,\ldots,s_n),\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right)\mapsto\sum_{v\in\mathbb{N}^n} a_vs_1^{v_1}\cdots s_n^{v_n}$$

Let $s_1,\dots,s_n \in S$. We define *evaluation in* $(s_1,\dots,s_n)\in S^n$ as the map

$$\mathrm{ev}_{s_1,\dots,s_n} : R[x_1,\dots,x_n] \to S$$
$$f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v \mapsto \mathrm{ev}((s_1,\dots,s_n),f) = \sum_{v\in\mathbb{N}^n} a_v s_1^{v_1}\cdots s_n^{v_n}$$

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

$$\mathrm{ev}_{s_1,\dots,s_n}(r) = \mathrm{ev}_{s_1,\dots,s_n}\left(r\mathbf{x}^{(0,\dots,0)}\right) = rs_1^0\cdots s_n^0 = r.$$

□

Remark 2.9.23. Given a commutative R-algebra S and elements $s_1,\dots,s_n \in S$, we note that the R-algebra generated by these elements, i.e. $R[s_1,\dots,s_n]$ is given as the image of $\mathrm{ev}_{s_1,\dots,s_n} : R[x_1,\dots,x_n] \to S$.

Lemma 2.9.24. *Let R be a ring and $J \subset R[y_1,\dots,y_m]$ be an ideal. Consider $\overline{f_1} := f_1 + J,\dots,\overline{f_n} := f_n + J \in R[\mathbf{y}]/J$. Then for each $f \in R[\mathbf{x}]$*

$$\mathrm{ev}_{\overline{f_1},\dots,\overline{f_n}}(f) = \mathrm{ev}_{f_1,\dots,f_n}(f) + J$$

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 $v_1,\dots,v_k \in \mathbb{N}^n$ and $a_{v_i} \in R$. Then

$$\begin{aligned}\mathrm{ev}_{\overline{f_1},\dots,\overline{f_n}}(f) &= \sum_1^k a_{v_i}(f_1+J)^{v_{i1}}\cdots(f_n+J)^{v_{in}} = \sum_1^k a_{v_i}\left(f_1^{v_{i1}}+J\right)\cdots\left(f_n^{v_{in}}+J\right)\\ &= \left[\sum_1^k a_{v_i} f_1^{v_{i1}}\cdots f_n^{v_{in}}\right] + J = \mathrm{ev}_{f_1,\dots,f_n}(f) + J\end{aligned}$$

□

Proposition 2.9.25. *Let S be a commutative R-algebra. Let $\sigma : R[x_1,\dots,x_n] \to S$ be an R-algebra homomorphism. Then $\sigma = \mathrm{ev}_{\sigma(x_1),\dots,\sigma(x_n)}$. Hence any element of $\mathrm{Hom}^{R-alg}(R[x_1,\dots,x_n],S)$ is uniquely determined by it's behavior on the variables of $R[x_1,\dots,x_n]$.*

Proof. Let $f = \sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v \in R[x_1,\dots,x_n]$. Then using the multiplicativity and additivity of σ we have that

$$\sigma(f) = \sigma\left(\sum_{v\in\mathbb{N}^n} a_v\mathbf{x}^v\right) = \sum_{v\in\mathbb{N}^n}\sigma(a_v)\sigma\left(x_1^{v_1}\cdots x_n^{v_n}\right) = \sum_{v\in\mathbb{N}^n} a_v\sigma\left(x_1^{v_1}\right)\cdots\sigma\left(x_n^{v_n}\right) = \mathrm{ev}_{\sigma(x_1),\dots,\sigma(x_n)}(f)$$

□

Corollary 2.9.26. *Let S be a commutative R-algebra. Then*

$$\mathrm{Hom}^{R-Alg}(R[x_1,\dots,x_n],S) = \left\{\mathrm{ev}_{s_1,\dots,s_n} : (s_1,\dots,s_n) \in S^n\right\}.$$

Corollary 2.9.27. *Let $I \subset R[x_1,\dots,x_n]$, $J \subset R[y_1,\dots,y_n]$ be ideals. Then $\mathrm{Hom}^{R-Alg}(R[\mathbf{x}]/I, R[\mathbf{y}]/J)$ is equal to*

$$\left\{f + I \mapsto \mathrm{ev}_{f_1,\dots,f_n}(f) + J : f_1,\dots,f_n \in R[\mathbf{y}], \mathrm{ev}_{f_1,\dots,f_n}(f) = 0 \text{ for every } f \in I\right\} =: F$$

Proof. It is easy to check that $F \subset \mathbf{Hom}^{K-\mathrm{Alg}}(R[\mathbf{x}]/I, R[\mathbf{y}]/J)$. Indeed, consider such a map σ for given $f_1, \ldots, f_n \in R[\mathbf{y}]$. Consider

$$\sigma' : R[\mathbf{x}] \to R[\mathbf{y}]/J, f \mapsto \mathrm{ev}_{f_1+J,\ldots,f_n+J}(f) = \mathrm{ev}_{f_1,\ldots,f_n}(f) + J,$$

is clearly a ring homomorphism satisfying $\sigma'(r) = r$ for $r \in R$, since it is equal to $\pi \circ \mathrm{ev}_{f_1,\ldots,f_n}$, where $\pi : R[\mathbf{y} \to R[y]/J$ is the canonical surjection. Since $\mathrm{ev}_{f_1,\ldots,f_n}(f) = 0$ 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 $\sigma \in \mathbf{Hom}^{R-\mathrm{Alg}}(R[\mathbf{x}]/I, R[\mathbf{y}]/J)$. Then $\sigma \circ \pi \in \mathbf{Hom}^{R-\mathrm{Alg}}(R[\mathbf{x}], R[\mathbf{y}]/J)$, where $\pi : R[\mathbf{x}] \to R[\mathbf{x}]/I$ is the canonical surjection. Hence by the prior corollary, $\sigma \circ \pi = \mathrm{ev}_{f_1+J,\ldots,f_n+J}$ for suitable $f_1 + J, \ldots, f_n + J \in R[\mathbf{y}]/J$. Let $f + I \in R[\mathbf{x}]/I$. Then by Lemma 2.9.24

$$\sigma(f + I) = \sigma \circ \pi(f) = \mathrm{ev}_{f_1+J,\ldots,f_n+J}(f) = \mathrm{ev}_{f_1,\ldots,f_n}(f) + J,$$

hence $\sigma \in F$. □

Lemma 2.9.28. *Let $S \supset R$ be a ring extension and $s_1, \ldots, s_n$ be algebraically independent over R. When $R[x_1, \ldots, x_n]$ denotes the polynomial over R in n variables then $R[s_1, \ldots, s_n] \simeq R[\mathbf{x}]$.*

Proof. $\mathrm{ev}_{s_1,\ldots,s_n} : R[\mathbf{x}] \to R[s_1, \ldots, s_n]$ defines a surjective ring homomorphism. Let $f \in R[\mathbf{x}]$. By the definition of algebraic independence

$$\mathrm{ev}_{s_1,\ldots,s_n}(f) = 0 \iff f = 0.$$

. $\mathrm{ev}_{s_1,\ldots,s_n}$ is therefor injective, implying $R[s_1, \ldots, s_n] \simeq R[\mathbf{x}]$ □

Corollary 2.9.29. *Consider the ring extension $R[x_1, \ldots, x_n, y_1, \ldots, y_m] \supset R$. Then the subring of $R[\mathbf{x}, \mathbf{y}]$ generated by $x_1, \ldots, x_n$ is isomorphic to the polynomial ring in n variables.*

Corollary 2.9.30. *Consider the ring extension $R[x_1, \ldots, x_n, y_1, \ldots, y_m] \supset R$. Then $R[\mathbf{x}, \mathbf{y}] = R[\mathbf{x}][\mathbf{y}]$. Furthermore $R[\mathbf{x}, \mathbf{y}] \simeq R[z_1, \ldots, z_n][w_1, \ldots, w_m]$.*

Definition 2.9.31. Let $f \in R[x_1, \ldots, x_n] \setminus \{0\}$, we define the *degree* of $f = \sum_{v \in \mathbb{N}^n} a_v \mathbf{x}^v$, denoted $\deg f$, as the non-negative integer

$$\max\{|v| : v \in \mathbb{N}^n, a_v \neq 0\}.$$

Remark 2.9.32. For a polynomial $f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v \in K[x_1,\dots,x_n]$ with $d := \deg\ f$ we may write

$$f = \sum_{v\in\mathbb{N}^n:|v|\le d} a_v \mathbf{x}^v.$$

Definition 2.9.33. Let $f \in R[x_1,\dots,x_n]\setminus\{0\}$, 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 $R[x_1,\dots,x_n]$ 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 $d,d' \ge 0$, $a_d \ne 0 \ne b_{d'}$. Then

$$fg = \sum_{k=0}^{d+d'} \left(\sum_{0\le i\le d, 0\le j\le d': i+j=k} a_i b_j \right) x^k.$$

