Algebra II

William Crawley-Boevey

Fakultät für Mathematik, Universität Bielefeld

Summer Semester 2024

We cover a selection of important topics in algebra:

1. Preparations. We introduce the notions of categories and functors, which are useful everywhere in pure mathematics, and we discuss Zorn's lemma, a version of the axiom of choice.

2. Modules. These are generalizations of vector spaces, where the field is replaced by an arbitrary ring. For example, additive groups are the same as modules for the ring $\mathbb{Z}$ of integers. In particular, we discuss free modules, semisimple modules and rings, and Noetherian rings and modules.

3. Multilinear algebra. An important construction for modules is the notion of a tensor product. This allows one to introduce exterior algebras, which provide a coordinate-free way to treat determinants. We also study Clifford algebras.

4. Representations of finite groups. A representation of a group is the occurrence of the group as symmetries. We are interested in linear representations, which are actions of the group on a vector space, or equivalently modules for a suitable ring, the group algebra. The character table contains all the information about its complex representations.

5. Commutative algebra. Commutative rings play an essential role in number theory and algebraic geometry. We discuss localization, integral extensions, the Nullstellensatz, which establishes a correspondence between subsets of affine space and ideals in a polynomial ring, and dimension theory.

Some suggested books:

- J. C. Jantzen und J. Schwermer, Algebra, Springer 2006.
- J. J. Rotman, Advanced Modern Algebra, Third Edition, Part 1, Amer. Math. Soc. 2015.
- M. Artin, Algebra, Birkhäuser 1998.
- J. A. Beachy, Introductory Lectures on Rings and Modules, CUP 2012.
- M. Brandenburg, Einführung in die Kategorientheorie : Mit ausführlichen Erklärungen und zahlreichen Beispielen, Springer 2017.
- P. M. Cohn, Basic algebra : groups, rings and fields, Springer 2005.
- D. J. H. Garling, Clifford algebras : an introduction, CUP 2011.
- G. James and M. Liebeck, Representations and characters of groups, Second Edition, CUP 2001.

Contents

1 Preparations

1.1 Zorn's Lemma

Definition. A **partial order** on a set S is a relation $\leq$ that is reflexive, transitive and antisymmetric, which means that $x \leq y$ and $y \leq x \Rightarrow x = y$. For elements in S we write $x \geq y$ for $y \leq x$, $x < y$ for $x \leq y$ and $x \neq y$, etc.

It is a **total order** if, in addition, for all $x, y \in S$, $x \leq y$ or $y \leq x$ holds.

A **partially ordered set** is a set equipped with a partial order.

Let S be a partially ordered set.

A **chain** is a subset C of S that is totally ordered.

A **largest element** in S is an element $c \in S$ with $x \leq c$ for all $x \in S$. Likewise a **smallest element**. If there is a largest element, it is unique.

An element $x \in S$ is called **maximal** if there is no $y \in S$ with $x < y$. Likewise **minimal**. (A largest element is maximal, but the opposite need not be true.)

An **upper bound** for a subset $X \subseteq S$ is an element $b \in S$ with $x \leq b$ for all $x \in X$.

A partially ordered set is **well-ordered** if every nonempty subset has a smallest element. Considering two-element subsets, we see that a well-ordered set is always totally ordered.

Theorem. *The following are equivalent:*

(i) The axiom of choice: Given a set I and nonempty sets X_i for each $i \in I$, the product $\prod_{i \in I} X_i$ is not empty.

(ii) Zorn's lemma: Let S be a partially ordered set. If every chain in S has an upper bound, then S has a maximal element.

(iii) Every set can be well-ordered.

Remarks. (1) In naive set theory, the axiom of choice seems obvious, so we assume it is true. Therefore, we assume that the equivalent conditions are also true.

The well-ordering property is not obvious. For example, the usual ordering of the real numbers is not a well-ordering. Can you find an ordering?

We will use Zorn's lemma several times.

(2) There are two hierarchies of infinities, one based only on sets, the other on well-ordered sets.

(a) We consider two sets to be equivalent if there is a bijection between them. The **cardinal numbers** are the equivalence classes of sets under this equivalence relation.

The cardinal number of the set $\mathbb{N} = \{0, 1, 2, \dots\}$ is denoted by $\aleph_0$, where aleph is the first letter of the Hebrew alphabet. Thus, a set X has cardinality $\aleph_0$ if and only if there is a bijection between X and $\mathbb{N}$. That is, X is countable and infinite.

(b) We consider two well-ordered sets to be equivalent if there is a bijection between them that preserves the order. The **ordinals** are the equivalence classes.

The ordinal number of the set $\mathbb{N}$ is denoted by ω. For every ordinal S there is a successor $S + 1$ obtained by appending a new largest element to S. For example, $\omega + 1$ corresponds to the well-ordered set $\{0, 1, 2 \dots, \omega\}$.

(3) For further discussion, including a proof of (ii)⇒(iii), I recommend P. M. Cohn, Basic algebra, §1.2.

Theorem. *Every proper ideal I in a ring R is contained in a maximal ideal.*

Proof. Remember: We only consider rings with one. We use: An ideal I is proper if and only if $1 \notin I$. Let S be the set of all proper ideals in R that contain I. It is partially ordered by inclusion. We are looking for a maximal element of S.

Suppose C is a chain in S. If $C = \emptyset$, then I is an upper bound for C. If $C \neq \emptyset$ then

$$K = \bigcup_{J \in C} J$$

is a subset of R. It is an ideal. For example, if $a, b \in K$, then $a \in J$ and $b \in J'$ for $J, J' \in C$. Since C is a chain, $J \subseteq J'$ or $J' \subseteq J$. In the first case, $a + b \in J' \subseteq K$ and in the second case, $a + b \in J \subseteq K$. Now K is a proper ideal, because if $1 \in K$, then $1 \in J$ for some $J \in C$, and then J is not a proper ideal. Thus $K \in S$, and it is an upper bound for C.

By Zorn's lemma, S contains a maximal element. □

Definition. Let V be a vector space over a field K. Let $(v_i)_{i \in I}$ be a tuple of elements of V, where I is a set that is not necessarily finite.

The tuple is **linearly independent** if for all elements $\lambda_i \in K$, all but finitely many of them zero,

$$\sum_{i \in I} \lambda_i v_i = 0 \Rightarrow \lambda_i = 0 \text{ for all } i.$$

The tuple is a **spanning set** for V if for every $v \in V$ there are elements $\lambda_i \in K$, all but finitely many of them zero, such that

$$v = \sum_{i \in I} \lambda_i v_i.$$

The tuple is a **basis** of V if it is linearly independent and a spanning set.

Theorem. *Every vector space V has a basis. If $(v_i)_{i \in I}$ is a spanning set for V and $I' \subseteq I$ is a subset with $(v_i)_{i \in I'}$ linearly independent, then there is a subset I'' with $I' \subseteq I'' \subseteq I$ such that $(v_i)_{i \in I''}$ is a basis.*

Proof. The second statement implies the first: The tuple $(v_i)_{i \in I}$ with $I = V$ and $v_i = i \in V$ is a spanning set for V and if $I' = \emptyset$, then $(v_i)_{i \in I'}$ is linearly independent.

Let S be the set of all subsets A of I with $I' \subseteq A$ such that $(v_i)_{i \in A}$ is linearly independent. S is partially ordered by inclusion.

Suppose C is a chain in S. If $C = \emptyset$, then I' is an upper bound for C. If $C \neq \emptyset$, then

$$B = \bigcup_{A \in C} A$$

is a subset of I. Obviously $I' \subseteq B$. Suppose there is a linear relation

$$\sum_{i \in B} \lambda_i v_i = 0.$$

The set $N = \{i \in B : \lambda_i \neq 0\}$ is finite. Suppose N is not empty. If $i \in N$, then $i \in B$, so $i \in A_i$ for some $A_i \in C$. Since C is a chain and N is finite, there is a $j \in N$ such that $A_i \subseteq A_j$ for all $i \in N$. Then $N \subseteq A_j$. However, this is impossible because $(v_i)_{i \in A_j}$ is linearly independent. Thus $N = \emptyset$, so $(v_i)_{i \in B}$ is linearly independent. So $B \in S$. Thus every chain has an upper bound.

According to Zorn's lemma, there is a maximal element I'' in S. Suppose $(v_i)_{i \in I''}$ is not a spanning set for V. Since $(v_i)_{i \in I}$ is a spanning set, there is $j \in I$ such that v_j cannot be written as a linear combination

$$v_j = \sum_{i \in I''} \lambda_i v_i.$$

But then $(v_i)_{i \in I'' \cup \{j\}}$ is linearly independent. So $I'' \cup \{j\} \in S$. A contradiction to maximality. □

1.2 Categories

Remark. Because of Russell's paradox, there is no set of all sets. One solution is to use classes. A **class** is a collection of things, possibly a set but not necessarily, defined by a property that all things in the class satisfy. There is a class of all sets.

Definition. A **category** $\mathcal{C}$ consists of a class $\mathrm{Ob}(\mathcal{C})$ of **objects** and for each pair $X, Y \in \mathrm{Ob}(\mathcal{C})$ a set $\mathrm{Hom}(X, Y)$ of **morphisms** from X to Y, together with a law of composition

$$\mathrm{Hom}(Y, Z) \times \mathrm{Hom}(X, Y) \to \mathrm{Hom}(X, Z), \quad (g, f) \mapsto g\ f,$$

for all $X, Y, Z \in \mathrm{Ob}(\mathcal{C})$, satisfying the following conditions.

(i) The composition of morphisms is associative, that is, if $f \in \mathrm{Hom}(X,Y)$, $g \in \mathrm{Hom}(Y,Z)$ and $h \in \mathrm{Hom}(Z,U)$, then $h(gf) = (hg)f$.

(ii) For all $X \in \mathrm{Ob}(\mathcal{C})$ there is an identity morphism $\mathrm{Id}_X \in \mathrm{Hom}(X,X)$ such that for all $f \in \mathrm{Hom}(X,Y)$ $f\ \mathrm{Id}_X = f = \mathrm{Id}_Y\ f$. Note that Id_X is unique.

Other notation: Instead of $\mathrm{Hom}(X,Y)$ sometimes $\mathrm{Hom}_{\mathcal{C}}(X,Y)$ or $\mathcal{C}(X,Y)$. Also $f : X \to Y$ means $f \in \mathrm{Hom}(X,Y)$.

A morphism $f : X \to Y$ is an **isomorphism** if there exists a morphism $f' : Y \to X$ with $f'f = \mathrm{Id}_X$ and $ff' = \mathrm{Id}_Y$. If so, then f' is unique, it is called the **inverse** of f and is denoted by f^{-1}.

Examples. (a) Concrete categories:

Set. The objects are sets, the morphisms are mappings $\theta : X \to Y$. It is easy to see that the isomorphisms are the bijective mappings.

Grp. The objects are groups. Morphisms are group homomorphisms. It is easy to see that isomorphisms correspond to the usual definition, i.e. bijective homomorphisms.

Ring. Similar.

K-Vec, where K is a field. The objects are vector spaces over K, morphisms are linear maps. Isomorphisms correspond to the usual definition.

(b) The **opposite category** $\mathcal{C}^{op}$ is given by $\mathrm{Ob}(\mathcal{C}^{op}) = \mathrm{Ob}(\mathcal{C})$ and

$$\mathrm{Hom}_{\mathcal{C}^{op}}(X,Y) = \mathrm{Hom}_{\mathcal{C}}(Y,X)$$

(c) The **product** $\mathcal{C} \times \mathcal{D}$ of two categories has as objects the pairs (X,Y) with X an object in $\mathcal{C}$ and Y an object in D, and

$$\mathrm{Hom}((X,Y),(X',Y')) = \mathrm{Hom}_{\mathcal{C}}(X,X') \times \mathrm{Hom}_{cD}(Y,Y').$$

(d) If G is a group, then there is a category with only one object X, $\mathrm{Hom}(X,X) = G$, and composition given by the multiplication for G. Every morphism is an isomorphism. The same applies to a ring in which only multiplication is used. The isomorphisms correspond to the units.

(e) If S is a partially ordered set, then there is a category with objects of the elements $s \in S$ and

$$\mathrm{Hom}(s,t) = \begin{cases} \{i_{st}\} & (s \leq t) \\ \emptyset & (\text{else}). \end{cases}$$

Definition. A morphism $f : X \to Y$ is a **monomorphism** if for all objects U and $\alpha, \beta : U \to X$, if $f\alpha = f\beta$, then $\alpha = \beta$.

A morphism $f : X \to Y$ is an **epimorphism** if for all objects Z and $\alpha, \beta : Y \to Z$, if $\alpha f = \beta f$, then $\alpha = \beta$.

Remark. For the categories Set and K-Vec, monomorphisms are the same as injective morphisms and epimorphisms are the same as surjective morphisms.

In the category Ring, every surjective morphism is an epimorphism, but the inclusion $f : \mathbb{Z} \to \mathbb{Q}$ is an epimorphism that is not surjective. Namely, if $\alpha, \beta : \mathbb{Q} \to R$ satisfy $\alpha f = \beta f$, then $\alpha(n) = \beta(n)$ for all $n \in \mathbb{Z}$. But since α and β are ring homomorphisms, we have

$$\alpha(n/m) = \alpha(n)\alpha(m^{-1}) = \alpha(n)\alpha(m)^{-1} = \beta(n)\beta(m)^{-1} = \beta(n/m)$$

for $n, m \in \mathbb{Z}$ with $m \neq 0$, so $\alpha = \beta$.