Note that for $i \le d$ and $j \le d'$, $i + j = k$ if and only if $i = d$ and $j = d'$, hence $\sum_{0\le i\le d,0\le j\le d':i+j=k} a_i b_j = a_d b_{d'} \ne 0$, using that R is a domain. Since $\{x^i : i \in \mathbb{N}\}$ is linearly independent over R it follows that $fg \ne 0$. Suppose $R[x_1,\dots,x_n]$ is a domain for some $n \ge 1$. Then by the one variable case $R[x_1,\dots,x_{n+1}] \simeq (R[x_1,\dots,x_n])[x_{n+1}]$ is a domain. □

Lemma 2.9.35. *The function*

$$\deg : R[x_1,\dots,x_n]\setminus 0 \to \mathbb{N}$$
$$f \mapsto \deg\ f$$

has the following properties

1. *The degree function is sub-additive for pairs of distinct polynomials, i.e.* $\deg\ f+g \le \max(\deg\ f, \deg\ g)$ *for every* $f,g \in R[\mathbf{x}]\setminus 0$ *with* $f \ne g$.
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.* $\deg\ fg \le \deg\ f + \deg\ g$ *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 $f = \sum_{v\in\mathbb{N}^n:|v|\le d} a_v\mathbf{x}^v$, $g = \sum_{v\in\mathbb{N}^n:|v|\le d'} b_v\mathbf{x}^v$
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

$$\max\{|v| : v \in \mathbb{N}^n, a_v + b_v = 0\} \le \max(d,d')$$

2. From 1. we have that $\deg\ f+g \leq \max(d,d') = d$, 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

$$\deg\ fg = \max\left\{ |u| : \sum_{v,w\in\mathbb{N}^n : v+w=u} a_v b_w \neq 0 \right\} \leq d + d' = \deg\ f + \deg\ g.$$

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 $\deg\ f'g' = d + d'$. 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 $\deg\ r_f < d$ and $\deg\ r_g < d'$, hence by 3.

$$\deg\ f'r_g \leq d + \deg\ r_g < d + d',$$
$$\deg\ g'r_f \leq d' + \deg\ r_f < d + d',$$
$$\deg\ r_f r_g \leq \deg\ r_f + \deg\ r_g < d + d'.$$

We thus get that

$$\deg\ f'r_g + g'r_f + r_f r_g \leq \max(f'r_g, g'r_f, r_f, r_g) < d + d' = \deg\ f'g'.$$

By 2. we get

$$\deg fg = \deg (f'+r_f)(g'+r_g) = \deg f'g' + (f'r_g + g'r_f + r_f r_g) = \deg f'g' = d + d' = \deg f + \deg g.$$

□

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 $h(x_1,\ldots,x_n,f_1,\ldots,f_m) \in I$ and hence is also an element of J. It follows that

$$h + J = h(\mathbf{x}, f_1,\ldots,f_m) + J = 0 + J \Rightarrow\ h \in J.$$

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 $f,g \in R[x_1,\ldots,x_n]$ 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 $f \in R[x_1,\ldots,x_n] \setminus 0$ and $f_1,\ldots,f_n \in R[y_1,\ldots,y_m] \setminus 0$ with $d_i := \deg\ f_i$. Then*

$$\deg\ f(f_1,\ldots,f_n) \leq \deg\ f(x_1^{d_1},\ldots,x_n^{d_n})$$

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

$$\deg\ M_v(x_1^{d_1},\dots,x_n^{d_n}) = \sum_1^n v_i d_i = \sum_1^n \deg\ f_i^{v_i} = \deg f_1^{v_1}\cdots {f_n}^{v_n} = \deg\ M_v(f_1,\dots,f_n).$$

Note that the map $(v_1,\dots,v_n)\mapsto(v_1d_1,\dots,v_nd_n)$ is injective hence,

$$\deg f(f_1,\dots,f_n) \leq \max_{v\in\mathbb{N}^n:a_v\neq 0} \deg M_v(f_1,\dots,f_n) = \max_{v\in\mathbb{N}^n:a_v\neq 0} \deg M_v(x_1^{d_1},\dots,x_n^{d_n}) = \deg f(x_1^{d_1},\dots,x_n^{d_n}).$$

□

Remark 2.9.45. The above result doesn't always hold with equality. take for instance $f = x_1x_2 - x_3$, take $f_1 = y_1$, $f_2 = y_2$ and $f_3 = -y_1y_3$. Then $f(f_1,f_2,f_3) = 0$, while $f(x_1,x_2,x_3^2) = x_1x_2 - x_3^2$.

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 $\{x^i + I : i \in \{0,\dots,d-1\}\}$.*

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 $\{x^i + I : i \in \{0,\dots,d-1\}\}$ 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 $i \in \{1,\dots,d-1\}$. So $\{x^i + I : i \in \{0,\dots,d-1\}\}$ 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 $f \in R[x_1,\dots,x_n]$. Then $f = 0$ if and only if $f(v) = 0$ for every $v \in K^n$.*

Proof. "⇒": This is trivial "⇐":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 $h \in R[x_1,\dots,x_n]\setminus 0$, 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

$$f(v) = f'(v_{n+1}) \neq 0.$$

□

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

$$g = f_{n+1} - \underbrace{\sum_{i=1}^n a_i x^{d_{n+1}-d_i} f_i}_{h}.$$

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

$$\begin{aligned} g &= a(n+1)x^{d_{n+1}} - \sum_{i=1}^{n} a_i x^{d_{n+1}-d_i} \sum_{j=1}^{d_i} a_j^{(i)} x^j + \underbrace{\sum_{i=1}^{d_{n+1}-1} a_i^{(n+1)} x^i}_{r} \\ &= a(n+1)x^{d_{n+1}} - \sum_{i=1}^{n} a_i a(i) x^{d_{n+1}-d_i} x^{d_i} - \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{d_i-1} a_j^{(i)} x^i}_{r'} + r \\ &= a(n+1)x^{d_{n+1}} - \left(\sum_{i=1}^{n} a_i a(i) \right) x^{d_{n+1}} + r' + r = a(n+1)x^{d_{n+1}} - a(n+1)x^{d_{n+1}} + r' + r \\ &= \sum_{i=1}^{n+1} \sum_{j=1}^{d_i-1} a_j^{(i)} x^i. \end{aligned}$$

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 $R(\!(x_1, \ldots, x_n)\!)$ to mean $Q(R[\![x_1, \ldots, x_n]\!])$.

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*

$$\sigma : R \to K[\![x]\!]$$
$$z \mapsto \sum_{i \in \mathbb{N}} \lambda_i x^i$$

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

$$z + w = \sum_0^n (\lambda_i + \mu_i) t^i + (w_n + z_n) \Rightarrow \sigma(z + w) = \sum_{i \in \mathbb{N}} (\lambda_i + \mu_i) x^i$$
$$= \sum_{i \in \mathbb{N}^n} \lambda_i x^i + \sum_{i \in \mathbb{N}^n} \mu_i x^i = \sigma(z) + \sigma(w).$$

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 $\mathbf{ord}(r) \geq n + 1$, 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 $\leq$ 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 $\leq$ is a total order, we can define $s_{\min} := \mathbf{min}_{i \in \{1,\ldots,m\}} \, s_i$. 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,

$$0 \leq v \text{ and } s_{\min} \leq s_j \Rightarrow s_{\min} \leq s_{\min} + v \leq s_j + v = s,$$

hence $s_{\min}$ is a least element of S, implying $\leq$ 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 $x^v \leq x^w$ if $v \leq w$. This total order will have the property that $x^{v_1} \leq x^{v_2} \Rightarrow x^{v_1+v} \leq x^{v_2+v}$ and $1 \leq x^v$. 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 $\leq$ a term order on $\mathbb{N}^n$. We define the *initial term of f with respect to* $\leq$ to be the monomial

$$\mathbf{in}_{\leq} f := \max_{v \in \mathbb{N}^n : a_v \neq 0} a_v \mathbf{x}^v.$$

Lemma 2.9.66. *Let $f, g \in R[x_1, \ldots, x_n] \setminus 0$. Then one finds that*

1. $(\mathbf{in}_{\leq} f + g) \leq \max(\mathbf{in}_{\leq} f, \mathbf{in}_{\leq} g)$
2. *If* $\mathbf{in}_{\leq} f < \mathbf{in}_{\leq} g$, *then* $(\mathbf{in}_{\leq} f + g) = \mathbf{in}_{\leq} g$.

3. *If the leading terms of* f *and* g *are equal then* $\mathrm{in}_{\leq}(f-g) < \mathrm{in}_{\leq} f = \mathrm{in}_{\leq} g$.

4. $\mathrm{in}_{\leq} fg \leq (lm_{\leq} f)(lm_{\leq} g)$.

5. *Suppose* $\boldsymbol{R}$ *is an integral domain. Then* $\mathrm{in}_{\leq} fg = (\mathrm{in}_{\leq} f)(\mathrm{in}_{\leq} g)$.

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 $a_v \mathbf{x}^w = \mathrm{in}_{\leq} f$ and $b_w \mathbf{x}^u = \mathrm{in}_{\leq} g$. 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

$$(\mathrm{in}_{\leq} f + g) = \max_{v \in \mathbb{N}^n : a_v + b_v \neq 0} (a_v + b_v)\mathbf{x}^v \leq \max(\mathrm{in}_{\leq} f, \mathrm{in}_{\leq} g).$$

2. When $\mathrm{in}_{\leq} f < \mathrm{in}_{\leq} g$, $(\mathrm{in}_{\leq} f + g) \leq \max(\mathrm{in}_{\leq} f, \mathrm{in}_{\leq} g)$. Note that $a_u + b_u = b_u \neq 0$, hence $(\mathrm{in}_{\leq} f + g) = \mathrm{in}_{\leq} g$.
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

$$\mathrm{in}_{\leq}(f-g) = \max_{v \in \mathbb{N}^n : a_v - b_v \neq 0} (a_v - b_v)\mathbf{x}^v < \mathrm{in}_{\leq} f.$$

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

$$\mathrm{in}_{\leq} fg = \max_{v \in \mathbb{N}^n : \sum_{v_1, v_2 \in \mathbb{N}^n : v_1 + v_2 = v} a_{v_1} b_{v_2} \neq 0} \left[\sum_{v_1, v_2 \in \mathbb{N}^n : v_1 + v_2 = v} a_{v_1} b_{v_2} \right] \mathbf{x}^v = \mathbf{x}^{w+u} = (\mathrm{in}_{\leq} f)(\mathrm{in}_{\leq} g).$$

If $\boldsymbol{R}$ is an integral domain we get that

$$\mathrm{in}_{\leq} fg = \left[\sum_{v_1, v_2 \in \mathbb{N}^n : v_1 + v_2 = w + u} a_{v_1} b_{v_2} \right] \mathbf{x}^{w+u} = \left(a_w \mathbf{x}^w\right)\left(b_u \mathbf{x}^u\right) = (\mathrm{in}_{\leq} f)(\mathrm{in}_{\leq} g).$$

□

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 $\mathrm{in}_{\leq}\ s^{(j)} = a_v \mathbf{x}^v$ and $\mathrm{in}_{\leq}\ f_i = \beta_w \mathbf{x}^w$, 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 $i \in \{1,\ldots,m\}$ and $f^F := r^{(N)}$. For each $j \geq 0$ for each $i \in \{1,\ldots,m\}$ 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 $\mathrm{in}_{\leq}\ f_i$, and hence f^F is either 0 or not divisible by any $\mathrm{in}_{\leq}\ f_i$. □

2.9.11 Gröbner Bases and Buchbergers Algorithm

For the exploration of Gröbner bases we fix a field K.

Definition 2.9.68. Let $I \subset K[x_1,\ldots,x_n]$ 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 $G = \{f_1,\ldots,f_m\} \subset K[\mathbf{x}] \setminus 0$ is called a Gröbner basis with respect to $\leq$ if it is a Gröbner basis for $\langle f_1,\ldots,f_m\rangle$ with respect to $\leq$.

Proposition 2.9.69. *Let $G = \{f_1, \ldots, f_m\} \subset I$ be a Gröbner basis for an ideal $I \subset K[x_1, \ldots, x_n]$ 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 $\lambda_1, \ldots, \lambda_m \in K[x_1, \ldots, x_n]$ 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 $\lambda_1, \ldots, \lambda_m \in K[\mathbf{x}]$ 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 $i \in \{1, \ldots, m\}$ such that $\mathbf{in}_{\leq} f_i \mid f^G$, 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 $I \subset K[x_1, \ldots, x_n]$ 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 $\lambda_1, \ldots, \lambda_m \in K[\mathbf{x}]$ 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 $I \subset K[x_1, \ldots, x_n]$ be a non-zero ideal and $\leq$ a term order. Then there is a Gröbner basis $G = \{f_1, \ldots, f_m\} \subset I$ for I with respect to $\leq$, hence $I = \langle f_1, \ldots, f_m \rangle$ by the prior corollary.*

Proof. Put $S = \{v \in \mathbb{N}^n : \mathbf{x}^v = \mathbf{in}_{\leq} f \text{ for some } f \in I\}$. Clearly $S \neq \emptyset$, hence by Dickson's lemma we may find $v_1, \ldots, v_m \in S$ such that

$$S \subset \bigcup_1^m (v_i + \mathbb{N}^n)$$

Let $f_i \in I$ be given such that $\mathrm{in}_{\leq} f_i = \mathbf{x}^{v_i}$ for $i \in \{1,\ldots,m\}$ and put $G = \{f_1,\ldots,f_m\} \subset I$. 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 $j \in \{1,\ldots,m\}$ 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 $K[x_1,\ldots,x_n]$ and $G = \{f_1,\ldots,f_m\} \subset K[\mathbf{x}]\setminus 0$ 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 $f_1,\ldots,f_m$ is equal to f^G. Furthermore, the remainder outputted by the division algorithm remains unchanged after a permutation of $f_1,\ldots,f_m$.*

Proof. Let $\lambda_1,\ldots,\lambda_m \in K[\mathbf{x}]$ 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 $\lambda_1',\ldots,\lambda_m' \in K[\mathbf{x}]$ 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 $j \in \{1,\ldots,m\}$ 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 $\lambda_1',\ldots,\lambda_m',(f^G)' \in K[\mathbf{x}]$ be the outcome of the division with respect to $\leq$ of f with $f_{\omega(1)},\ldots,f_{\omega(m)}$. 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* F if there are $\lambda_1,\dots,\lambda_m \in K[\mathbf{x}]$ such that

$$f = \sum_1^m \lambda_i f_i$$

and $\mathrm{in}_\leq\ \lambda_i f_i \leq \mathrm{in}_\leq\ f$ 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 $b_i \mathbf{x}^{v_i} = \mathrm{in}_\leq\ \lambda_i$ and $c_i \mathbf{x}^{w_i} = \mathrm{in}_\leq\ f_i$. set

$$\kappa = \max\ \left\{ v_i + w_i \in \mathbb{N}^n : i \in \{1,\dots,m\}, \lambda_i \neq 0 \right\}.$$

Then $v \leq \kappa$ and the following statements hold

1. *$v = \kappa \iff \mathrm{in}_\leq\ f_i \lambda_i \leq \mathrm{in}_\leq\ f$ for every $i \in \{1,\dots,m\}$ such that $\lambda_i \neq 0$.*
2. *$v = \kappa \Rightarrow \mathrm{in}_\leq\ f_i \mid \mathrm{in}_\leq f$ 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

$$\mathrm{in}_\leq\ f > \max_{i \in \{1,\dots,m\} : \lambda_i \neq 0} \mathrm{in}_\leq\ \lambda_i f_i = \mathrm{in}_\leq \left(\sum_1^m \lambda_i f_i \right) \Rightarrow f \neq \sum_1^m \lambda_i f_i.$$

It thus follows that since we assumed $f = \sum_1^m \lambda_i f_i$, we get the bound $v \leq \kappa$.
1. For every $i \in \{1,\dots,m\}$ such that $\lambda_i f_i \neq 0$ we have that

$$b_i c_i \mathbf{x}^{v_i + w_i} = \mathrm{in}_\leq \lambda_i \leq \mathrm{in}_\leq\ f = a_v \mathbf{x}^v \iff \mathbf{x}^{v_i + w_i} \leq \mathbf{x}^v \iff v_i + w_i \leq v$$

hence

$$\kappa = v \iff \mathrm{in}_\leq\ \lambda_i f_i \leq .$$

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

$$\mathrm{in}_\leq f = a_v \mathbf{x}^v = \left[\sum_1^l b_i c_i \right] \mathbf{x}^{v_1} \mathbf{x}^{w_1} = \left(\left[\sum_1^l b_i \frac{c_i}{c_1} \right] \mathbf{x}^{v_1} \right) c_1 \mathbf{x}^{w_1} = \left(\left[\sum_1^l b_i \frac{c_i}{c_1} \right] \mathbf{x}^{v_1} \right) \mathrm{in}_\leq f_i \Rightarrow \mathrm{in}_\leq f_i \mid \mathrm{in}_\leq f.$$

for some $l \leq m$.

□

Lemma 2.9.75. *Let $F = \{f_1, \ldots, f_m\} \subset K[x_1, \ldots, x_n] \setminus 0$ 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 $\lambda_1, \ldots, \lambda_m \in K[\mathbf{x}]$ such that

$$f = \sum_1^m \lambda_i f_i,$$

and $\mathrm{in}_\leq \lambda_i f_i \leq \mathrm{in}_\leq f$ for every $i \in \{1, \ldots, m\}$ with $\lambda_i \neq 0$. Then by the above lemma

$$\kappa := \max \{\mathrm{in}_\leq \lambda_i f_i : i \in \{1, \ldots, m\}, \lambda_i \neq 0\} = \mathrm{in}_\leq f$$

and by the same lemma we then have that $\mathrm{in}_\leq f_i \mid \mathrm{in}_\leq f$ for some $i \in \{1, \ldots, m\}$, hence F is a Gröbner basis.
2. "$\Rightarrow$": If $f^F = 0$ then there are $\lambda_1, \ldots, \lambda_i \in K[\mathbf{x}]$ such that

$$f = \sum_1^m \lambda_i f_i$$

and $\mathrm{in}_\leq \lambda_i f_i \leq \mathrm{in}_\leq f$ for $i \in \{1, \ldots, m\}$ 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 $f, g \in K[x_1, \ldots, x_n] \setminus 0$. Pick $w \in \mathbb{N}^n$ such that $\mathbf{x}^w = \mathrm{lcm}(\mathrm{in}_\leq f, \mathrm{in}_\leq g)$. We define *the S-polynomial* or *the syzygy* of f and g to be

$$S(f, g) := \frac{\mathbf{x}^w}{\mathrm{in}_\leq f} f - \frac{\mathbf{x}^w}{\mathrm{in}_\leq g} g.$$

Remark 2.9.77. Recall that $w = \left(\max\left(w_1^f, w_1^g\right), \ldots, \max\left(w_n^f, w_n^g\right)\right)$, where $a_{w^f} \mathbf{x}^{w^f} = \mathrm{in}_\leq f$ and $b_{w^g} \mathbf{x}^{w^g} = \mathrm{in}_\leq g$.