Definition. A **subcategory** $\mathcal{D}$ of a category $\mathcal{C}$ is a category such that:

- Every object of $\mathcal{D}$ is an object of $\mathcal{C}$,

- $\mathrm{Hom}_{\mathcal{D}}(X, Y) \subseteq \mathrm{Hom}_{\mathcal{C}}(X, Y)$ for all $X, Y \in \mathrm{Ob}(\mathcal{D})$.

- The composition of morphisms in $\mathcal{D}$ corresponds to the composition in $\mathcal{C}$ and $\mathrm{Id}_X \in \mathrm{Hom}_{\mathcal{D}}(X, X)$ for all $X \in \mathrm{Ob}(\mathcal{D})$.

A subcategory is **full** if $\mathrm{Hom}_{\mathcal{D}}(X, Y) = \mathrm{Hom}_{\mathcal{C}}(X, Y)$ for all $X, Y \in \mathrm{Ob}(\mathcal{D})$.

Thus a full subcategory of $\mathcal{C}$ is determined by a subclass $\mathrm{Ob}(\mathcal{D})$ of $\mathrm{Ob}(\mathcal{C})$.

Examples. The category Ab of abelian groups is a full subcategory of Grp.

The category CRing of commutative rings is a full subcategory of Ring.

The category K-vec of finite-dimensional vector spaces is a full subcategory of K-Vec.

Consider a group G as a category with an object. Any subgroup $H \leq G$ gives a subcategory.

The subcategory of Set, which is given by all sets and injective (or surjective, or bijective) mappings.

1.3 Functors

Definition. Let $\mathcal{C}$ and $\mathcal{D}$ be categories. A **(covariant) functor** $F : \mathcal{C} \to \mathcal{D}$ is given by

- For every object $X \in \mathrm{Ob}(\mathcal{C})$, an object $F(X) \in \mathrm{Ob}(\mathcal{D})$

- For every morphism $\theta : X \to Y$ in $\mathcal{C}$ a morphism $F(\theta) : F(X) \to F(Y)$ in $\mathcal{D}$

such that the following hold:

(i) $F(\mathrm{Id}_X) = \mathrm{Id}_{F(X)}$ for all $X \in \mathrm{Ob}(\mathcal{C})$.

(ii) $F(g\ f) = F(g)F(f)$ for morphisms g and f in $\mathcal{C}$ that are composable (i.e. $f : X \to Y$ and $g : Y \to Z$).

A **contravariant functor** from $\mathcal{C}$ to $\mathcal{D}$ is a covariant functor $G : \mathcal{C}^{op} \to \mathcal{D}$. So if $\theta : X \to Y$ is a morphism in $\mathcal{C}$, then $G(\theta) : G(Y) \to G(X)$.

Note that if F is a functor and θ is an isomorphism, then $F(\theta)$ is an isomorphism with inverse $F(\theta^{-1})$.

A functor $F : \mathcal{C} \to \mathcal{D}$ is called **full** (respectively **faithful**) if for all $X, Y \in \mathrm{Ob}(\mathcal{C})$ the mapping $F : \mathrm{Hom}_{\mathcal{C}}(X, Y) \to \mathrm{Hom}_{\mathcal{D}}(F(X), F(Y))$, $\theta \mapsto F(\theta)$ is surjective (respectively injective).

A functor $F : \mathcal{C} \to \mathcal{D}$ is said to be **dense** if every object in $\mathcal{D}$ is isomorphic to one of the form $F(X)$ for some $X \in \mathrm{Ob}(\mathcal{C})$.

A functor $F : \mathcal{C} \to \mathcal{D}$ is called an **equivalence** if it is full, faithful and dense.

Examples. (1) The identity functor $\mathrm{Id}_{\mathcal{C}} : \mathcal{C} \to \mathcal{C}$. It's an equivalence!

(2) If $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{E}$ are functors, the composition is a functor $GF : \mathcal{C} \to \mathcal{E}$. If F and G are both full, faithful or dense, the same applies to GF.

(3) The **inclusion functor** of a subcategory into a category, for example CRing → Ring. Inclusion functors are faithful and full for full subcategories.

(4) **Forgetful functors**, e.g. Grp → Set or Ring → Ab forget some or all of the structure. They are faithful.

A concrete category is a category $\mathcal{C}$ with a faithful functor $\mathcal{C} \to$ Set. This makes it possible to consider the objects of the category as sets with additional structure and their morphisms as structure-preserving maps.

(5) Let G be a group. The commutator of $a, b \in G$ is $[a, b] = aba^{-1}b^{-1} \in G$. The commutator group of G is

$$G' = \langle \{[a, b] : a, b \in G\} \rangle \trianglelefteq G.$$

Thus there is a factor group G/G', and it is abelian. If $\theta : G \to H$, then $\theta(G') \subseteq H'$, so there is an induced homomorphism $G/G' \to H/H'$. This defines a functor Grp → Ab, $G \mapsto G/G'$.

(6) Let K be a field. Let $\mathcal{C}$ be the category with $\mathrm{Ob}(\mathcal{C}) = \mathbb{N} = \{0, 1, 2, \dots\}$, $\mathrm{Hom}(n, m) = M_{m\times n}(K)$ and composition given by matrix multiplication. Consider the functor $F : \mathcal{C} \to K$-vec given by $F(n) = K^n$ and for $A \in \mathrm{Hom}(n, m) = M_{m\ timesn}(K)$, $F(A)$ is the corresponding linear map $L_A : K^n \to K^m$ of left multiplication by A. Then F is an equivalence.

(7) The dual of a K vector space V is $V^* = \mathrm{Hom}(V, K)$. If $f : V \to W$ is a linear map, then there is a linear map $f^* : W^* \to V^*$ given by $f^*(h)(v) = h(f(v))$.

This gives a functor $K\text{-Vec}^{op} \to K\text{-Vec}$, $V \mapsto V^*$.

Thus there is a contravariant functor from K-Vec to itself.

For finite-dimensional vector spaces there is an equivalence $K\text{-vec}^{op} \to K\text{-vec}$, $V \mapsto V^*$.

The double dual is again a covariant functor $K\text{-Vec} \to K\text{-Vec}$, $V \mapsto V^{**}$.

This gives an equivalence $K\text{-vec} \to K\text{-vec}$, $V \mapsto V^{**}$

(8) Given $U \in \mathrm{Ob}(\mathcal{C})$ there is a representable functor $\mathrm{Hom}_{\mathcal{C}}(U,-) : \mathcal{C} \to \mathrm{Set}$ that sends an object X to $\mathrm{Hom}_{\mathcal{C}}(U,X)$.

There is also a contravariant representable functor $\mathrm{Hom}(-,U) : \mathcal{C}^{op} \to \mathrm{Set}$.

Definition. Let $F, G : \mathcal{C} \to \mathcal{D}$ be functors. A **natural transformation** $\alpha : F \to G$ is given by morphisms $\alpha_X : F(X) \to G(X)$ for all $X \in \mathrm{Ob}(\mathcal{C})$, such that for every morphism $f : X \to Y$ in $\mathcal{C}$ the following diagram commutes:

$$\begin{tikzcd}
F(X) \arrow[r, "\alpha_X"] \arrow[d, "F(f)"'] & G(X) \arrow[d, "G(f)"] \\
F(Y) \arrow[r, "\alpha_Y"] & G(Y).
\end{tikzcd}$$

It is a **natural isomorphism** if the α_X are all isomorphisms.

Theorem. *A functor $F : \mathcal{C} \to \mathcal{D}$ is an equivalence if and only if there is a functor $G : \mathcal{D} \to \mathcal{C}$ such that FG is naturally isomorphic to* $\mathrm{Id}_{\mathcal{D}}$ *and GF is naturally isomorphic to* $\mathrm{Id}_{\mathcal{C}}$.

Proof. Suppose G and natural isomorphisms $\alpha : GF \to \mathrm{Id}_{\mathcal{C}}$ and $\beta : FG \to \mathrm{Id}_{\mathcal{D}}$ exist. For every $U \in \mathrm{Ob}(\mathcal{D})$ we have $F(G(U)) \cong U$, so F is dense.

Let $\theta, \theta' : X \to Y$ be morphisms in $\mathcal{C}$ with $F(\theta) = F(\theta')$. We have a commutative square

$$\begin{tikzcd}
GF(X) \arrow[r, "\alpha_X"] \arrow[d, "GF(\theta)"'] & X \arrow[d, "\theta"] \\
GF(Y) \arrow[r, "\alpha_Y"] & Y
\end{tikzcd}$$

so

$$\theta = \alpha_Y GF(\theta)\alpha_X^{-1} = \alpha_Y GF(\theta')\alpha_X^{-1} = \theta'.$$

Thus F is faithful. By symmetry G is also faithful.

Let $\phi : F(X) \to F(Y)$. Let $\theta = \alpha_Y G(\phi)\alpha_X^{-1}$. Then we have a commutative square

$$\begin{tikzcd}
GF(X) \arrow[r, "\alpha_X"] \arrow[d, "G(\phi)"'] & X \arrow[d, "\theta"] \\
GF(Y) \arrow[r, "\alpha_Y"] & Y
\end{tikzcd}$$

But we also have the commutative square above, so

$$GF(\theta) = \alpha_Y^{-1}\theta\alpha_X = G(\phi).$$

Now since G is faithful, $\phi = F(\theta)$. So F is full.

Conversely, let us assume that F is an equivalence. We define $G : \mathcal{D} \to \mathcal{C}$, α, β as follows. Since F is dense, for each object $U \in \mathrm{Ob}(\mathcal{D})$ we can choose an object $G(U) \in \mathrm{Ob}(\mathcal{C})$ with $F(G(U))$ isomorphic to U and an isomorphism $\beta_U : F(G(U)) \to U$. Here we use the Axiom of Choice!

If $\phi : U \to W$ is a morphism in $\mathcal{D}$, then $\beta_W^{-1}\phi\beta_U$ is a morphism $FG(U) \to FG(W)$. Since F is full and faithful, there is a unique morphism $\theta : G(U) \to G(W)$ such that $F(\theta) = \beta_W^{-1}\phi\beta_U$. We define $G(\phi) = \theta$.

If $X \in \mathrm{Ob}(\mathcal{C})$, then $\beta_{F(X)} : FGF(X) \to F(X)$, and since F is full and faithful, there is an isomorphism $\alpha_X : GF(X) \to X$ with $\beta_{F(X)} = F(\alpha_X)$. Now it is easy to check that G, α, β satisfy the conditions. □

Remark. As further reading, you could look at M. Brandenburg, Einführung in die Kategorientheorie. There are many interesting topics, such as limits, adjoint functors and Yoneda's Lemma.

2 Modules

2.1 Basics

Recall that a ring R is given by an additive group $(R, +)$ and a multiplication

$$R \times R \to R, \quad (r, s) \mapsto rs$$

which is associative, distributive over addition and has a one, denoted 1 or 1_R.

Definition. Let R be a ring. A **(left) R-module** is an additive group $(M, +)$ together with an operation

$$R \times M \to M, \quad (r, m) \mapsto rm$$

called the **action**, satisfying the following:

(1) For all $r, s \in R$ and $m \in M$ we have $r(sm) = (rs)m$,

(2) $1m = m$ for all $m \in M$,

(3) $r(m+m') = rm+rm'$ and $(r+r')m = rm+r'm$ for all $r, r' \in R$ and $m, m' \in M$.

We sometimes write ${}_RM$ to indicate that M is a left R-module.

A **submodule** of an R-module M is a subset L of M which is a module under the same operations. Equivalently L is an additive subgroup of M and $rx \in L$ for all $r \in R$ and $x \in L$.

If M and N are left R-modules, an **R-module homomorphism** $\theta : M \to N$ is a homomorphism of additive groups with $\theta(rm) = r\theta(m)$ for all $r \in R$ and $m \in M$.

Using composition of mappings, this gives a category R-Mod of left R-modules. The isomorphisms correspond to bijective homomorphisms.

We denote by $\mathrm{Hom}_R(M, N)$ the set of all R-module homomorphisms from M to N. It is naturally an additive group under

$$(\theta + \phi)(m) = \theta(m) + \phi(m)$$

for $\theta, \phi \in \mathrm{Hom}_R(M, N)$ and $m \in M$.

We define $\mathrm{End}_R(M) = \mathrm{Hom}_R(M, M)$. It is naturally a ring under composition.

Examples. (a) If K a field, then left K-modules = K-vector spaces, submodules = subspaces and K-module homomorphisms = linear maps. Thus K-Mod = K-Vec.

(b) Any additive group becomes a $\mathbb{Z}$-module in a unique way, submodules = subgroups and $\mathbb{Z}$-module homomorphisms = group homomorphisms. Thus we can identify $\mathbb{Z}$-Mod and Ab.