Note this simple fact about S-polynomials

Lemma 2.9.78. *Let $f, g \in K[x_1, \ldots, x_n] \setminus 0$. Pick $w \in \mathbb{N}^n$ such that $\mathbf{x}^w = \mathrm{lcm}(\mathrm{in}_\leq f, \mathrm{in}_\leq g)$. Then $\mathrm{in}_\leq S(f, g) < \mathbf{x}^w$. In other words, the initial term of $\frac{\mathbf{x}^w}{\mathrm{in}_\leq f} f$ cancels with the initial term of $-\frac{\mathbf{x}^w}{\mathrm{in}_\leq g} g$.*

Proof. Indeed, note that

$$\begin{aligned}\mathrm{in}_{\leq}\ \frac{\mathbf{x}^w}{\mathrm{in}_{\leq}\ f}f-\mathrm{in}_{\leq}\ \frac{\mathbf{x}^w}{\mathrm{in}_{\leq}\ g}g&=\left(\mathrm{in}_{\leq}\ \frac{\mathbf{x}^w}{\mathrm{in}_{\leq}\ f}\right)(\mathrm{in}_{\leq}\ f)-\left(\mathrm{in}_{\leq}\ \frac{\mathbf{x}^w}{\mathrm{in}_{\leq}\ g}\right)(\mathrm{in}_{\leq}\ g)\\&=\frac{\mathbf{x}^w}{\mathrm{in}_{\leq}\ f}\mathrm{in}_{\leq}\ f-\frac{\mathbf{x}^w}{\mathrm{in}_{\leq}\ g}\mathrm{in}_{\leq}\ g=\mathbf{x}^w-\mathbf{x}^w=0\end{aligned}$$

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 $\boldsymbol{F}$.

Lemma 2.9.79. *Let* $F=\{f_1,\ldots,f_m\}\subset K[x_1,\ldots,x_n]\setminus 0$ *and let* $1\leq l\leq m$ *and* $\sum_1^m\lambda_i f_i\in\langle F\rangle$ *with* $b_i\mathbf{x}^{v_i}=\mathrm{in}_{\leq}\lambda_i$, $c_i\mathbf{x}^{w_i}=\mathrm{in}_{\leq}\ f_i$ *be given such that*

$$b_ic_ix^{v_i+w_i}=\mathrm{in}_{\leq}\ \lambda_if_i=\kappa:=\max_{j\in\{1,\ldots,m\}}\ \mathrm{in}_{\leq}\ \lambda_jf_i$$

for every $i\in\{1,\ldots,l\}$ *and* $\sum_1^l b_ic_i\neq 0$. *Define*

$$\mu_{\lambda_1,\ldots,\lambda_m,F}:=\sum_1^m(\mathrm{in}_{\leq}\ \lambda_i)f_i.$$

Then $\mu_{\lambda_1,\ldots,\lambda_m,F}\in I:=\langle S(f_1,f_2),S(f_2,f_3),\ldots,S(f_{l-1},f_l)\rangle$, *and* $\mathrm{in}_{\leq}\ \mu_{\lambda_1,\ldots,\lambda_m,F}<\kappa$.

Proof. Put $g_i:=\mathbf{x}^{v_i}\frac{f_i}{c_i}$ for $i\in\{1,\ldots,l\}$. Then

$$\begin{aligned}\mu_{\lambda_1,\ldots,\lambda_m,F}&=\sum_1^l b_ic_i\left(\mathbf{x}^{v_i+w_i}+\ldots\right)=\sum_1^l b_ic_ig_i\\&=\left[\sum_{j=1}^{l-1}\left[\sum_{i=1}^{j}b_ic_i\right](g_j-g_{j+1})\right]+\underbrace{\left[\sum_1^l b_ic_i\right]}_{=0}g_l.\end{aligned}$$

Put $x^{u_{ij}}:=\mathrm{lcm}(\mathbf{x}^{w_i},\mathbf{x}^{w_j})$ for $i,j\in\{1,\ldots,l\}$ and note that

$$g_i-g_j=\frac{\mathbf{x}^{v_i}}{c_i}f_i-\frac{\mathbf{x}^{v_j}}{c_j}f_j=\frac{\mathbf{x}^{v_i+w_i}}{c_i\mathbf{x}^{w_i}}f_i-\frac{\mathbf{x}^{v_j+w_j}}{c_j\mathbf{x}^{w_j}}f_j\overset{(*)}{=}\mathbf{x}^{\xi_{ij}}\left(\frac{\mathbf{x}^{u_{ij}}}{\mathrm{in}_{\leq}\ f_i}f_i-\frac{\mathbf{x}^{u_{ij}}}{\mathrm{in}_{\leq}\ f_j}f_j\right)=\mathbf{x}^{\xi_{ij}}S(f_i,f_j),$$

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

$$\mu_{\lambda_1,\ldots,\lambda_l,F}=\sum_1^{l-1}\mathbf{x}^{\xi_i}S(f_i,f_{i+1})\in I.$$

Set $u_i = u_{i(i+1)}$. Then additionally we have that

$$\mathrm{in}_{\leq}\ \mu_{\lambda_1,\ldots,\lambda_m,F} = \max_{i\in\{1,\ldots,l-1\}}\ x^{\xi_i}\mathrm{in}_{\leq}\ S(f_i,f_{i+1}) < \max_{i\in\{1,\ldots,l-1\}}\ \mathbf{x}^{\xi_i+u_i} = \max_{i\in\{1,\ldots,l-1\}}\ \mathbf{x}^{v_i+w_i} = \kappa.$$

□

Theorem 2.9.80. *Let $F = \{f_1,\ldots,f_m\} \subset K[x_1,\ldots,x_n]\setminus 0$. If $S(f_i,f_j) \to_F 0$ for every $i,j\in\{1,\ldots,m\}$, 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 $\mu_{\lambda_1,\ldots,\lambda_m,F} \in \langle S(f_1,f_2), S(f_2,f_3),\ldots,S(f_{l-1},f_l)\rangle$ and $\mathrm{in}_{\leq}\ \mu_{\lambda_1,\ldots,\lambda_m} < \kappa$. By assumption there are $\psi_1^{(i)},\ldots,\psi_m^{(i)} \in K[\mathbf{x}]$ 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 $i\in\{1,\ldots,l-1\}$ and $j\in\{1,\ldots,m\}$. This means that

$$\mu_{\lambda_1,\ldots,\lambda_m,F} = \sum_{j=1}^m \underbrace{\left[\sum_{i=1}^{l-1} \mathbf{x}^{\xi_i}\psi_j^{(i)}\right]}_{\chi_j} f_j$$

with

$$\mathrm{in}_{\leq}\ \chi_j f_j = \max_{i\in\{1,\ldots,l-1\}}\ \mathbf{x}^{\xi_i}\left(\mathrm{in}_{\leq}\ \psi_j^{(i)}\right)\left(\mathrm{in}_{\leq}\ f_j\right) \leq \max_{i\in\{1,\ldots,l-1\}}\ \mathbf{x}^{\xi_i}\mathrm{in}_{\leq}\ S(f_i,f_{i+1}) = \mathrm{in}_{\leq}\ \mu_{\lambda_1,\ldots,\lambda_m,F} < \kappa.$$

Now note that

$$f = \mu_{\lambda_1,\ldots,\lambda_m,F} + \sum_1^l (\lambda_i - \mathrm{in}_{\leq}\ \lambda_i) f_i + \sum_{l+1}^m \lambda_i f_i,$$

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

$$\mathbf{in}_{\leq} f = \kappa' := \max_{i \in \{1,\ldots,m\}} \lambda_i' f_i,$$

we have $\mathbf{in}_{\leq} \lambda_i' f_i \leq \mathbf{in}_{\leq} f$ 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 $m(0) \geq 1$ be a generating set for I. For $i \geq 0$ 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 $\{t_i\}_{i \geq 0} \subset K[x_1,\ldots,x_n]$ 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 $N \geq 0$ for every $i \geq N$, $t_j \mid t_i$ 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 $t_i = \mathbf{in}_{\leq} f_i$ 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 $\mathbf{in}_{\leq} f_q$ 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,

$$\mathbf{lcm}(\mathbf{in}_{\leq} f, \mathbf{in}_{\leq} g) = (\mathbf{in}_{\leq} f)(\mathbf{in}_{\leq} g) = \mathbf{in}_{\leq} fg.$$

Put $r = f - \mathbf{in}_{\leq} f$ and $s = g - \mathbf{in}_{\leq} g$. Then $\mathbf{in}_{\leq} r < \mathbf{in}_{\leq} f$ or $\mathbf{in}_{\leq} r = 0$ and $\mathbf{in}_{\leq} s < \mathbf{in}_{\leq} g$ or $\mathbf{in}_{\leq} s = 0$. Then

$$S(f, g) = (\mathbf{in}_{\leq} g) f - (\mathbf{in}_{\leq} f) g = (g - s) f - (f - r) g = rg - sf.$$

Suppose $r = 0$, then $S(f, g) = -sf$, implying $\mathbf{in}_{\leq} f \leq \mathbf{in}_{\leq} S(f, g)$ and hence $S(f, g) \to_{\{f,g\}} 0$. Suppose $\mathbf{in}_{\leq} r < \mathbf{in}_{\leq} f$. Suppose for a contradiction

$$(\mathbf{in}_{\leq} r)(\mathbf{in}_{\leq} g) = (\mathbf{in}_{\leq} s)(\mathbf{in}_{\leq} f).$$

Then $\mathbf{in}_{\leq} f \mid \mathbf{in}_{\leq} r$ (since $\mathbf{in}_{\leq} g \nmid \mathbf{in}_{\leq} f$), but then $\mathbf{in}_{\leq} f \leq \mathbf{in}_{\leq} r$ leading to a contradiction. We then find that

$$\mathbf{in}_{\leq} S(f,g) = \mathbf{in}_{\leq}(rg - sf) = \max(\mathbf{in}_{\leq} rg, \mathbf{in}_{\leq} sf)$$

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. $\mathbf{in}_{\leq} f_i \nmid \mathbf{in}_{\leq} f_j$ for every $i, j \in \{1, \ldots, m\}$ with $i \neq j$.
2. $\mathbf{in}_{\leq} f_i = \mathbf{x}^{v_i}$ 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 $\mathbf{in}_{\leq} f_j$ 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 $\mathbf{in}_{\leq} g \mid \mathbf{in}_{\leq} f$, 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 $\mathbf{in}_{\leq} f_j \mid \mathbf{in}_{\leq} f_i$, 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 $l \in \{1, \ldots, m+1\}$, $\mathbf{in}_{\leq} f_l \mid \mathbf{in}_{\leq} f$. If $l = i$, then $\mathbf{in}_{\leq} f_j \mid \mathbf{in}_{\leq} f$, 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 $0 \neq I \subset K[x_1,\dots,x_n]$ has a unique reduced Gröbner basis.*

Proof. **Uniqueness:** Consider two reduced Gröbner bases $G = \{f_1,\dots,f_m\}$ and $G' = \{f'_1,\dots,f'_{m'}\}$. We first check that the cardinality of these Gröbner bases match. Let $i \in \{1,\dots,m\}$, then for some $\tau(i) \in \{1,\dots,m'\}$, $\mathbf{in}_{\leq} f'_{\tau(i)} \mid \mathbf{in}_{\leq} f_i$. Let $j \in \{1,\dots,m'\}$ then for some $\omega(j) \in \{1,\dots,m\}$, $\mathbf{in}_{\leq} f_{\omega(j)} \ \mathbf{in}_{\leq} f'_j$. Then we have that

$$\mathbf{in}_{\leq} f_{\omega(\tau(i))} \mid \mathbf{in}_{\leq} f'_{\tau(i)} \text{ and } \mathbf{in}_{\leq} f'_{\tau(i)} \mid \mathbf{in}_{\leq} f_i \Rightarrow \mathbf{in}_{\leq} f_{\omega(\tau(i))} \mid \mathbf{in}_{\leq} f_i$$

by minimality of the Gröbner bases $i = \omega(\tau(i))$. A similar argument shows that $\tau(\omega(j))$ for every $j \in \{1,\dots,m'\}$, 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 $\mathbf{in}_{\leq} f'_{\tau(i)} = \mathbf{in}_{\leq} f_i$ for every $i \in \{1,\dots,m\}$ since the coefficient of every initial term is 1. Let $i \in \{1,\dots,m\}$. Since $\mathbf{in}_{\leq} f_i = \mathbf{in}_{\leq} f'_{\tau(i)}$, either $f_i = f_{\tau(i)}$ or $\mathbf{in}_{\leq}\left(f_i - f'_{\tau(i)}\right) < \mathbf{in}_{\leq} f_i$. 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 $\mathbf{in}_{\leq} f_i$. 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, $\mathbf{in}_{\leq} f_j$ does not divide such a term for any $j \in \{1,\dots,m\} \setminus \{i\}$. 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 $G = \{f_1,\dots,f_m\}$ for I. Define $g_i := f_i^{\{g_1,\dots,g_{i-1},f_i,\dots,f_m\}\setminus\{f_i\}}$ for every $i \in \{1,\dots,m\}$. Since $\mathbf{in}_{\leq} f_i$ is divisible by any $\mathbf{in}_{\leq} f_j$ for $j \in \{i+1,\dots,m-1\}$ we see that g_i is of the form $\mathbf{in}_{\leq} f_i + \dots$. Thus if $f \in I$, there is some g_j such that $\mathbf{in}_{\leq} g_j = \mathbf{in}_{\leq} f_j \mid \mathbf{in}_{\leq} f$, meaning that each set $\{g_1,\dots,g_{\{}k-1\},f_k,\dots,f_m\}$ and in particular $G' := \{g_1,\dots,g_m\}$ is a Gröbner basis. The g_i's being residues following the division algorithm by a set of polynomials with initial terms coming from $\{f_1,\dots,f_m\} \setminus \{f_i\}$ implies that no term in g_i is divisible by any $\mathbf{in}_{\leq} f_j$ 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 $I \subset K[x_1,\dots,x_n]$ with respect to the lexicographic term order with $x_1 < \dots < x_n$. Then $G \cap K[x_1,\dots,x_i] \subset K[x_1,\dots,x_i]$ is a Gröbner basis for the ideal $I \cap K[x_1,\dots,x_i] \subset K[x_1,\dots,x_i]$ with respect to the lexicographic term order with $x_1 < \dots < x_i$ for every $i \in \{1,\dots,i\}$.*

Proof. Let $G' = G \cap K[x_1,\ldots,x_i]$. Let $f \in I' = I \cap K[x_1,\ldots,i]$. For some $g \in G$, $\mathbf{in}_{\leq}\ g \mid \mathbf{in}_{\leq}\ f$. Let $t = a\mathbf{x}^v$ be a term of g and write $b\mathbf{x}^w = \mathbf{in}_{\leq}\ f$. 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 $j \in \{i+1,\ldots,m\}$, then $w <_{\text{lex}} u$. Hence we conclude that $v_j = 0$ for every $j \in \{i+1,\ldots,m\}$, implying $t \in K[x_1,\ldots,x_i]$ and ultimately that $g \in K[x_1,\ldots,x_i]$. □

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 $p \in R[x_1,\ldots,x_n]$ 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 $s := \max\ \{i \in \{1,\ldots,k\} : p \nmid a_i\}$ and $t := \max\ \{i \in \{1,\ldots,h\} : o \nmid b_i\}$. 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

$$p \nmid \sum_{i=1,\ldots,h,j=1,\ldots,ki+j,=s+t} a_i b_j \Rightarrow p \nmid fg.$$

□

Definition 2.9.93. A polynomial $f = \sum_0^n a_i x^i \in R[x]$ is said to be *primitive* if $\mathbf{gcd}(a_0,\ldots,a_n) = 1$.

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,

$$\begin{aligned} x_{n+1}^d f\left(\frac{x_1}{x_{n+1}},\dots,\frac{x_n}{x_{n+1}}\right) &= x_{n+1}^d \sum_{v\in\mathbb{N}^n} a_v \frac{x_1^{v_1}}{x_{n+1}^{v_1}} \cdots \frac{x_n^{v_n}}{x_{n+1}^{v_n}} \\ &= x_{n+1}^d \sum_0^d \sum_{v\in\mathbb{N}^n:|v|=i} a_v \frac{x_1^{v_1}\cdots x_n^{v_n}}{x_{n+1}^{v_1+\cdots+v_n}} \\ &= \sum_0^d \sum_{v\in\mathbb{N}^n:|v|=i} x_{n+1}^{d-i} a_v \mathbf{x}^v = \sum_0^d x_{n+1}^{d-i} f_i. \end{aligned}$$

Note also $x_{n+1}^{d-i} f_i$ is homogeneous of degree $d - i + i = d$, hence $f^* = x_{n+1}^d f_0 + x_{n+1}^{d-1} f_1 + \cdots + f_d$ 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

$$\begin{aligned} x_{n+1}^{\deg\ f+\deg\ g-\deg(f+g)}(f+g)^* &= x_{n+1}^{\deg\ f+\deg\ g-\deg(f+g)+\deg(f+g)}\left[f\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right)+g\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right)\right] \\ &= x_{n+1}^{\deg\ g}\left[x_{n+1}^{\deg\ f} f\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right)\right]+x_{n+1}^{\deg\ f}\left[x_{n+1}^{\deg\ g} g\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right)\right] \\ &= x_{n+1}^{\deg\ g} f^* + x_{n+1}^{\deg f} g^*. \end{aligned}$$