(c) If $\phi : S \to R$ is a ring homomorphism, and M is an R-module, then we can turn the additive group $(M, +)$ into an S-module with the action

$$S \times M \to M, \quad (s, m) \mapsto \phi(s)m.$$

We denote this S-module by ${}_S M$. If $f : M \to N$ is a R-module homomorphism, then it gives an S-module homomorphism ${}_S M \to {}_S N$. Thus we obtain a functor R-Mod $\to$ S-Mod. It is called **restriction** to S via ϕ.

(d) If V is a K-vector space and $\theta : V \to V$ is a linear map, then V becomes a $K[X]$-module via

$$(a_0 + a_1 X + a_2 X^2 + \dots)v = a_0 v + a_1 \theta(v) + a_2 \theta^2(v) + \dots.$$

A subset W of V is a $K[X]$-submodule if and only if W is a θ-invariant subspace. we have

$$\operatorname{End}_{K[X]}(V) = \{f \in \operatorname{End}_K(V) : \theta f = f\theta\}.$$

(e) If R is a ring, then there is a ring $M_n(R)$ of $n \times n$ matrices with entries on R. Let R^n be the set of n-tuples of elements of R, written as column vectors. Then R^n is naturally a left $M_n(R)$-module by the usual product of a matrix and a column vector.

Remarks. (1) If M is a left R-module and $r \in R$, then the mapping of left multiplication by r,

$$\lambda_r : M \to M, \quad \lambda_r(m) = rm$$

is not in general an R-module homomorphism, unless R is commutative, but it is a homomorphism of additive groups, so of $\mathbb{Z}$-modules. This gives a mapping

$$\lambda : R \to \operatorname{End}_{\mathbb{Z}}(M), \quad r \mapsto \lambda_r$$

which is a ring homomorphism.

Conversely, given an additive group $(M, +)$ and a ring homomorphism $R \to \operatorname{End}_{\mathbb{Z}}(M)$, we can turn M into an R-module with the action $rm = \lambda(r)(m)$.

Thus a left R-module con either be thought of as an additive group M together with an action $R \times M \to M$, or as an additive group M together with a ring homomorphism $R \to \operatorname{End}_{\mathbb{Z}}(M)$.

(2) Instead of left modules one can define a **right module**, with an action $M \times R \to M$, $(m, r) \mapsto mr$ satisfying $m(rs) = (mr)s$. We write Mod-R for the category of right R-modules.

If R is commutative, then left R-modules correspond to right R-modules, via $rm = mr$.

If R is not commutative, then modules for R on one side, correspond to modules for R^{op} on the other side, where R^{op} is the **opposite ring** to R, given by the same additive group, but with multiplication $\cdot_{op}$ given by

$$r \cdot_{op} s = rs.$$

For example if M is a right R-module, then setting $rm = mr$ we have

$$r(sm) = r(ms) = (ms)r = m(sr) = (sr)m = (r \cdot_{op} s)m$$

so M is a left R^{op}-module.

(3) Multiplication turns any ring R into a left R-module ${}_R R$.

A **left ideal** in R is a submodule of ${}_R R$. Thus it is an additive subgroup L of $(R, +)$ such that $rx \in L$ for all $r \in R$ and $x \in L$.

Similarly R gives a right R-module R_R, and the submodules are called **right ideals** in R. An ideal in R is a subset of R which is a left ideal and a right ideal.

If R is commutative, ideal = left ideal = right ideal.

Properties. (a) If $\theta : M \to N$ is a homomorphism of R-modules, then

$$\operatorname{Ker}\theta = \{m \in M : \theta(m) = 0\}$$

is a submodule of M and

$$\operatorname{Im}\theta = \{\theta(m) : m \in M\}$$

is a submodule of N. More generally, if M' is a submodule of M then $\theta(M')$ is a submodule of N and if N' is a submodule of N, then $\theta^{-1}(N')$ is a submodule of M. [This is the same as for vector spaces.]

(b) If M is an R-module and L is a submodule of M, then the factor group M/L becomes a module with the action $r(L + m) = L + rm$. It is called the **factor** or **quotient** module. The canonical map $M \to M/L$ is a module homomorphism which is surjective and has kernel L. [This is the same as for vector spaces.]

(c) If M is a left R-module and $m \in M$, then the mapping

$$\rho_m : R \to M, \quad \rho_m(r) = rm$$

is an R-module homomorphism since $\rho_m(rs) = rsm = r\rho_m(s)$ for $r, s \in R$.

It gives a mapping

$$\rho : M \to \operatorname{Hom}_R(R, M), \quad m \mapsto \rho_m,$$

This is a homomorphism of additive groups, and it is an isomorphism since if $\rho(m) = 0$ then $\rho_m = 0$, so $\rho_m(1) = 0$, so $1m = 0$, so $m = 0$, and if $f \in \operatorname{Hom}_R(R, M)$ and we take $m = f(1)$, then $\rho_m(r) = rm = rf(1) = f(r1) = f(r)$, so $f = \rho(m)$.

The mapping ρ becomes an isomorphism of R-modules if we turn $\mathrm{Hom}_R(R,M)$ into an R-module with the action

$$R \times \mathrm{Hom}_R(R,M) \to \mathrm{Hom}_R(R,M), \quad (r,\theta) \mapsto r\theta, \quad (r\theta)(r') = \theta(r'r).$$

Namely, $(r\rho(m))(r') = (r\rho_m)(r') = \rho_m(r'r) = r'rm = \rho_{rm}(r') = \rho(rm)(r')$, so $r\rho(m) = \rho(rm)$.

(d) Applying this to the module $M = {}_R R$ gives a ring isomorphism

$$\rho : R^{op} \to \mathrm{End}_R(R).$$

Namely $(\rho_r \rho_s)(t) = \rho_r(\rho_s(t)) = \rho_r(ts) = tsr = \rho_{sr}(t)$.

(e) If M is a left R-module and $m \in M$, then the set

$$Rm = \{rm : r \in R\}$$

is a submodule of M. It is the image of the map ρ_m.

(f) If I is a set and M_i are submodules of M for $i \in I$, then $\bigcap_{i\in I} M_i$ is a submodule of M.

(g) If $M_1, \dots, M_n$ are submodules of M, we define

$$M_1 + \dots + M_n = \{x_1 + \dots + x_n : x_i \in M_i\}.$$

More generally, if I is a set and M_i are submodules of M for $i \in I$, we define

$$\sum_{i\in I} M_i = \{\sum_{i\in I} x_i : x_i \in M_i, \text{ all but finitely many zero}\}.$$

This is a submodule of M.

In Algebra I there were Homomorphism Theorems and Isomorphism Theorems for groups (§1.5) and for rings (§3.3). In just the same way there are versions for modules.

Theorem (Homomorphism Theorem). *Let $\theta : M \to N$ be a homomorphism of left R-modules.*

(1) If L is a submodule of M with $L \subseteq \mathrm{Ker}\,\theta$, then there is a unique homomorphism $\overline{\theta} : M/L \to N$ with $\overline{\theta}(L+m) = \theta(m)$ for $m \in M$.

(2) There is an isomorphism $\overline{\theta} : M/\mathrm{Ker}\,\theta \to \mathrm{Im}\,\theta$ with $\overline{\theta}(\mathrm{Ker}\,\theta + m) = \theta(m)$ for $m \in M$.

Theorem (First Isomorphism Theorem). *Let M be a left R-module and L, N submodules of M. Then there is an isomorphism*

$$L/(L \cap N) \to (L+N)/N, \quad (L\cap N) + x \mapsto N + x.$$

Theorem (Second Isomorphism Theorem)**.** *Let L be a submodule of a left R-module M.*

(1) If N is a submodule of M and $L \subseteq N \subseteq M$, then N/L is a submodule of M/L.

(2) Every submodule U of M/L is of this form for a unique N with $L \subseteq N \subseteq M$, namely $N = \{m \in M : L + m \in U\}$.

(3) In this case there is an isomorphism

$$M/N \to (M/L)/(N/L), \quad N + m \mapsto (N/L) + (L + m).$$

Remark. We denote by 0 the zero additive group $\{0\}$ or zero R-module $\{0\}$.

If M and N are additive groups or R-modules, then any homomorphism $M \to N$ sends 0 to 0. Thus $\mathrm{Hom}_R(M, 0) = 0$ and $\mathrm{Hom}_R(0, M) = \{0\}$. Thus 0 is an initial object in R-Mod and in R-Mod^{op}, so a final object in R-Mod.

Definition. Let R be a ring. A sequence of R-modules and homomorphisms

$$\dots \to X \xrightarrow{f} Y \xrightarrow{g} Z \to \dots$$

is said to be **exact** at Y if $\mathrm{Ker}\, g = \mathrm{Im}\, f$. It is **exact** if it is exact at every module which has homomorphisms in and out.

A **short exact sequence** is one of the form $0 \to X \to Y \to Z \to 0$.

For example a homomorphism $\theta : M \to N$
- is injective if and only if $0 \to M \xrightarrow{\theta} N$ is exact,
- is surjective if and only if $M \xrightarrow{\theta} N \to 0$ is exact.
- is an isomorphism if and only if $0 \to M \xrightarrow{\theta} N \to 0$ is exact.

Any submodule L of M gives a short exact sequence

$$0 \to L \to M \to M/L \to 0.$$

If

$$0 \to X \xrightarrow{f} Y \xrightarrow{g} Z$$

is exact, then f induces an isomorphism $X \to \mathrm{Ker}\, g$. If

$$X \xrightarrow{f} Y \xrightarrow{g} Z \to 0$$

is exact, then g induces an isomorphism $Y/\mathrm{Im}\, f \to Z$ by the Homomorphism Theorem. We call $Y/\mathrm{Im}\, f$ the **cokernel** of f and denote it $\mathrm{Coker}\, f$.

An exact sequence

$$\cdots W \to X \xrightarrow{f} Y \to Z \cdots$$

can be broken into two exact sequences

$$\cdots W \to X \xrightarrow{f} \operatorname{Im} f \to 0 \quad 0 \to \operatorname{Im} f \to Y \to Z \cdots$$

Any morphism $\theta : M \to N$ gives an exact sequence

$$0 \to \operatorname{Ker}\theta \to M \to N \to N/\operatorname{Im}\theta \to 0,$$

so short exact sequences

$$0 \to \operatorname{Ker}\theta \to M \to \operatorname{Im}\theta \to 0 \quad \text{and} \quad 0 \to \operatorname{Im}\theta \to N \to N/\operatorname{Im}\theta \to 0.$$

Definition. For any R-module M, we get a representable functor

$$\operatorname{Hom}_R(M, -) : R - \operatorname{Mod} \to \operatorname{Ab}.$$

It sends an R-module X to the additive group $\operatorname{Hom}_R(M, X)$, and it sends a morphism $f \in \operatorname{Hom}_R(X, Y)$ to the mapping $\operatorname{Hom}_R(M, X) \to \operatorname{Hom}_R(M, Y)$, $\theta \mapsto f\theta$.

Similarly, we get a contravariant representable functor, so a functor

$$\operatorname{Hom}_R(-, M) : R - \operatorname{Mod}^{op} \to \operatorname{Ab}$$

sending X to $\operatorname{Hom}_R(X, M)$ and a morphism $f \in \operatorname{Hom}_R(X, Y)$ to the mapping $\operatorname{Hom}_R(Y, M) \to \operatorname{Hom}_R(X, M)$, $\theta \mapsto \theta f$.

Proposition. *If M is an R-module and $0 \to X \xrightarrow{f} Y \xrightarrow{g} Z \to 0$ is an exact sequence of R-modules, then the following sequences are exact:*

(i) $0 \to \operatorname{Hom}_R(M, X) \to \operatorname{Hom}_R(M, Y) \to \operatorname{Hom}_R(M, Z)$*, and*

(ii) $0 \to \operatorname{Hom}_R(Z, M) \to \operatorname{Hom}_R(Y, M) \to \operatorname{Hom}_R(X, M)$.

Proof. (i) If $\theta \in \operatorname{Hom}_R(M, X)$ is sent to zero, then $f\theta = 0$. Thus $f(\theta(m)) = 0$ for all $m \in M$. But f is injective, so $\theta(m) = 0$ for all m, so $\theta = 0$.

If $\phi \in \operatorname{Hom}_R(M, Y)$ is sent to zero, then $g\phi = 0$, so $\phi(m) \in \operatorname{Ker} g = \operatorname{Im} f$. Thus for each $m \in M$ there is a unique $x_m \in X$ with $f(x_m) = \phi(m)$. We define $\theta(m) = x_m$, so $f\theta = \phi$. It remains to check that θ is an R-module homomorphism, so that ϕ is in the image of the map $\operatorname{Hom}_R(M, X) \to \operatorname{Hom}_R(M, Y)$. This is straightforward. For example $\phi(rm) = r\phi(m) = rf(x_m) = f(rx_m)$, so by uniqueness $x_{rm} = rx_m$, so $\theta(rm) = r\theta(m)$.

(ii) If $\theta \in \operatorname{Hom}_R(Z, M)$ is send to zero, then $\theta g = 0$. Thus $\theta(g(m)) = 0$ for all $m \in M$. Since g is surjective, it follows that $\theta(z) = 0$ for all $z \in Z$, so $\theta = 0$.

Suppose $\phi \in \operatorname{Hom}_R(Y, M)$ is sent to zero, so $\phi f = 0$. We define a homomorphism $\theta \in \operatorname{Hom}_R(Z, M)$ as follows. If $z \in Z$, then $z = g(y)$ for some $y \in Y$, and we define $\theta(z) = \phi(y)$. This is well-defined, for if $z = g(y) = g(y')$, then $y - y' \in \operatorname{Ker} g = \operatorname{Im} f$, so $y - y' = f(x)$ for some x, so $\phi(y) - \phi(y') = \phi(y - y') = \phi(f(x)) = 0$. Finally we need that $\theta \in \operatorname{Hom}_R(Z, M)$. For example if $z, z' \in Z$ and $z = g(y)$, $z' = g(y')$, then $z + z' = g(y + y')$, so $\theta(z + z') = \phi(y + y') = \phi(y) + \phi(y') = \theta(z) + \theta(z')$. □

2.2 Finitely generated and noetherian modules

Definition. Given a subset S of a module M, we define the submodule of M **generated** by S as the intersection of all submodules containing S. It is the unique smallest submodule containing S. Clearly it is equal to

$$\sum_{m\in S} Rm.$$

A module M which can be generated by one element m is called a **cyclic**. Thus $M = Rm$. A module is **finitely generated** (f.g.) if it can be generated by a finite set $\{m_1, \dots, m_n\}$, so

$$M = Rm_1 + \cdots + Rm_n.$$

Proposition. *A module is cyclic if and only if it is isomorphic to a quotient R/L with L a left ideal in R.*

Proof. The module R/L is cyclic, generated by the element $L + 1_R$. Say M is cyclic, so $M = Rm$. The mapping $R \to M$, $r \mapsto rm$ is a module homomorphism, and it is surjective. The kernel L is a left ideal in R, and by the Homomorphism Theorem, $M \cong R/L$. □

Theorem. *Any proper submodule L of a finitely generated module M is contained in a maximal (proper) submodule.*

Proof. We use Zorn's Lemma. Let S be the set of all proper submodules of M containing L and let $m_1, \dots, m_n$ be a generating set of M.

Let C be a chain in S. We want to show that C has an upper bound in S. If C is empty, then L is an upper bound. Thus suppose C is not empty. Let

$$B = \bigcup_{N\in C} N.$$

Since C is a chain, it is easy to see that B is a submodule of M. Moreover it is proper, for if $B = M$, then $m_1, \dots, m_n \in B$. Then each $m_i \in N_i$ for some $N_i \in C$. Since C is a chain, there is some j with $N_i \subseteq N_j$ for all i. Then all $m_i \in N_j$, so $N_j = M$, which is a contradiction.

Thus by Zorn's Lemma S has a maximal element. □

Lemma. *Suppose M is a module and N a submodule.*

(i) If M is f.g. then so is M/N.

(ii) If N and M/N are f.g., so is M.

[In general, if M is f.g., it does not follow that N is f.g.]

Proof. (i) is clear.

(ii) Suppose $n_1, \ldots, n_r$ are generators of N and $N + m_1, \ldots, N + m_s$ are generators of M/N. We claim that $n_1, \ldots, n_r, m_1, \ldots, m_s$ are generators of M.

Let X be the submodule generated by these elements. Since it contains the n_i, it contains N. Also X/N contains the $M + m_i$, so $X/N = M/N$. Thus $X = M$. □

Theorem. *Let M be a left R-module. The following are equivalent.*

*(i) M is **noetherian**, that is, any ascending chain of submodules of M*

$$M_1 \subseteq M_2 \subseteq \ldots$$

***breaks off**, meaning that there is some n such that $M_n = M_{n+1} = \ldots$. [This is also called the **ascending chain condition** on submodules of M.]*

(ii) Any non-empty set S of submodules of M has a maximal element.

(iii) Every submodule of M is finitely generated.

Proof. (i)⇒(ii) Suppose S has no maximal element. Choose $M_1 \in S$. Since M_1 isn't maximal, there is $M_2 \in S$ with $M_1 \subset M_2$, $M_1 \neq M_2$. Since M_2 isn't maximal, there is $M_3 \in S$ with $M_2 \subset M_3$, $M_2 \neq M_3$. This gives an ascending chain which doesn't break off.

(ii)⇒(iii) Let N be a submodule M. Let S be the set of f.g. submodules of N. It is nonempty since $\{0\} \in S$. Thus it has a maximal element L. If $L \neq N$ then there is $x \in N \setminus L$, and $L + Rx \in S$ and L is a proper submodule of $L + Rx$, contradicting maximality of L.

(iii)⇒(i) Let

$$M_1 \subseteq M_2 \subseteq \cdots \subseteq M$$

be an ascending chain. Then $N = \sum M_i = \bigcup M_i$ is a submodule of M, so N is f.g., say by $n_1, \ldots, n_r$. Now each n_i belongs to some M_{j_i}. Then all $n_1, \ldots, n_r \in M_j$ where $j = \max\{j_1, \ldots, j_r\}$. But then $N \subseteq M_j \subseteq N$. Thus $M_j = M_{j+1} = \ldots$, so the ascending chain breaks off. □

Definition. A ring R is **left noetherian** if ${}_R R$ is noetherian.
A commutative ring R is **noetherian** if ${}_R R$ is noetherian.

Thus a commutative ring is noetherian
⇔ each ideal is finitely generated
⇔ any ascending chain of ideals breaks off
⇔ any non-empty set of ideals has a maximal element.

Examples: $\mathbb{Z}$, since any ideal is principal, and fields K, since they have no ideals except 0 and K.

Example. Let $R = K[X_1, X_2, \dots]$, a polynomial ring over a field K in infinitely many indeterminates. It is not noetherian since it has an ascending chain of ideals

$$I_1 \subseteq I_2 \subseteq \dots, \quad I_n = \sum_{i=1}^{n} RX_i$$

which does not break off (since $X_{n+1} \in I_{n+1} \setminus I_n$).

Then R is a f.g. (even cyclic) R-module, but it has submodules which are not f.g.

Theorem. *Let M be a left R-module..*

(i) If N is a submodule of M, then M is noetherian if and only if N and M/N are noetherian.

(ii) If L and N are noetherian submodules of M, so is $L + N$.

(iii) If R is a left noetherian ring, then M is noetherian if and only if it is f.g.

Proof. (i) Suppose M is noetherian. Clearly N is noetherian. Also any ascending chain of submodules of M/N is of the form

$$M_1/N_1 \subseteq M_2/N \subseteq \dots$$

for some ascending chain $M_1 \subseteq M_2 \subseteq \dots$, so it breaks off.

Now suppose N and M/N are noetherian. If M_i is an ascending chain of submodules of M, then $M_i \cap N$ is an ascending chain of submodules of N, and $(N + M_i)/N$ is an ascending chain of submodules of M/N. Thus there is n such that $M_n \cap N = M_{n+1} \cap N = \dots$ and $(N + M_n)/N = (N + M_{n+1})/N = \dots$. Then $N + M_n = N + M_{n+1} = \dots$. Then if $x \in M_{n+1}$, we have $x = n + y$ with $n \in N$ and $y \in M_n$, so $n = x - y \in N \cap M_{n+1} = N \cap M_n$, so $y = x - (x - y) \in M_n$. Thus $M_n = M_{n+1} = \dots$.

(ii) $(L + N)/N \cong L/(L \cap N)$ by the First Isomorphism Theorem, and this is noetherian by (i), hence so is $L + N$.

(iii) Suppose R is left noetherian and M is f.g.. If $m \in M$ then Rm is isomorphic to a quotient of ${}_RR$, so it is noetherian by (i). Then $Rm_1 + \dots + Rm_n$ is noetherian by (ii) and induction. □

Theorem (Hilbert's Basis Theorem)**.** *If K is a commutative noetherian ring, then so is the polynomial ring $R = K[X]$.*

Proof. Suppose that I is an ideal in R which is not finitely generated.

Take a nonzero polynomial $f_1(X)$ of least degree in I. By induction, suppose we have fixed $f_1(X), \dots, f_k(X) \in I$. Since I is not finitely generated, $\sum_{i=1}^{k} Rf_i(X)$ is a proper subset of I, so we can choose $f_{k+1}(X) \in I$ of least degree not in this subset. Thus we obtain an infinite sequence of polynomials $f_1(X), f_2(X), \dots$. Let $f_i(X)$

have degree d_i and leading coefficient a_i. We have $d_1 \leq d_2 \leq \dots$ by construction. Consider the ascending chain

$$Ka_1 \subseteq Ka_1 + Ka_2 \subseteq \dots$$

of ideals in K. If it breaks off, then for some n we have $a_{n+1} = \sum_{i=1}^n \lambda_i a_i$ with $\lambda_i \in K$. But then

$$f_{n+1}(X) - \sum_{i=1}^n \lambda_i X^{d_{n+1}-d_i} f_i(X)$$

would be an element of $I \setminus \sum_{i=1}^n Rf_i(X)$ of degree less than d_{n+1}. This contradicts the minimality of the degree of $f_{n+1}(X)$. □

Examples. The following are commutative noetherian rings:

- $\mathbb{Z}$ and any field K,

- $K[X_1, \dots, X_n]$ with K a commutative noetherian ring,

- R/I with R a commutative noetherian ring and I an ideal in R.

2.3 Products, direct sums and free modules

Definition. Suppose we are given a left R-module M_i for all i in a set I.

The product $\prod_{i\in I} M_i$ becomes an R-module with the operations

$$(m_i) + (m_i') = (m_i + m_i'), \quad r(m_i) = (rm_i).$$

The **(external) direct sum** is the submodule

$$\bigoplus_{i\in I} M_i = \{(m_i) \in \prod_{i\in I} M_i : m_i = 0 \text{ for all but finitely many } i\}.$$

For a finite index set we have equality:

$$M_1 \times M_2 \times \cdots \times M_n = M_1 \oplus M_2 \oplus \cdots \oplus M_n.$$

In case all terms M_i are equal to M, we use the notation

$$M^I = \prod_{i\in I} M, \quad M^{(I)} = \bigoplus_{i\in I} M.$$

The first can be identified with the set of mappings $I \to M$, the second with the mappings that send all but finitely many elements of I to 0. (Of course $M^\emptyset = M^{(\emptyset)} = \{0\}$.)

For $j \in I$ there are R-module homomorphisms

$$\pi_j : \prod_{i\in I} M_i \to M_j, \quad (m_i) \mapsto m_j$$

and

$$\mu_j : M_j \to \bigoplus_{i\in I} M_i, \quad m \mapsto (m_i) \text{ where } m_j = m \text{ and } m_i = 0 \text{ for } i \neq j$$

Proposition. *For any family of R-modules M_i $(i \in I)$ and R-module N we have isomorphisms of additive groups*

$$\operatorname{Hom}_R(N, \prod_{i\in I} M_i) \to \prod_{i\in I} \operatorname{Hom}_R(N, M_i), \quad \theta \mapsto (\pi_i\theta)$$

and

$$\operatorname{Hom}_R(\bigoplus_{i\in I} M_i, N) \to \prod_{i\in I} \operatorname{Hom}_R(M_i, N), \quad \phi \mapsto (\phi\mu_i).$$

Proof. It is easy to see that the mappings are homomorphisms of additive groups.

The first is a bijection since it has inverse mapping

$$\prod_{i\in I} \operatorname{Hom}_R(N, M_i) \to \operatorname{Hom}_R(N, \prod_{i\in I} M_i)$$

sending (θ_i) to the map θ with $\theta(n) = (\theta_i(n))$.

The second is a bijection since it has inverse mapping

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

sending (ϕ_i) to the map ϕ with $\phi((m_i)) = \sum_{i\in I} \phi_i(m_i)$ for $(m_i) \in \bigoplus_{i\in I} M_i$. The sum makes sense since only finitely many terms are nonzero. □

Definition. Let M_i $(i \in I)$ be a family of submodules of a module M. By taking the inclusion map $M_i \to M$ for each i, we obtain a homomorphism

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

The image is $\sum_{i\in I} M_i$. If this homomorphism is an isomorphism, we say that M is the **(internal) direct sum** of the modules M_i $(i \in I)$ and write

$$M = \bigoplus_{i\in I} M_i.$$

Note that if M is the external direct sum of of M_i, then by the homomorphisms μ_i we can identify each M_i as a submodule of M, and then M is the internal direct sum of these copies of M_i.

Proposition. *If L and N are submodules of M then $M = L \oplus N$ if and only if $M = L + N$ and $L \cap N = \{0\}$.*

Proof. The kernel of the map $L \oplus N \to M$, $(\ell, n) \mapsto \ell + n$ is

$$\{(x, -x) : x \in L \cap N\}.$$

□

Lemma. *Suppose $M = \bigoplus_{i\in I} M_i$. Then M is f.g. if and only if the modules M_i are all f.g. and all but finitely many of them are zero.*

Proof. Exercise. □

Definition. Let M be a module and let (x_i) be a tuple of elements of M indexed by a set I.

Recall that the submodule of M generated by the x_i is $\sum_{i\in I} Rx_i$.

Thus the x_i generate M if and only if each element of M can be written in the form $\sum_{i\in I} r_i x_i$ with $r_i \in R$, all but finitely many zero.

We say that (x_i) is a **basis** of M if each element of M can be written uniquely in the form $\sum_{i\in I} r_i x_i$ with $r_i \in R$, all but finitely many zero.

A module is **free** if it has a basis.