For the second see that

$$\begin{aligned} (fg)^* &= x_{n+1}^{\deg\ f+\deg\ g} f\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right) g\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right) \\ &= x_{n+1}^{\deg\ f} f\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right) x_{n+1}^{\deg\ g} g\left(\frac{x_1}{x_{n+1}},\ldots,\frac{x_n}{x_{n+1}}\right) = f^* g^*. \end{aligned}$$

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 $Q \in R[x_1,\ldots,x_{n+1}]$. 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

$$d' = \deg\ F(x_1,\ldots,x_n,1) = \max_{v\in\mathbb{N}^{n+1}:a_v\neq 0} \sum_1^n v_i = \max_{v\in\mathbb{N}^{n+1}:a_v\neq 0} d - v_{n+1}.$$

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* $F \in R[x_1,\ldots,x_{n+1}]$ *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 $F = y^r (F_*)^* = a y^r \prod_1^l (x - a_i y)^{r_i}$. □

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

$$\sum_{0}^{d+1} \binom{k+n}{n} = \binom{m+n+1}{n+1} + \binom{m+n+1}{n} = \binom{(m+1)+n+1}{n+1}.$$

□

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

$$\#\Delta_{n+1,d+1} = \sum_{0}^{d+1} \#\Delta_j = \sum_{0}^{d+1} \binom{j+n-1}{n-1} = \binom{d+n+1}{n} = \binom{(d+1)+(n+1)-1}{(n+1)-1}.$$

□

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

$$\dim_K V_K(d,n) = \#\Delta_{n,d} = \binom{d+n-1}{n-1}$$

□

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_ij 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 $R[x_1,\ldots,x_n]$, 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 $I,J\subset R[x_1,\ldots,x_n]$ 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 $I\subset R[x_1,\ldots,x_n]$ 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 $I\subset R[x_1,\ldots,x_n]$ be a homogeneous polynomial. Let $\alpha\in R[\mathbf{x}]/I$. Then for some unique $d\geq 0$, there are unique $\alpha_i\in R[\mathbf{x}]/I$, $i\in\{0,\ldots,d\}$ 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* $(d_1,\dots,d_m) \in \mathbb{Z}_{\geq 0}^m$, 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 $\mathbf{char}\ R = p > 0$, 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 $\mathbf{char}\, R[x_1,\ldots,x_n] = p$, $R[x_1,\ldots,x_{n+1}] \simeq R[x_1,\ldots,x_n][x_{n+1}]$ 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 $S = R[s_1, \ldots, s_n]$ for suitable $s_1, \ldots, s_n \in S$.

Proposition 2.10.3. *Let $S \supset R$ be a ring-finite ring extension. Then*

$$S \simeq R[x_1, \ldots, x_n]/I,$$

for some $n \geq 1$ and some ideal $I \subset R[x_1, \ldots, x_n]$.

Proof. For suitable $s_1, \ldots, s_n \in S$, $S = R[s_1, \ldots, s_n]$, hence $\mathrm{ev}_{s_1,\ldots,s_n} : R[x_1, \ldots, x_n] \to S$ is a surjective R-algebra homomorphism. Then by the first isomorphism theorem $S \simeq K[\mathbf{x}]/\ker\ \mathrm{ev}_{s_1,\ldots,s_n}$. □

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 $s_1, \ldots, s_n \in S$, $S = \sum_1^n R s_i$. We prove that $S = R[s_1, \ldots, s_n]$. We already know that $S \supset R[s_1, \ldots, s_n]$. Let $s \in S$. Then $s = \sum_1^n r_i s_i$ for suitable $r_1, \ldots, r_n \in R$, hence $s \in R[s_1, \ldots, s_n]$. □

Example 2.10.5. The converse implication of the above proposition is clearly not true. Consider for instance $R[x_1, \ldots, x_n] \supset R$. given $f_1, \ldots, f_m \in R[\mathbf{x}]$. If these are all 0, clearly $R[\mathbf{x}] \supsetneq \sum_1^m R f_i$. Otherwise putting $D = \max_{i \in \{1,\ldots,m\}} \{\deg\ f_i\}$, we see that for $r_1, \ldots, r_m \in R$,

$$\deg \sum_1^m r_i f_i \leq D < D + 1 = \deg\ x_1^{D+1},$$

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 $s_1, \ldots, s_n \in S$, $R[s_1, \ldots, s_n]$ is also an integral domain. We denote the fraction field of $R[s_1, \ldots, s_n]$ by $R(s_1, \ldots, s_n)$

Definition 2.10.7. A field extension $L \supset K$ is said to be *finite*, if there exist $a_1, \ldots, a_n$ such that $L = K(a_1, \ldots, a_n)$.

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

$$t = \sum_1^m \sum_1^n b_{ij}t_i s_j \in \sum_1^n \sum_1^m Rt_i s_j.$$

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 $s^n \in \sum_1^{n-1} R s^i$. By a simple induction argument it follows that $s^{n+j} \in \sum_1^{n-1} R s^i$ for every $j \geq 0$. Let $\sum_1^m b_j s^j \in R[s]$. Then by the considerations prior to this, $\sum_1^m b_j s^j \in \sum_1^{n-1} R s^i$, hence $R[s] = \sum_1^{n-1} R s^i$.
"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

$$s t_i = \sum_1^n a_{ij} t_i$$

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*

$$\{s \in S : s \text{ is integral over } R\}$$

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 $s_1,\dots,s_n \in S$ such that $S = R[s_1,\dots,s_n]$. 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 $S = R[s_1,\dots,s_n] \supset R$ 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 $a_0,\dots,a_{n-1} \in K[x]$ 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 $K(x) \supsetneq R := K[z_1,\dots,z_m]$ for any $z_1,\dots,z_m \in K(x)$. 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 $K[a_1] \overset{\sigma}{\cong} K[x]$ 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 $\sigma(a_1)^N \sigma(z)$ 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, $\{1+I, x+I, \ldots, x^{d-1}+I\}$ 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 $\{1, a, \ldots, a^{d-1}\}$ is a basis of $K[a]$ □

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* $a_1, \ldots, a_n \in L$ *that are algebraic over* K*.*

Proof. "⇒": By Lemma 2.10.21 $L = K[a_1, \ldots, a_n]$ for suitable $a_1, \ldots, a_n \in L$ that are algebraic over K, hence $L \subset K[a_1, \ldots, a_n] \subset K(a_1, \ldots, a_n) \subset L$, hence $L = K(a_1, \ldots, a_n)$. "⇐": Suppose $L = K(a_1, \ldots, a_n)$ 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 $L' = K(a_1, \ldots, a_n)$ 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. □

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

$$z = \frac{g(a)}{h(a)} = \frac{\sigma(g)}{\sigma(h)} = \frac{\overline{\sigma}(g+I)}{\overline{\sigma}(h+I)} = \overline{\sigma}\left((g+I)(h+I)^{-1}\right) \in \mathrm{im}\ \overline{\sigma} = \mathrm{im}\ \sigma.$$

□

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 $\deg\ f = 1$ 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 $\gcd(f,h) = x-a \in K'[x]$, 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. □

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. "⇒": 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 $K(a_1,\dots,a_d)$.
"⇐": 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 $g := cf \in K[a_1,\ldots,a_d][y] \setminus 0$, where c is a common denominator of the b_i's, hence

$$g = \sum_0^m g_i(a_1,\ldots,a_d)y^i,$$

for suitable $g_i \in K[x_1,\ldots,x_d]$, with $g_m = c \neq 0$. It follows that

$$h := \sum_0^m g_i y^i \in K[x_1,\ldots,x_d,y],$$

such that $h(a_1,\ldots,a_d,a) = g(a) = 0$, 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 $\mathbf{trdeg}_K\ L \geq \mathbf{trdeg}_K\ L'$.*

Proof. Let $\alpha \in L'$ and set $n := \mathbf{trdeg}\ L$. Pick $\alpha_1,\ldots,\alpha_n \in L'$ and pick $\beta,\beta_1,\ldots,\beta_n \in L$ such that $\sigma(\beta),\sigma(\beta_i) = \alpha$. Pick $f \in K(\beta_1,\ldots,\beta_n)[x] \setminus 0$ 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 $K(\beta_1,\ldots,\beta_n) \to K(\alpha_1,\ldots,\alpha_n)$. 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 $K(\alpha_1,\ldots,\alpha_n)$. So every sequence of n elements in L' are algebraically dependent over K by the prior lemma. It follows that $\mathbf{trdeg}_K\ L' \leq n = \mathbf{trdeg}_K\ L$. □

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 $a_1,\ldots,a_d \in L$.*

1. *There is a subset of $X := \{a_1,\ldots,a_r\}$ that is a transcendence basis for L over K.*

2. *Let $Y = \{a \in X : a$ is transcendental over $K(b_1,\ldots,b_s)\}$. If $b_1,\ldots,b_s \in L$ are algebraically independent over K, then for some $Z \subset Y$, $\{b_1,\ldots,b_s\} \cup Y$ 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 $\{a_1,\ldots,a_r\}$ whose elements are algebraically

independent over K. Let $Y \subset \{a_1,\dots,a_r\}$ be a maximal subset of elements that are algebraically independent over K. After a permutation, we can write $Y = \{a_1,\dots,a_k\}$ for some $k < r$. Then $a_1,\dots,a_k,a_i$ are algebraically dependent over K for each $i \in \{k+1,\dots,r\}$. Then by Lemma 2.10.42, a_i is algebraic over $K(a_1,\dots a_k)$ for each $i \in \{k+1,\dots,r\}$. Then L is algebraic over $K(a_1,\dots,a_k)$, implying Y is a transcendence basis of L over K.