Examples. (i) Over a field K, the notion of a basis agrees with the definition for vector spaces, so every K-module has a basis.

(ii) R is a free R-module with basis (1_R).

(iii) If I is a set, then the R-module $R^{(I)}$ is free with basis (e_i), where e_i is the tuple with ith entry 1_R and the other entries 0. Namely, the element $(r_i) \in R^{(I)}$ can be written uniquely in the form $\sum_{i\in I} r_i e_i$.

(iv) If $n > 0$ then the $\mathbb{Z}$-module $\mathbb{Z}/\mathbb{Z}n$ is not free, since for any $x \in \mathbb{Z}/\mathbb{Z}n$ we have the relation $nx = 0 = 0x$, so x cannot be part of a basis.

(v) The $\mathbb{Z}$-module $\mathbb{Q}$ is not free. It does not have a basis (x) with one element, since there is no $x \in \mathbb{Q}$ with $\mathbb{Q} = \{nx : n \in \mathbb{Z}\}$. Also if $x = a/n$ and $y = b/m$ are two non-zero elements of $\mathbb{Q}$, then $bnx = ab = amy$, so x and y cannot both be in a basis.

Theorem. *(i) An R-module is free if and only if it is isomorphic to $R^{(I)}$ for some I. Explicitly, if $(x_i)_{i\in I}$ is a basis of M, then the mapping*

$$R^{(I)} \to M, \quad (r_i) \mapsto \sum_{i\in I} r_i x_i$$

is an isomorphism of R-modules.

(ii) If M is a free module with basis $(x_i)_{i\in I}$ and N is any module, then the mapping

$$\operatorname{Hom}_R(M, N) \cong N^I, \quad \theta \mapsto (\theta(x_i))$$

is an isomorphism of additive groups.

(iii) A free module is finitely generated if and only if it has a finite basis.

(iv) Every (f.g.) module M can be written as a quotient of a (f.g.) free module.

Proof. (i) If M is free, then the stated map is clearly an isomorphism. Conversely if $\theta : R^{(I)} \to M$ is an isomorphism, then $(\theta(e_i))$ is a basis.

(ii) We have

$$\operatorname{Hom}_R(M,N) \cong \operatorname{Hom}_R(R^{(I)},N) = \operatorname{Hom}_R(\bigoplus_{i\in I} R, N)$$

$$\cong \prod_{i\in I} \operatorname{Hom}_R(R,N) \cong \prod_{i\in I} N = N^I.$$

(iii) Follows from the lemma about finitely generated direct sums, for if $R \neq 0$ and I is infinite, then $R^{(I)}$ is not f.g..

(iv) Every module M can be generated by some subset, e.g. M itself, and by the lemma, a generating set gives a surjective homomorphism $\theta : F \to M$ with F a free module. (This means that $M \cong F/\operatorname{Ker}\theta$ so it is a quotient of F.) □

Definition. Given a set X, we denote by RX the set of formal sums

$$\sum_{x\in X} r_x x$$

with $r_x \in R$, all but finitely many zero. It is a free R-module with basis identified with the set X. Suppose we have elements $c_i \in RX$ for i in a set I. They generate a submodule $\sum_{i\in I} Rc_i$ of RX. The quotient

$$RX/\sum_{i\in I} Rc_i$$

is called the **module generated by X subject to the relations $c_i = 0$ $(i \in I)$**. We denote the image of the basis element $x \in X$ in this module also by x.

2.4 Semisimple modules and rings

Definition. A ring R is a **division ring** or **skew field** if $R \neq 0$ and every non-zero element is invertible. Thus a commutative division ring is a field.

An R-module S is **simple** or **irreducible** if it has exactly two submodules, namely 0 and S. In particular, a simple module is nonzero.

If S is simple, then it is cyclic, generated by any non-zero element of S.

Lemma (Schur's Lemma)**.** *(i) If M and N are simple R-modules, then any nonzero homomorphism $\theta : M \to N$ is an isomorphism.*

(ii) If M and N are non-isomorphic simple modules, then $\mathrm{Hom}_R(M, N) = 0$.

(iii) If M is simple, then $\mathrm{End}_R(M)$ *is a division ring.*

Proof. (i) If $\theta \neq 0$ then $\mathrm{Im}\,\theta$ is nonzero, so it must be N, and $\mathrm{Ker}\,\theta$ is not M, so it must be 0.

(ii),(iii) Follow. □

Definition. An R-module M is **semisimple** or **completely reducible** if every submodule has a complement. That is, for every submodule N of M, there is a submodule C with $N \oplus C = M$.

Proposition. *Any submodule or quotient of a semisimple module is semisimple.*

Proof. If N is a submodule of a semisimple module M and L is a submodule of N, then L has a complement C in M, so $L \oplus C = M$. We claim that $L \oplus (C \cap N) = N$.

Since $L \cap C = 0$ we have $L \cap (C \cap N) = 0$.

Clearly $L + (C \cap N) \subseteq N$.

Also if $n \in N$, then we can write $n = \ell + c$ for some $\ell \in L$ and $c \in C$. Then $c = n - \ell \in N$, so $c \in C \cap N$. Thus $n \in L + (C \cap N)$. Thus $N = L + (C \cap N)$.

Thus N is semisimple.

Now suppose M is semisimple and consider a quotent M/N. Then N has a complement C, so $M = N \oplus C$. But then the map $C \to M/N$ is an isomorphism, and C is semismple, hence so is M/N. □

Theorem. *For an R-module M the following conditions are equivalent*

(a) M is semisimple

(b) M is a sum of simple submodules

(c) M is a direct sum of simple modules.

Proof. (a)⇒(b). Let N be the sum of all simple submodules of M. Suppose $N \neq M$. Now N has a complement $C \neq 0$. Let $0 \neq c \in C$. Then Rc is a non-zero f.g. submodule of C. Thus Rc has a maximal submodule L. Then Rc is semisimple, so L has a complement D, and it must be simple since L is maximal. Thus $D \subseteq N$. Contradiction.

(c)⇒(b) is trivial.

Suppose (b) holds and N is a submodule of M. We show that N has a complement in M which is a direct sum of simple modules. This proves (a) and (c).

Let I be the set of simple submodules of M. So $M = \sum_{S\in I} S$. Consider the set X of subsets J of I such that $N \oplus \bigoplus_{S\in J} S$ is a direct sum.

Any chain Y in X has an upper bound $J' = \bigcup_{J\in Y} J$ in X, for if the sum $N + \sum_{S\in J'} S$ is not direct, then the same holds for a finite subset of J', and this is contained in an element of Y.

Thus by Zorn's lemma there is a maximal element $J \in X$. Now if $N \oplus \bigoplus_{S\in J} S \neq M$, then there is some simple submodule T not contained in $N \oplus \bigoplus_{S\in J} S$. But then $J \cup \{T\} \in X$, contradiction.

Thus $\bigoplus_{S\in J} S$ is a complement to N which is a direct sum of simple modules. □

Corollary. *A direct sum of semisimple modules is semisimple.*

Proof. If $M = \bigoplus_{i\in I} M_i$ and each $M_i = \bigoplus_{j\in I_i} S_i$, then $M = \bigoplus_{j\in J} S_j$ where J is the disjoint union of the sets I_i with $i \in I$. □

Theorem. *A semisimple R-module M is f.g. if and only if it is a direct sum of finitely many simple submodules. Moreover if*

$$M = S_1 \oplus \cdots \oplus S_n = T_1 \oplus \cdots \oplus T_m$$

are two such decompositions as direct sums of simple submodules, then $n = m$ and there is a permutation σ such that $S_i \cong T_{\sigma(i)}$ for all i.

Proof. The first statement follows from the lemma about f.g. direct sums.

We show the rest by induction on n. Let $N = S_1 \oplus \dots S_{n-1}$. Then $M/N \cong S_n$, so it is simple. For some i we must have $T_i \not\subseteq N$, so $(N + T_i)/N$ is a nonzero submodule of M/N, so it equals M/N, so $N + T_i = M$. Thus the homomorphism $N \to M/T_i$, $x \mapsto T_i + x$ is surjective. It is also injective, for the kernel is $N \cap T_i$, and if this is not zero, then it is equal to T_i, so $T_i \subseteq N$, a contradiction. Thus we have an isomorphism between $N = S_1 \oplus \dots S_{n-1}$ and $M/T_i \cong \bigoplus_{j\neq i} T_j$. By induction we get $n - 1 = m - 1$ and a bijection $\pi : \{1, \dots, i-1, i+1, \dots, m\} \to \{1, \dots, n-1\}$ with $T_j \cong S_{\pi(j)}$, hence $n = m$ and we get σ. □

The theory of vector spaces extends to division rings.

Theorem. *Every module for a division ring is semisimple and free. Moreover the modules with a finite basis are exactly the f.g. modules, and they have a well defined dimension, the number of elements in any basis.*

Proof. If R is a division ring, then it has no left ideals other than 0 and R, so ${}_RR$ is a simple module.

Any non-zero cyclic module is isomorphic to R/L with L a proper left ideal, so isomorphic to ${}_RR$, so simple.

Any simple module is cyclic, so isomorphic to ${}_RR$.

Now every module is the sum of its cyclic submodules $M = \sum_{m \in M} Rm$, so semisimple. Thus it is isomorphic to a direct sum of copies of simple modules, so free.

Now any basis gives a decomposition as a direct sum of copies of the simple module R, and the number of terms in the decomposition is uniquely determined by the last theorem. □

Theorem (Artin-Wedderburn Theorem)**.** *For a ring R, the following are equivalent, in which case we call R a **semisimple ring** (or more precisely, a **semisimple artinian ring**).*

(i) ${}_RR$ is semisimple.

(ii) Every left R-module is semisimple.

(iii) R is isomorphic to a product of matrix rings over division rings,

$$R \cong M_{n_1}(D_1) \times \cdots \times M_{n_r}(D_r)$$

with $n_i \geq 1$ and the D_i division rings.

Proof. (i)⇒(ii) Using that a semisimple module is one which is a sum of simple modules, we see that any direct sum of semisimple modules is semisimple. Thus (i) implies that any free module is semisimple, and then any module is a quotient of a free module, so semisimple.

(ii)⇒(i) Clear

(i)⇒(iii) Given R-modules $M_1, \dots, M_n$ we have

$$\operatorname{End}(M_1 \oplus \cdots \oplus M_n) \cong \prod_{i=1}^{n} \prod_{j=1}^{n} \operatorname{Hom}_R(M_j, M_i).$$

We can display elements as matrices

$$\begin{pmatrix} \theta_{11} & \dots & \theta_{1n} \\ \vdots & \ddots & \vdots \\ \theta_{n1} & \dots & \theta_{nn} \end{pmatrix}$$

with $\theta_{ij} \in \operatorname{Hom}_R(M_j, M_i)$, and then composition corresponds to matrix multiplication.

Now write R as a direct sum of simple submodules. Since ${}_RR$ is a finitely generated module, only finitely many terms can appear in the direct sum. Collecting terms, we have

$${}_RR \cong \underbrace{S_1 \oplus \cdots \oplus S_1}_{n_1} \oplus \cdots \oplus \underbrace{S_r \oplus \cdots \oplus S_r}_{n_r}$$

for pairwise non-isomorphic simple modules S_i.

Let $E_i = \mathrm{End}_R(S_i)$. By Schur's Lemma these are division rings and elements of $\mathrm{End}_R(R)$ correspond to matrices which have the block form

$$\begin{pmatrix} A_1 & 0 & \dots \\ 0 & A_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

with $A_i \in M_{n_i}(E_i)$. Thus

$$R^{op} \cong \mathrm{End}_R(R) \cong M_{n_1}(E_1) \times \cdots \times M_{n_r}(E_r).$$

Then using the transposes of the matrices we get an isomorphism

$$R \cong M_{n_1}(D_1) \times \cdots \times M_{n_r}(D_r)$$

where $D_i = E_i^{op}$.

(iii)⇒(i) Suppose $R = M_{n_1}(D_1) \times \cdots \times M_{n_r}(D_r)$. Consider the elements which are zero except for the jth column of the ith factor $M_{n_i}(D_i)$. They give a left ideal in R isomorphic to the module $S_i = D_i^{n_i}$. This module is simple since D_i is a division ring, as on the exercise sheet for a field. Now R is the direct sum of these simple left ideals, so ${}_RR$ is semisimple. □

3 Algebras and multilinear algebra

3.1 Algebras

Definition. Let K be a commutative ring, e.g. $\mathbb{Z}$ or a field. An **(associative and unital) algebra over** K or K**-algebra** is a ring R which is also a K-module, with the same addition, and such that multiplication is bilinear, that is,

$$(\lambda x + \lambda' x')y = \lambda(xy) + \lambda'(x'y), \quad x(\lambda y + \lambda' y') = \lambda(xy) + \lambda'(xy')$$

for all $x, x', y, y' \in R$ and $\lambda, \lambda' \in K$.

An **algebra homomorphism** $R \to S$ is a ring homomorphism which is also a homomorphism of K-modules. This gives a category of K-algebras. A **subalgebra** is a subring which is also a K-submodule.

Remarks. (a) The centre of a ring R is the set

$$Z(R) = \{r \in R : rs = sr \text{ for all } s \in R\}.$$

It is a subring of R and it is commutative.

If R is a K-algebra, then the mapping $K \to R$, $\lambda \mapsto \lambda 1_R$ is a ring homomorphism with image contained in $Z(R)$.

Conversely, if we are given a ring R and a homomorphism $\theta : K \to R$ with image contained in $Z(R)$, then ${}_R R$ becomes a K-module via restriction and this turns R into a K-algebra.

(b) If R is a K-algebra, then any R-module M becomes a K-module ${}_K M$ by restriction using the homomorphism $K \to R$.

If M and N are R-modules, then $\mathrm{Hom}_R(M, N)$ becomes a K-module via

$$(\lambda\theta)(m) = \lambda\theta(m)$$

for $\lambda \in K$, $\theta \in \mathrm{Hom}_R(M, N)$ and $m \in M$.

(c) To give an R-module, it is equivalent to give a K-module M and an algebra homomorphism $R \to \mathrm{End}_K(M)$. (For the case $K = \mathbb{Z}$, see Remark (1) in §2.1.)

(d) If R is a free K-module with basis $(x_i)_{i \in I}$, then any bilinear multiplication is uniquely determined by its **structure coefficients** $c_{ijk} \in K$, where

$$x_i x_j = \sum_{k \in I} c_{ijk} x_k.$$

Examples. (1) Every ring is a $\mathbb{Z}$-algebra in a unique way.

(2) Any commutative ring R is naturally an R-algebra using the module structure ${}_R R$ or equivalently the identity homomorphism $R \to R$.

(3) If L/K is a field extension, then L is naturally a K-algebra via the inclusion $K \to L$.

(4) $M_n(K)$ is a K-algebra. As a K-module it is free, with basis E^{ij} for $1 \leq i, j \leq n$, where E^{ij} is the matrix with 1 in the (i, j) position and zero elsewhere, so $E^{ij}E^{pq} = \delta_{jp}E^{iq}$.

(5) Hamilton's **Quaternions** is the $\mathbb{R}$-algebra given by the 4-dimensional real vector space

$$\mathbb{H} = \{a1 + bi + cj + dk : a, b, c, d \in \mathbb{R}\}$$

with multiplication defined by expanding out, bringing coefficients to the front and using the rules

$$\begin{array}{ccc} i^2 = -1 & ij = k & ik = -j \\ ji = -k & j^2 = -1 & jk = i \\ ki = j & kj = -i & k^2 = -1 \end{array}$$

See Linear Algebra II §10.4. For example

$$(1+2i+j)(i-3k) = 1i+2i^2+ji-3k-6ik-3jk = i-2-k-3k+6j-3i = -2-2i+6j-4k.$$

It is a non-commutative division ring, with

$$q^{-1} = \frac{1}{|q|^2}\,\overline{q}$$

for $q \neq 0$, where $\overline{a + bi + cj + dk} = a - bi - cj - dk$ and $|a + bi + cj + dk| = \sqrt{a^2 + b^2 + c^2 + d^2}$.

(6) The polynomial ring $K[X_1, \dots, X_n]$ is naturally a K-algebra. A monomial is an expression of the form $m = X_1^{r_1}X_2^{r_2}\dots X_n^{r_n}$ with $r_i \geq 0$. We can omit any term with $r_i = 0$, and we write the monimal with all $r_i = 0$ as $m = 1$. Then a polynomial is an expression

$$\sum_m a_m m$$

with coefficients $a_m \in K$, all but finitely many zero, and the sum runs over all monomials m. The multiplication is given by expanding out, bringing coefficients to the front, and using

$$(X_1^{r_1}X_2^{r_2}\dots)(X_1^{s_1}X_2^{s_2}\dots) = X_1^{r_1+s_1}X_2^{r_2+s_2}\dots.$$

(7) The **free algebra** $K\langle X_1, \dots, X_n\rangle$. A word of length r is a sequence $X_{i_1}X_{i_2}\dots X_{i_r}$ with $i_1, i_2, \dots, i_r \in \{1, \dots, n\}$. The word of length $r = 0$ is denoted 1. An element of $K\langle X_1, \dots, X_n\rangle$ is an expression

$$\sum_w a_w w$$

with coefficients $a_w \in K$, all but finitely many zero, and the sum runs over all words w. The multiplication is given by expanding out, bringing coefficients to the front, and using the concatenation of words:

$$(X_{i_1}X_{i_2}\dots X_{i_r})(X_{j_1}X_{j_2}\dots X_{j_r}) = X_{i_1}X_{i_2}\dots X_{i_r}X_{j_1}X_{j_2}\dots X_{j_r}.$$

For example in $K\langle X,Y\rangle$ we have

$$(a+bX)(c+dX+eYX) = ac+(ad+bc)X+aeYX+bdX^2+beXYX.$$

for $a,b,c,d,e \in K$. The elements are sometimes called noncommuting polynomials, since $XY \neq YX$ in $K\langle X,Y\rangle$.

When working with algebras over a field K, life is often simplest if the field is algebraically closed.

Definition. A field K is **algebraically closed** if it satisfies the following equivalent conditions.

(i) Any non-constant polynomial in $K[X]$ has a root in K.

(ii) Any non-constant polynomial in $K[X]$ splits into linear factors.

(iii) Any irreducible polynomial in $K[X]$ is linear.

(iv) If L/K is a field extension and $a \in L$ is algebraic over K, then $a \in K$.

(v) If L/K is a finite field extension then $L = K$.

It is easy to see that these are equivalent. We know that $\mathbb{C}$ is algebraically closed.

Theorem. *(i) If K is an algebraically closed field, then the only finite-dimensional division algebra over K is K itself.*

(ii) (Frobenius) The only finite-dimensional division algebras over $\mathbb{R}$ are $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$.

Proof. (i) Let R be a f.d. division algebra over K and let $r \in R$. The powers of r cannot be linearly independent, so there is some monic polynomial $f(X) \in K[X]$ with $f(r) = 0$. We can factorize into monic linear factors $f(X) = f_1(X)\dots f_n(X)$, and then $f_1(r)\dots f_n(r) = 0$. Since R is a division ring, it has no zero divisors, so some $f_j(r) = 0$. But $f_j(X)$ is linear, of the form $X - \lambda$, so $r = \lambda 1$, so $R = K1 \cong K$.

(ii) Let R be a f.d. division algebra over $\mathbb{R}$. We make several steps.

(a) If $r \in R$, then either $r \in \mathbb{R}1$ or $r = \lambda 1 + \mu i$ for some $\lambda, \mu \in \mathbb{R}$ and some element $i \in R$ with $i^2 = -1$.

Proof. Again $f(r) = 0$ for some monic $f(X) \in \mathbb{R}[X]$.

We can write $f(X) = f_1(X)\dots f_n(X)$, a product of monic linear and irreducible quadratic factors, and some $f_j(r) = 0$.

If $f_j(X)$ is linear, say $X - \lambda$, then $r = \lambda 1 \in \mathbb{R}1$.

If $f_j(X)$ is quadratic, then since it is irreducible it has the form $X^2 + bX + c$ with $d = b^2 - 4c < 0$. Then r has the required form with $i = (c - b^2/4)^{-1/2}(r + b/2)$.

(b) For $r \in R$ let $\hat{r} : R \to R$ be left multiplication by R. It is a K-linear map. Then $R = \mathbb{R}1 \oplus T$ where $T = \{r \in R : \operatorname{tr} \hat{r} = 0\}$. Moreover T contains any element $i \in R$ with $i^2 = -1$.

Proof. If $r \in \mathbb{R}1 \cap T$ then $r = \lambda 1$, so $\operatorname{tr} \hat{r} = \lambda \operatorname{tr} \hat{1} = \lambda \dim R$, so $\lambda = 0$.

If $r \in R$ then $r = \lambda 1 \oplus (r - \lambda 1)$, and $\operatorname{tr} \widehat{r - \lambda 1} = \operatorname{tr} \hat{r} - \lambda \dim R$. Taking $\lambda = (\operatorname{tr} \hat{r})/ \dim R$ this gives that $r \in \mathbb{R}1 + T$.

Suppose $i^2 = -1$. Then 1 and i span a subalgebra of R, which we can identify with $\mathbb{C}$.

We can't necessarily consider R as an algebra over $\mathbb{C}$, but we can at least consider it as a vector space.

If $(e_1, \ldots, e_n)$ is a $\mathbb{C}$-basis of R, then $(e_1, ie_1, e_2, ie_2, \ldots, e_n, ie_n)$ is a basis for R as a vector space over $\mathbb{R}$. With respect to this basis, the matrix of $\hat{i}$ consists of diagonal blocks

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Thus it has trace 0.

(c) $R \cong \mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$.

Proof. If $0 \neq r \in T$, then r^2 is real and negative, for by (a) we can write $r = \lambda 1 + \mu i$, and then $0 = \operatorname{tr} \hat{r} = \lambda \operatorname{tr} \hat{1} + \mu \operatorname{tr} \hat{i} = \lambda \dim R$, so $\lambda = 0$, so $a = \mu i$, so $a^2 = -\mu^2$.

It follows that T becomes a real inner product space with inner product

$$\langle r, s \rangle = -\frac{1}{2}(rs + sr) = \frac{1}{2}(r^2 + s^2 - (r + s)^2) \in \mathbb{R}.$$

Choose an orthonormal basis $(e_1, \ldots, e_n)$ of T. Then $e_i^2 = -1$ and $e_i e_j = -e_j e_i$ for $i \neq j$.

If $n = 0$ or 1 then clearly $R \cong \mathbb{R}$ or $R \cong \mathbb{C}$, with e_1 corresponding to i, so suppose $n \geq 2$.

Suppose $n = 2$. We have $(e_1e_2)^2 = e_1e_2e_1e_2 = -e_1e_1e_2e_2 = -(-1)(-1) = -1$, so by (b) we have $e_1e_2 \in T$. Thus $e_1e_2 = \lambda e_1 + \mu e_2$ for some $\lambda, \mu \in \mathbb{R}$. Then $e_1(e_1e_2) = -e_2 \in T$, but $e_1(\lambda e_1 + \mu e_2) = -\lambda + \mu e_1e_2$, so $\lambda = 0$. Similarly $\mu = 0$, which is nonsense since $e_1e_2 \neq 0$.

Thus suppose $n \geq 3$. If $3 \leq i \leq n$, then e_i commutes with e_1e_2, so

$$(e_i - e_1e_2)(e_i + e_1e_2) = e_i^2 - (e_1e_2)^2 = 0.$$

So, since R is a division algebra, $e_i = \pm e_1 e_2$. Since the e_i are linearly independent, this gives a contradiction unless $n = 3$, and then there is an isomorphism $R \cong \mathbb{H}$, with i corresponding to e_1, j to e_2 and k to $e_1 e_2 = \pm e_3$. □

Example. Up to isomorphism, the 9-dimensional semisimple K-algebras for K algebraically closed are

$$\underbrace{K \times \cdots \times K}_{9},\ M_2(K) \times \underbrace{K \times \cdots \times K}_{5},\ M_2(K) \times M_2(K) \times K,\ M_3(K).$$

The 4-dimensional semisimple $\mathbb{R}$-algebras are

$$\mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R},\ \mathbb{R} \times \mathbb{R} \times \mathbb{C},\ \mathbb{C} \times \mathbb{C},\ \mathbb{H},\ M_2(\mathbb{R}).$$

3.2 Tensor products

Definition. Let R be a ring. The **tensor product** of a right R-module X and a left R-module Y is a $\mathbb{Z}$-module $X \otimes_R Y$, equipped with a mapping

$$X \times Y \to X \otimes_R Y, \quad (x, y) \mapsto x \otimes y$$

such that the mapping is a homomorphism of additive groups in each argument, that is

$$x \otimes y + x' \otimes y = (x + x') \otimes y,\ x \otimes y + x \otimes y' = x \otimes (y + y'),$$

for $x, x' \in X$ and $y, y' \in Y$, the mapping is R-balanced, that is

$$xr \otimes y = x \otimes ry$$

for all $x \in X$, $y \in Y$ and $r \in R$, and such that it is universal for this property. That is, if

$$f : X \times Y \to M$$

is a mapping to an additive group which is a homomorphism in each argument and is R-balanced, then there is a unique homomorphism $\alpha : X \otimes_R Y \to M$ such that $f(x, y) = \alpha(x \otimes y)$.

Theorem. *The tensor product exists and is unique up to isomorphism. It is the $\mathbb{Z}$-module generated by symbols $x \otimes y$ for $x \in X$ and $y \in Y$, subject to the relations*

$$x \otimes y + x' \otimes y - (x + x') \otimes y = 0,$$

$$x \otimes y + x \otimes y' - x \otimes (y + y') = 0,$$

$$xr \otimes y - x \otimes ry = 0$$

for $x, x' \in X$, $y, y' \in Y$ and $r \in R$.