2. We proceed by induction in k: if a_i is algebraic over $K(b_1,\dots,b_s)$ for every $i \in \{1,\dots,r\}$, then L is algebraic over $K(b_1,\dots,b_s)$, hence $\{b_1,\dots,b_s\}$ is a transcendence basis for L over K by Lemma 2.10.42.

Suppose the statement is true for some $k < n$. Consider WLOG $Y = \{a_1,\dots,a_{k+1}\}$. By Lemma 2.10.42 $b_1,\dots,b_s,a_1$ are algebraically independent over K. Every element in $X \setminus Y$ is algebraic over $K(b_1,\dots,b_s,a_1)$, hence $Y' := \{a \in X : a \text{ is transcendental over } K\} \subset Y$. It thus follows by induction that for some $Z' \subset Y'$,

$$\{b_1,\dots,b_s\} \cup \underbrace{\{a_1\} \cup Z'}_{=:Z}$$

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 $X = \{a_1,\dots,a_d\}$ if and only if* $\operatorname{trdeg}_K L = d$

Proof. "⇒": **To prove this we make the following claim**: L is algebraic over $K(b_1,\dots,b_k,a'_{k+1},\dots,a'_d)$ for any $k \in \{0,\dots,d\}$ for some subset $\{a'_{k+1},\dots,a'_d\} \subset \{a_1,\dots,a_d\}$ with $d-k$ elements. We prove this using induction in r. For $k = 0$, we have that $L \supset K(a_1,\dots,a_d)$ is algebraic by the assumption that $\{a_1,\dots,a_d\}$ is a transcendence basis for L over K. Suppose that we the statement holds for some $k \in \{0,\dots,d-1\}$. Then by induction hypothesis L is algebraic over $M := K(b_1,\dots,b_k,a'_{r+1},\dots,a'_d)$ for a suitable subset $\{a'_{k+1},\dots,a'_d\} \subset \{a_1,\dots,a_d\}$. This means b_{k+1} is algebraic over M, hence by Lemma 2.10.42 $b_1,\dots,b_{k+1},a'_{r+1},\dots,a_d$ are algebraically dependent over K. By Lemma 2.10.45 there is an integer $r \le s < d$ and a subset $\{a''_{k+1},\dots,a''_{s+1}\} \subset \{a'_{k+1},\dots,a'_d\}$ such that $b_1,\dots,b_{k+1},a''_{k+2},\dots,a''_{s+1}$ are algebraically independent over K and $b_1,\dots,b_{r+1},a''_{k+1},\dots,a''_{s+1}$ are algebraically dependent over K. Again by Lemma 2.10.42, we have that a''_{k+1} is algebraic over $M' := K(b_1,\dots,b_{k+1},a''_{k+2},\dots,a''_d)$. 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 $b_1,\dots,b_{d+1} \in L$ that are algebraically independent over K. Then $b_1,\dots,b_d$ are algebraically independent over K. But then by the claim $L \supset K(b_1,\dots,b_d)$ is algebraic, hence b_{d+1} is

algebraic over $K(b_1,\dots,b_d)$, but then $b_1,\dots,b_{d+1}$ are algebraically dependent over K.
" $\Leftarrow$": There are algebraically independent $a_1,\dots,a_d \in L$ and for every $a \in L$ $a,a_1,\dots,a_d$ are algebraically dependent, hence by Lemma 2.10.42 $\{a_1,\dots,a_d\}$ is a transcendence basis of L over K. □

Remark 2.10.47. Note the claim of " $\Leftarrow$", simply proves that every algebraically independent $b_1,\dots,b_d$ forms a transcendence basis of L over K

Corollary 2.10.48. *Two transcendence bases have the same cardinality.*

Example 2.10.49. Consider $L = K(x_1,\dots,x_n) = Q(K[x_1,\dots,x_n])$. The elements $x_1,\dots,x_n \in K[x_1,\dots,x_n]$ are algebraically independent over K, hence $x_1,\dots,x_n \in K(x_1,\dots,x_n)$ are algebraically independent over K. It follows that $\{x_1,\dots,x_n\}$ 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 $X = \{a_1,\dots,a_n\} \subset L$ 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 $K(a_1,\dots,a_n)$. Note that $M \supset L$ being module-finite, implies $M \supset L$ is algebraic (cf. Lemma 2.10.21), hence $M \supset K(a_1,\dots,a_n)$ is algebraic. □

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 $a_1,\dots,a_n \in L$ are algebraically independent over K, then so are $\sigma(a_1),\dots,\sigma(a_n)$. It follows that $\mathbf{trdeg}_K\ L \leq \mathbf{trdeg}_K\ M$.*

Proof. Let $a_1,\dots,a_n \in L$ be algebraically independent over K. Let $f \in K[x_1,\dots,x_n] \setminus 0$. Note that $\sigma(k) = k$ for every $k \in K$ and $f(a_1,\dots,a_n) \neq 0$ by the assumption that $a_1,\dots,a_n$ are algebraically independent over K. It follows that

$$f(\sigma(a_1),\dots,\sigma(a_n)) = \sigma(f(a_1,\dots,a_n)) \neq 0,$$

hence $\sigma(a_1),\dots,\sigma(a_n)$ are algebraically independent. Hence for any $n \leq \mathbf{trdeg}\ L$ there are algebraically independent $b_1,\dots,b_n \in M$, hence $\mathbf{trdeg}\ L \leq \mathbf{trdeg}\ M$. □

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 $d_i e_i \in \mathcal{S}$ 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

$$B := \{g_1^{q_1}\cdots g_n^{q_n}x_1^{r_1}\cdots x_n^{r_n} : 0 \le r_i < d_i, \sum_1^n q_i d_i + r_i \le N\}$$

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 $\boldsymbol{B}=\{\Lambda_v:v\in\Delta\}$. 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 $(v,i)\neq(d_1e_1,1)\dots,(d_ne_n,n)$ 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)

$$\sigma(c_{vw})=\begin{cases}1 & \text{if } w=v\\ 0 & \text{if } w\neq v\end{cases}$$

Let $\boldsymbol{D}$ 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,

$$\sigma(\det\ {}_{\mathcal{V}}T_B)=\det\ \sigma({}_{\mathcal{V}}T_B)=\det\ I_D=1\neq 0\Rightarrow\det\ {}_{\mathcal{V}}T_B\neq 0.$$

This means ${}_{\mathcal{V}}T_B$ is invertible in $M_D(L)$, hence $\boldsymbol{B}$ 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

$$f = \sum_{v\in\Delta} a_v \Lambda_v = \sum_{(r_1,\dots,r_n)\in\mathbb{N}^n : r_i<d_i} \left[\sum_{(q_1,\dots,q_n):\sum_1^n (q_i d_i + r_i)\le N} a_{(q_1d_1+r_1,\dots,q_nd_n+r_n)} g_1^{q_1}\cdots g_n^{q_n} \right] x_1^{r_1}\cdots x_n^{r_n}.$$

Setting

$$f_{r_1,\dots,r_n} := \sum_{(q_1,\dots,q_n)\in\mathbb{N}^n : \sum_1^n (q_i d_i + r_i)\le N} a_{(q_1d_1+r_1,\dots,q_nd_n+r_n)} z_1^{q_1}\cdots z_n^{q_n} \quad (0 \le r_i < d_i)$$

These polynomials will satisfy property 1. Secondly,

$$\begin{aligned}\deg\ f_{r_1,\dots,r_n}(z_1^{d_1},\dots,z_n^{d_n}) + \sum_1^n r_i &\le \max_{(q_1,\dots,q_n):\sum_1^n (q_i d_i + r_i)} \deg\ z_1^{q_1d_1}\cdots z_n^{q_nd_n} + \sum_1^n r_i \\ &= \max_{(q_1,\dots,q_n):\sum_1^n (q_i d_i + r_i)} \sum_1^n (q_i d_i + r_i) \le N = \deg\ f.\end{aligned}$$

Lemma 2.10.64. *Let K be any field and $d_1,\dots,d_n > 0$, $\mathcal{S} \subset \mathbb{N}^n$ containing $d_i e_i \in \mathcal{S}$ 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. $\deg\ P_{ij}(z_1^{d_1},\dots,z_n^{d_n}) + \deg M_i \leq \deg\ g_{n+1}M_i = d_{n+1} + \deg\ M_i.$

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

$$\deg\ P_{ij}(z_1^{d_1},\dots,z_n^{d_n}) - \delta_{ij}z_{n+1}^{d_{n+1}} \leq d_{n+1} + \deg\ M_i - \deg\ M_j.$$

For an arbitrary permutation $\pi \in S_d$,

$$\begin{aligned}\deg\ \prod_1^d(P_{i\pi(i)}(z_1^{d_1},\dots,z_d^{d_n}) - \delta_{i\pi(i)}z_{n+1}^{d_{n+1}}) &\leq \sum_1^d(d_{n+1} + \deg\ M_{\pi(i)} - \deg\ M_i)\\ &= \sum_1^d d_{n+1} + \sum_1^d \deg\ M_i - \sum_1^d \deg\ M_i\\ &= dd_{n+1} = \prod_1^{n+1} d_i.\end{aligned}$$

□

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. □

Theorem 2.10.66. *(Perron's Theorem) Let K be any field and let $f_1,\dots,f_{n+1} \in K[x_1,\dots,x_n]$ and put $d_i := \deg\ f_i$ for each i. Then there is a $P \in K[z_1,\dots,z_{n+1}] \setminus 0$ satisfying*

1. $P(f_1,\dots,f_{n+1}) = 0$,
2. $\deg\ P(z_1^{d_1},\dots,z_{n+1}^{d_{n+1}}) \leq \prod_1^{n+1} d_i$.