Proof. The construction means the following. We take the free $\mathbb{Z}$-module

$$F_{X,Y} = \bigoplus_{(x,y)\in X\times Y} \mathbb{Z}(x\otimes y)$$

with basis the symbols $x \otimes y$. Let $S_{X,Y}$ be the $\mathbb{Z}$-submodule generated by the elements

$$\begin{gathered} x\otimes y + x'\otimes y - (x+x')\otimes y, \\ x\otimes y + x\otimes y' - x\otimes(y+y'), \\ xr\otimes y - x\otimes ry. \end{gathered}$$

for $x, x' \in X$, $y, y' \in Y$ and $r \in R$. The elements $x \otimes y$ in $F_{X,Y}$ induce elements in $F_{X,Y}/S_{X,Y}$ which should be denoted $S_{X,Y} + x\otimes y$ or $\overline{x\otimes y}$, but we shall just denote them $x\otimes y$. Then by construction the mapping $X\times Y \to F_{X,Y}/S_{X,Y}$ sending (x,y) to $x\otimes y$ is a homomorphism in each argument and is R-balanced.

To show that this is the tensor product, we need to show it has the universal property. Let $f : X\times Y \to M$ is a mapping which is a homomorphism in each argument and is R-balanced. We need to show that there is a unique homomorphism $\alpha : F_{X,Y}/S_{X,Y} \to M$ with $\alpha(x\otimes y) = f(x,y)$ for $(x,y) \in X\times Y$. Uniqueness holds since $F_{X,Y}/S_{X,Y}$ is generated by the elements $x\otimes y$. Since $F_{X,Y}$ is a free module, by part (ii) of the theorem in section 2.3, there is a unique $\mathbb{Z}$-module homomorphism $\phi : F_{X,Y} \to M$ with $\phi(x\otimes y) = f(x,y)$. Now ϕ sends the generators of $S_{X,Y}$ to zero, so $\phi(S_{X,Y}) = 0$, so it descends to a homomorphism $\bar\phi : F_{X,Y}/S_{X,Y} \to M$ and $\alpha = \bar\phi$ gives existence.

For uniqueness, suppose that $X \otimes'_R Y$ is another tensor product, equipped with the mapping $X\times Y \to X\otimes'_R Y$, $(x,y)\mapsto x\otimes' y$.

By the universal property of $X\otimes_R Y$ there is a unique homomorphism $\alpha : X\otimes_R Y \to X\otimes'_R Y$ with $x\otimes' y = \alpha(x\otimes y)$.

By the universal property of $X\otimes'_R Y$ there is a unique homomorphism $\alpha' : X\otimes'_R Y \to X\otimes_R Y$ with $x\otimes y = \alpha'(x\otimes' y)$.

Now the homomorphisms $\beta = \mathrm{Id}$ and $\beta = \alpha'\alpha$ both satisfy $\beta(x\otimes y) = x\otimes y$. Thus by the uniqueness part of the universal property for $X\otimes_R Y$ we have $\alpha'\alpha = \mathrm{Id}$. Similarly $\alpha\alpha' = \mathrm{Id}$. Thus α and α' are inverse isomorphisms. □

Lemma. *Let X be a right R-module and let Y be a left R-module.*

(i) For all $x\in X$ and $y \in Y$ we have $x\otimes 0 = 0\otimes y = 0$ and $n(x\otimes y) = (nx)\otimes y = x\otimes(ny)$ for $n\in\mathbb{Z}$.

(ii) Any element in $X\otimes_R Y$ can be written (non-uniquely) as a sum

$$x_1\otimes y_1 + x_2\otimes y_2 + \cdots + x_r\otimes y_r$$

(iii) Considering X as a left R^{op}-module and Y as a right R^{op}-module, there is an isomorphism of additive groups $X \otimes_R Y \to Y \otimes_{R^{op}} X$ with $x \otimes y \mapsto y \otimes x$.

(iv) If $\theta : X \to X'$ is a homomorphism of right R-modules and $\phi : Y \to Y'$ is a homomorphism of left R-modules, then there is a unique homomorphism of additive groups

$$\theta \otimes \phi : X \otimes_R Y \to X' \otimes_R Y'$$

with $(\theta \otimes \phi)(x \otimes y) = \theta(x) \otimes \phi(y)$.

Proof. (i) For fixed y, the map $X \to X \otimes_R Y$, $x \mapsto x \otimes y$ is a homomorphism of additive groups, and these are standard properties for such homomorphisms. Similarly for fixed x.

(ii) Follows from the definition and (i).

(iii) The definitions are the same, but with different notation.

(iv) Consider the mapping

$$f : X \times Y \to X' \otimes_R Y', \quad f(x,y) = \theta(x) \otimes \phi(y).$$

This is a homomorphism in each argument, for example $f(x+x', y) = \theta(x+x') \otimes y = (\theta(x) + \theta(x')) \otimes y = \theta(x) \otimes y + \theta(x') \otimes y$.

It is R-balanced, since $f(xr, y) = \theta(xr) \otimes \phi(y) = \theta(x)r \otimes \phi(y) = \theta(x) \otimes r\phi(y) = \theta(x) \otimes \phi(r) = f(x, ry)$.

Now the homomorphism $\theta \otimes \phi : X \otimes_R Y \to X' \otimes_R Y'$ follows from the universal property for $X' \otimes_R Y'$. □

Suppose that R is a K-algebra.

Any left R-module Y becomes a left K-module via $\lambda y = (\lambda 1_R) y$ for $\lambda \in K$ and $y \in Y$.

Any right R-module X becomes a left K-module via $\lambda x = x(\lambda 1_R)$ for $\lambda \in K$ and $x \in X$.

It turns out that tensor products are also left K-modules in this case.

Proposition. *Suppose that R is a K-algebra.*

(i) If X is a right R-module and Y is a left R-module, then $X \otimes_R Y$ has a unique structure as a K-module with

$$\lambda(x \otimes y) = (\lambda x) \otimes y = x \otimes (\lambda y)$$

for $\lambda \in K$, $x \in X$ and $y \in Y$.

(ii) The mapping

$$X \times Y \to X \otimes_R Y, \quad (x,y) \mapsto x \otimes y$$

is K-bilinear, that is, a K-module map in each argument, and R-balanced. Moreover it is universal for K-bilinear R-balanced maps.

Proof. (i) We define an action of K on $X \otimes Y$ by $\lambda m = (\theta_\lambda \otimes \mathrm{Id}_Y)(m)$ for $m \in X \otimes_R Y$, where

$$\theta_\lambda : X \to X, \quad x \mapsto \lambda x,$$

which is a homomorphism of right R-modules since $(\lambda x)r = x(\lambda 1_R)r = xr(\lambda 1_R) = \lambda(xr)$. Then $\lambda(x \otimes y) = \theta_\lambda(x) \otimes y = (\lambda x) \otimes y = (x(\lambda 1_R)) \otimes y = x \otimes (\lambda 1_R)y) = x \otimes (\lambda y)$.

(ii) By (i) the mapping is K-bilinear, and by definition it is R-balanced. To show the universal property, suppose that $f : X \times Y \to M$ is a K-bilinear R-balanced map to a K-module. By the definition of the tensor product, there is a unique homomorphism of additive groups $\alpha : X \otimes_R Y \to M$ with $\alpha(x \otimes y) = f(x, y)$. Now α is a K-module homomorphism since

$$\alpha(\lambda(x \otimes y)) = \alpha((\lambda x) \otimes y) = f(\lambda x, y) = \lambda f(x, y) = \lambda\alpha(x \otimes y).$$

□

Theorem. *Suppose R is a K-algebra. Let X and X_i be right R-modules and let Y and Y_i be left R-modules.*

(i) The map $x \mapsto x \otimes 1$ is an isomorphism of K-modules $X \to X \otimes_R R$. The inverse sends $x \otimes r$ to xr. Similarly the map $y \to 1 \otimes y$ is an isomorphism $Y \to R \otimes_R Y$, and the inverse sends $r \otimes y$ to ry.

(ii) $\left(\bigoplus_{i\in I} X_i\right) \otimes_R Y \cong \bigoplus_{i\in I}(X_i \otimes_R Y)$ and $X \otimes_R \left(\bigoplus_{i\in I} Y_i\right) \cong \bigoplus_{i\in I}(X \otimes_R Y_i)$.

(iii) If X is a free right R-module with basis $(x_i)_{i\in I}$ then there is an isomorphism of additive groups

$$Y^{(I)} \to X \otimes_R Y, \quad (y_i) \mapsto \sum_{i\in I} x_i \otimes y_i.$$

Similarly, if Y is a free left R-module with basis $(y_j)_{j\in J}$ then there is an isomorphism of additive groups

$$X^{(J)} \to X \otimes_R Y, \quad (x_j) \mapsto \sum_{j\in J} x_j \otimes y_j.$$

(iv) If X' is a submodule of X, then $(X/X') \otimes_R Y$ is isomorphic to the quotient of $X \otimes_R Y$ by the K-submodule generated by all elements of the form $x' \otimes y$ with $x' \in X'$ and $y \in Y$. Similarly for $X \otimes_R (Y/Y')$.

(v) If $X_1 \to X_2 \to X_3 \to 0$ is an exact sequence of right R-modules, then

$$X_1 \otimes_R Y \to X_2 \otimes_R Y \to X_3 \otimes_R Y \to 0$$

is exact. Similarly, if $Y_1 \to Y_2 \to Y_3 \to 0$ *is an exact sequence of left* R*-modules, then*

$$X \otimes_R Y_1 \to X \otimes_R Y_2 \to X \otimes_R Y_3 \to 0$$

is exact.

Proof. We use the universal property.

(i) We have a homomorphism $X \to X \otimes_R R$ given by $x \mapsto x \otimes 1$.

The map $X \times R \to X$, $(x, r) \mapsto xr$ is K-bilinear and R-balanced, so corresponds to a homomorphism $X \otimes_R R \to X$ with $x \otimes r \mapsto xr$.

These homomorphisms are clearly inverses, since $x \otimes r = xr \otimes 1$.

(ii) Let $\mu_j : X_j \to \bigoplus_{i\in I} X_i$ be the canonical map. Then $\mu_j \otimes \mathrm{Id}_Y : X_j \otimes_R Y \to (\bigoplus_{i\in I} X_i) \otimes_R Y$. Varying j, these give a homomorphism

$$\bigoplus_{j\in I}(X_j \otimes_R Y) \to (\bigoplus_{i\in I} X_i) \otimes_R Y.$$

On the other hand we have a mapping

$$(\bigoplus_{i\in I} X_i) \times Y \to \bigoplus_{i\in I}(X_i \otimes_R Y), \quad ((x_i), y) \mapsto (x_i \otimes y)$$

This is K-bilinear and R-balanced, so it corresponds to a homomorphism

$$(\bigoplus_{i\in I} X_i) \otimes_R Y \to \bigoplus_{i\in I}(X_i \otimes_R Y).$$

These homomorphisms are inverses.

(iii) Follows from (i) and (ii).

(iv) The canonical map $X \to X/X'$ induces a homomorphism $X \otimes_R Y \to (X/X') \otimes_R Y$, and this map kills the K-submodule S generated by elements of the form $x' \otimes y$, so it induces a homomorphism $(X \otimes_R Y)/S \to (X/X') \otimes_R Y$.

For $y \in Y$, if $x_1, x_2 \in X$ and $X' + x_1 = X' + x_2$, then $x_1 - x_2 \in X'$, so $x_1 \otimes y - x_2 \otimes y = (x_1 - x_2) \otimes y \in S$. Thus we have a well-defined mapping

$$(X/X') \times Y \to (X \otimes Y)/S, \quad (X' + x, y) \mapsto S + (x \otimes y).$$

This is K-bilinear and R-balanced, so corresponds to a homomorphism

$$(X/X') \otimes_R Y \to (X \otimes_R Y)/S.$$

These mappings are clearly inverses.

(v) Follows from (iv), taking $X = X_2$, X' the image of the map $X_1 \to X_2$, and identifying X_3 with X/X'. □

Example. We consider tensor products over $\mathbb{Z}$. Since $\mathbb{Z}$ is commutative, left and right $\mathbb{Z}$-modules are the same - they are additive groups. Now

$X \otimes_{\mathbb{Z}} Y \cong Y \otimes_{\mathbb{Z}} X$.

$X \otimes_{\mathbb{Z}} \mathbb{Z} \cong X$

$X \otimes_R \mathbb{Z}^n \cong X^n$.

$\mathbb{Z}^m \otimes \mathbb{Z}^n \cong (\mathbb{Z}^m)^n \cong \mathbb{Z}^{mn}$.

If $a \in \mathbb{Z}$ then $(\mathbb{Z}/\mathbb{Z}a) \otimes_{\mathbb{Z}} Y \cong Y/aY$.

If $a, b \in \mathbb{Z}$ then $a(\mathbb{Z}/\mathbb{Z}b)$ is the set of cosets $a(\mathbb{Z}b + x) = \mathbb{Z}b + ax$ with $x \in \mathbb{Z}$, so $a(\mathbb{Z}/\mathbb{Z}b) = (\mathbb{Z}b + \mathbb{Z}a)/\mathbb{Z}b$. Thus $(\mathbb{Z}/\mathbb{Z}a) \otimes_{\mathbb{Z}} (\mathbb{Z}/\mathbb{Z}b) \cong (\mathbb{Z}/\mathbb{Z}b)/((\mathbb{Z}b + \mathbb{Z}a)/\mathbb{Z}b) \cong \mathbb{Z}/(\mathbb{Z}b + \mathbb{Z}a) = \mathbb{Z}/\mathbb{Z}\,\mathrm{ggT}(a, b)$.