Proof. First a slight reformulation. Let M be some field and consider $H = \{h_1,\dots,h_{n+1}\} \subset M[x_1,\dots,x_n]$, $\delta_i := \deg\ h_i$. Set

$$\Delta_{\delta_1,\dots,\delta_{n+1}} := \left\{ v \in \mathbb{N}^{n+1} : \sum_1^{n+1} v_i\delta_i \leq \prod_1^{n+1} \delta_i \right\}.$$

Then define

$$B(H) := \left\{ h_1^{q_1}\cdots h_{n+1}^{q_{n+1}} : (q_1,\dots,q_{n+1}) \in \Delta_{\delta_1,\dots,\delta_{n+1}} \right\}.$$

And let $\mathcal{V}$ be the standard basis of $\{\mathbf{z}^v : v \in \Delta_{\delta_1,\dots,\delta_n}$. If ${}_{\mathcal{V}}T_{B(H)}$ is not invertible if and only if for suitable $a_v \in M$

$$\sum_{v\in\Delta} a_v h_1^{v_1}\cdots h_{n+1}^{v_{n+1}} = \mathrm{ev}_{h_1,\dots,h_{n+1}}\left(\underbrace{\sum_{v\in\Delta} a_v \mathbf{z}^v}_{P_H} \right),$$

where

$$\deg\ P_H(z_1^{\delta_1},\dots,z_{n+1}^{\delta_{n+1}}) \leq \max_{v\in\Delta} \deg\ \mathbf{z}^{(\delta_1 v_1,\dots,\delta_{n+1}v_{n+1})} = \max_{v\in\Delta}\ \sum_1^{n+1} \delta_i v_i \leq \prod_1^{n+1} \delta_i.$$

Set $F = \{f_1,\dots,f_{n+1}\}$ and $B := \{\mathbf{z}^v : v \in \Delta_{d_1,\dots,d_{n+1}}\}$. To prove the theorem we can equivalently prove that $\det\ {}_BT_{B(F)} = 0$. 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. $Q(g_1,\dots,g_n,f_{n+1}) = 0$,
2. $\deg\ Q(z_1^{d_1},\dots,z_{n+1}^{d_{n+1}}) \leq \prod_1^{n+1} d_i$.

Set $G := \{g_1,\dots,g_n,f_{n+1}\}$. By small easy lemma $\sigma({}_BT_{B(G)}) = {}_BT_{B(F)}$. It thus follows that

$$\det\ {}_BT_{B(F)} = \det\ \sigma({}_BT_{B(G)}) = \sigma(\det\ {}_BT_{B(G)}) = \sigma(0) = 0.$$

□

2.10.9 Noether Normalizations

Lemma 2.10.67. *Let K be a field and let $f = \sum_{v\in\mathbb{N}^n} a_v \mathbf{x}^v \in K[x_1,\ldots,x_n]$ be given with $d := \deg\ f > 0$. Then we have the following*

1. *There are elements $y_1,\ldots,y_{n-1} \in K[x_1,\ldots,x_n]$ such that $x_i = y_i + x_n^{r_i}$ for $i \in \{1,\ldots,n-1\}$ 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 $G_i \in K[y_1,\ldots,y_{n-1}]$.

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 $i \in \{1,\ldots,n-1\}$. 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

$$f = f\left(y_1 + x_n^{r_1},\ldots,y_{n-1} + x_n^{r_{n-1}}, x_n\right) = \sum_{v\in\mathbb{N}^n:|v|\le d} a_v \left(\prod_1^{n-1}\left(y_i + x_n^{r_i}\right)^{v_i}\right) x_n^{v_n}$$

$$= \sum_{v\in\mathbb{N}^n:|v|\le d} a_v \left(x_n^{v_n+\sum_1^{n-1} v_i r_i} + \underbrace{\cdots}_{\text{lower degree terms}}\right) = \sum_{v\in\mathbb{N}^n:|v|\le d} a_v x_n^{v_n+\sum_1^{n-1} v_i r_i} + \underbrace{\cdots}_{\text{lower degree terms}}$$

$$= a_{v_m} x_n^m + \underbrace{\cdots}_{\text{lower degree terms}}.$$

We can write the lower order terms on the form $\sum_1^{m-1} G_i x_n^i$ for suitable $G_i \in K[y_1,\ldots,y_{n-1}]$.
2. Write $f = \sum_0^d f_i$ with $f_d \neq 0$ for homogeneous $f_i \in K[x_1,\ldots,x_n]$ of degree i. By Lemma 2.9.136, $f_d(x_1,\ldots,x_{n-1},1) \neq 0$. Since $\#K = \infty$ we get that there are $a_1,\ldots,a_{n-1} \in K$ such that $f_d(a_1,\ldots,a_{n-1},1) \neq 0$. We now pick $y_i = x_i - a_i x_i$ for $i \in \{1,\ldots,n-1\}$. Note that

$$\sum_{v\in\mathbb{N}^n:|v|=d} a_v (a_1 x_n)^{v_1}\cdots(a_{n-1}x_n)^{v_{n-1}} x_n^{v_n} = \left[\sum_{v\in\mathbb{N}^n:|v|=d} a_v a_1^{v_1}\cdots a_{n-1}^{v_{n-1}}\cdot 1^{v_n}\right] x_n^d = f_d(a_1,\ldots,a_{n-1},1)x_n^d.$$

From this it follows that

$$\begin{aligned} f &= f(y_1 + a_1x_n, \ldots, y_{n-1} + a_{n-1}x_n, x_n) \\ &= \sum_{v \in \mathbb{N}^n : |v| = d} a_v (y_1 + a_1 x_n)^{v_1} \cdots (y_{n-1} + a_{n-1}x_n)^{v_{n-1}} x_n^{v_n} + \sum_1^{m-1} \sigma(F_i) \\ &\overset{(*)}{=} \left[\sum_{v \in \mathbb{N}^n : |v| = d} a_v a_1^{v_1} \cdots a_{n-1}^{v_{n-1}} \cdot 1^{v_n} \right] x_n^d + \cdots = f_d(a_1, \ldots, a_{n-1}, 1) x_n^d + \ldots. \end{aligned}$$

We then set $a = f_d(a_1, \ldots, a_{n-1}, 1)$. The $\ldots$ 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 $G_i \in K[y_1, \ldots, y_{n-1}]$. □

Theorem 2.10.68. *(Noether Normalization Theorem) Let $A = K[x_1, \ldots, x_n]/J$ for some field K and some ideal $J \subsetneq K[x_1, \ldots, x_n]$. Let also $I \subsetneq A$ be an ideal.*

(a) Then there are elements $y_1, \ldots, y_d \in A$, which are algebraically independent such that A is a finitely generated $K[y_1, \ldots, y_d]$-module. Furthermore, for some $\delta \leq d$, $I \cap K[y_1, \ldots, y_d] = \langle y_{\delta+1}, \ldots, y_d \rangle$.

(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 $i \in \{1, \ldots, n-1\}$, such that

$$y_n = f = a x_n^m + \sum_1^{m-1} G_i(y_1, \ldots, y_{n-1}) x_n^i \iff y_n - a x_n^m + \sum_1^{m-1} G_i(y_1, \ldots, y_{n-1}) x_n^i = 0,$$

for some $a \in K \setminus \{0\}$, $m > 0$, $G_i \in K[y_1, \ldots, y_{n-1}]$. Then x_n is integral over $K[y_1, \ldots, y_n]$. Since $x_i = y_i + x_n^{r_i} \in A$ we get that $A = K[y_1, \ldots, y_n][x_n]$, hence A is a finitely generated $K[y_1, \ldots, y_n]$-module.

We now claim that $y_1, \ldots, y_n$ are algebraically independent over K. Suppose for contradiction that this is not the case. Then $Y := \{y_1, \ldots, y_n\}$ is not a transcendence basis of $K(y_1, \ldots, y_n)$. However, by Lemma 2.10.45 (a) there is a subset of Y, say $y_{l_1}, \ldots y_{l_k}$ for $k < n$, which constitutes a transcendence basis for $K(y_1, \ldots, y_n)$. Then by Corollary 2.10.50

$$k = \mathbf{trdeg} K(y_1, \ldots, y_n) = \mathbf{trdeg} K(x_1, \ldots, x_n) = n > k,$$

leading to a contradiction.

Let $\lambda \in I \cap K[y_1, \ldots, y_n]$. Then $\lambda = gf = g y_n$ 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. □

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. $\delta = \mathbf{trdeg}_K Q(A) = d$. □

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.*