$(\mathbb{Z}/\mathbb{Z}2) \otimes_{\mathbb{Z}} (\mathbb{Z}/\mathbb{Z}3) \cong \mathbb{Z}/\mathbb{Z}1 = 0$.

For $a \neq 0$ we have $(\mathbb{Z}/\mathbb{Z}a) \otimes_{\mathbb{Z}} \mathbb{Q} = \mathbb{Q}/a\mathbb{Q} = \mathbb{Q}/\mathbb{Q} = 0$.

$\mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} \cong \mathbb{Q}$. Namely, there are homomorphisms $\mathbb{Q} \to \mathbb{Q} \otimes_{\mathbb{Z}} \mathbb{Q} \to \mathbb{Q}$ sending a to $a \otimes 1$ and sending $a \otimes b$ to ab. They are inverses since $a \otimes b$ is sent to ab and then to $ab \otimes 1$. But if $b = n/m$ with $n, m \in \mathbb{Z}$ and $m \neq 0$, then

$$ab \otimes 1 = a(n/m) \otimes 1 = (a/m)n \otimes m(1/m) = (a/m)m \otimes n(1/m) = a \otimes b.$$

Definition. Let S and R be rings. An **S-R-bimodule** is an additive group M which is both a left S-module and a right R-module, such that $(sm)r = s(mr)$ for all $m \in M$, $s \in S$ and $r \in R$.

Proposition. *Let X be an S-R-bimodule.*

(i) If Y is a left R-module, then $X \otimes_R Y$ becomes an S-module, with $s(x \otimes y) = (sx) \otimes y$. In fact $X \otimes_R -$ defines a functor from R-Mod *to* S-Mod.

(ii) If Z is a left S-module, then $\mathrm{Hom}_S(X, Z)$ *becomes a left R-module via $(r\phi)(x) = \phi(xr)$. In fact* $\mathrm{Hom}_S(X, -)$ *defines a functor from* S-Mod *to* R-Mod.

(iii) We have $\mathrm{Hom}_S(X \otimes_R Y, Z) \cong \mathrm{Hom}_R(Y, \mathrm{Hom}_S(X, Z))$. *(In fact this is a natural isomorphism of functors in Y and Z, which says that the functors $X \otimes_R -$ and* $\mathrm{Hom}_S(X, -)$ *are* adjoint.*)*

Proof. (i) The action of $s \in S$ is given by $\theta_s \otimes \mathrm{Id}_Y$, where $\theta_s : X \to X$ is the right R-module homomorphism with $\theta_s(x) = sx$.

It is straightforward to check that this turns $X \otimes_R Y$ into a left S-modules. Moreover if $\phi : Y \to Y'$ is a homomorphism of left R-modules, then $\mathrm{Id}_X \otimes \phi : X \otimes_R Y \to X \otimes_R Y'$ is a homomorphism of left S-modules.

(ii) Easy

(iii) We construct mappings in both directions. For the direction

$$\mathrm{Hom}_S(X \otimes_R Y, Z) \to \mathrm{Hom}_R(Y, \mathrm{Hom}_S(X, Z)),$$

suppose that $\theta \in \mathrm{Hom}_S(X \otimes_R Y, Z)$. For $y \in Y$, define

$$\phi_{y,\theta} : X \to Z, \quad \phi_{y,\theta}(x) = \theta(x \otimes y).$$

Then $\phi_{y,\theta} \in \mathrm{Hom}_S(X, Z)$ since

$$\phi_{y,\theta}(sx) = \theta((sx) \otimes y) = \theta(s(x \otimes y)) = s\theta(x \otimes y) = s\phi_{y,\theta}(x).$$

Let $\phi_\theta : Y \to \mathrm{Hom}_S(X, Z)$ be the mapping sending y to $\phi_{y,\theta}$. Then $\phi_\theta \in \mathrm{Hom}_R(Y, \mathrm{Hom}_S(X, Z))$ since

$$\phi_\theta(y+y')(x) = \phi_{y+y',\theta}(x) = \theta(x \otimes (y+y')) = \theta(x \otimes y) + \theta(x \otimes y') = \phi_\theta(y)(x) + \phi_\theta(y')(x)$$

so $\phi_\theta(y + y') = \phi_\theta(y) + \phi_\theta(y')$, and

$$\phi_\theta(ry)(x) = \theta(x \otimes ry) = \theta(xr \otimes y) = (\phi_\theta(y))(xr) = (r(\phi_\theta(y))(x)$$

so $\phi_\theta(ry) = r\phi_\theta(y)$. Thus we use the mapping $\theta \mapsto \phi_\theta$.

For the direction

$$\mathrm{Hom}_R(Y, \mathrm{Hom}_S(X, Z)) \to \mathrm{Hom}_S(X \otimes_R Y, Z),$$

suppose $\psi \in \mathrm{Hom}_R(Y, \mathrm{Hom}_S(X, Z))$. There is a mapping $f_\psi : X \otimes Y \to Z$ given by $f_\psi(x, y) = \psi(y)(x)$. This mapping is a homomorphism of additive groups in each argument and it is R-balanced, since

$$f_\psi(xr, y) = \psi(y)(xr) = (r(\psi(y)))(x) = \psi(ry)(x) = f_\psi(x, ry).$$

Thus there is a mapping $\alpha_\psi : X \otimes_R Y \to Z$ with $\alpha_\psi(x \otimes y) = f_\psi(x, y) = \psi(y)(x)$.

Now $\alpha_\psi \in \mathrm{Hom}_S(X \otimes_R Y, Z)$ since

$$\alpha_\psi(s(x \otimes y)) = \alpha_\psi((sx) \otimes y) = \psi(y)(sx) = s\psi(y)(x) = s\alpha_\psi(x \otimes y).$$

Thus we use the mapping $\psi \mapsto \alpha_\psi$.

Now it is easy to see that these two mappings are homomorphisms of additive groups and inverse to each other. □

3.3 Tensor, Clifford and exterior algebras

In this section K is a commutative ring and we consider left K-modules. We use the letter K because the most important case is when K is a field.

Let V and W be K-modules. Because K is commutative we can consider V as a right K-module and form the tensor product $V \otimes_K W$. We write it as $V \otimes W$.

Properties (of tensor products of modules for a commutative ring).

(i) $V \otimes W$ is naturally a K-module with $\lambda(v \otimes w) = (\lambda v) \otimes w = v \otimes (\lambda w)$ for $\lambda \in K$, $v \in V$, $w \in W$.

This comes from considering K as a K-algebra.

(ii) If V is free with basis $(v_i)_{i \in I}$, then every element in $V \otimes W$ can be written uniquely in the form

$$\sum_{i \in I} v_i \otimes w_i$$

with the $w_i \in W$, all but finitely many zero. Similarly if W is free.

This holds since $V \otimes W \cong K^{(I)} \otimes W \cong W^{(I)}$.

(iii) If V is free with basis $(v_i)_{i \in I}$ and W is free with basis $(w_j)_{j \in J}$ then $V \otimes W$ is a free K-module with basis $(v_i \otimes w_j)_{(i,j) \in I \times J}$. In particular, if K is a field, then $\dim V \otimes W = (\dim V)(\dim W)$.

By (ii) we can write each element of $V \otimes W$ uniquely in the form

$$\sum_{i \in I} v_i \otimes (\sum_{j \in J} a_{ij} w_j) = \sum_{(i,j) \in I \times J} a_{ij} v_i \otimes w_j$$

with the $a_{ij} \in K$, all but finitely many zero.

(iv) Symmetry. There is an isomorphism $V \otimes W \cong W \otimes V$, $v \otimes w \mapsto w \otimes v$.

This follows from a general result, since $K^{op} = K$.

(v) The map $V \times W \to V \otimes W$ is K-bilinear, and universal for K-bilinear maps $V \times W \to M$.

Any K-bilinear map $f : V \times W \to M$ is automatically K-balanced, since $f(x\lambda, y) = \lambda f(x,y) = f(x, \lambda y)$. Thus it follows from the universal property of the tensor product, the first proposition in §3.2.

(vi) The map

$$U \times V \times W \to (U \otimes V) \otimes W, \quad (u,v,w) \mapsto (u \otimes v) \otimes w$$

is K-multilinear, that is, a K-module map in each argument, and it is universal for K-multilinear maps $U \times V \times W \to M$.

Proof. The mapping is K-multilinear. Suppose $f : U \times V \times W \to M$ is K-multilinear. For fixed $w \in W$ the map $U \times V \to M$, $(u,v) \mapsto f(u,v,w)$ is K-bilinear, so there is a K-module map $\alpha_w : U \otimes V \to M$ with $\alpha_w(u \otimes v) = f(u,v,w)$. Now the map $(U \otimes V) \times W \to M$, $(\xi, w) \mapsto \alpha_w(\xi)$ is K-bilinear. Thus there is a

K-module map $\alpha : (U \otimes V) \otimes W \to M$ with $\alpha(\xi, w) = \alpha_w(\xi)$, so $\alpha((u \otimes v) \otimes w) = \alpha_w(u \otimes v) = f(u, w, z)$.

(vii) Associativity. There is an isomorphism of K-modules $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$ with $(u \otimes v) \otimes w$ sent to $u \otimes (v \otimes w)$.

Proof. A similar argument to (vi) shows that

$$U \times V \times W \to U \otimes (V \otimes W), \quad (u, v, w) \mapsto (u \otimes v) \otimes w$$

is also universal for K-multilinear maps $U \times Z \times W \to M$. Now we have uniqueness, similar to the uniqueness of the tensor product.

Definition. Let V be a K-module. For $d \geq 0$, the dth **tensor power** of V is

$$T^d(V) = \begin{cases} \underbrace{V \otimes \cdots \otimes V}_{d} & (d > 0) \\ K & (d = 0). \end{cases}$$

By the associativity property above, the tensor product can be computed with any bracketing, and for all d, e we have an isomorphism of K-modules

$$T^d(V) \otimes T^e(V) \to T^{d+e}(V)$$

$$(v_1 \otimes \cdots \otimes v_d) \otimes (w_1 \otimes \cdots \otimes w_e) \mapsto v_1 \otimes \cdots \otimes v_d \otimes w_1 \otimes \cdots \otimes w_e$$

Properties (of tensor powers)**.**

(i) The mapping $V^d \to T^d(V)$, $(v_1, \ldots, v_d) \mapsto v_1 \otimes v_2 \otimes \cdots \otimes v_n$ is K-multilinear, that is, it is a K-module map in each argument, and it is universal with this property.

This is proved by induction on d, with the inductive step being analogous to property (vi) of tensor products of modules for a commutative ring.

(ii) If $\theta : V \to W$ is a K-module homomorphism, there is a homomorphism

$$T^d(\theta) : T^d(V) \to T^d(W), \quad v_1 \otimes \cdots \otimes v_d \mapsto \theta(v_1) \otimes \cdots \otimes \theta(v_d).$$

Note that $T^0(\theta) : K \to K$ is the identity map.

To see this, apply the universal property to the map

$$V^d \to T^d(W), \quad (v_1, \ldots, v_d) \mapsto \theta(v_1) \otimes \cdots \otimes \theta(v_d).$$

(iii) If V has basis $(b_1, \ldots, b_n)$, then $T^d(V)$ has basis given by the elements $b_{i_1} \otimes \cdots \otimes b_{i_d}$ with $1 \leq i_1, \ldots, i_d \leq n$, so with n^d elements.

Definition. An algebra R over K is **graded** if it is equipped with a decomposition

$$R = \bigoplus_{d=0}^{\infty} R_d$$

as a K-module, such that if $x \in R_d$ and $y \in R_e$, then $xy \in R_{d+e}$ for all d, e. The elements of R_d are called the **homogeneous elements of degree** d. An element is **homogeneous** if it belongs to some R_d.

Example. The polynomial ring $R = K[X_1, \dots, X_n]$ is graded, with R_d being the polynomials only involving monomials $X_1^{d_1} X_2^{d_2} \dots X_n^{d_n}$ of degree $d_1 + d_2 + \dots + d_n = d$.

The free algebra $R = \langle X_1, \dots, X_n \rangle$ is graded, with R_d being the linear combinations of words of length d.

Definition. Given a K-module V, the **tensor algebra** is

$$T(V) = \bigoplus_{d=0}^{\infty} T^d(V).$$

It becomes a graded algebra with multiplication given by the mapping $T^d(V) \otimes T^e(V) \mapsto T^{d+e}(V)$.

Note that we can write the product of elements $v, v' \in V$ as either $v \otimes v'$ or as vv'.

Properties (of tensor algebras)**.**

(i) If R is a K-algebra and $\theta : V \to R$ is a K-module homomorphism, then there is a unique K-algebra homomorphism $\tilde{\theta} : T(V) \to R$ with $\tilde{\theta}(v) = \theta(v)$ for $v \in V$.

By the property (i) of tensor powers, for each d there is a K-module homomorphism

$$T^d(V) \to R, \quad v_1 \otimes \dots \otimes v_d \mapsto \theta(v_1) \dots \theta(v_d).$$

Combining them, using

$$\mathrm{Hom}_K(T(V), R) = \mathrm{Hom}_K(\bigoplus_{d=0}^{\infty} T^d(V), R) = \prod_{d=0}^{\infty} \mathrm{Hom}_K(T^d(V), R)$$

we get a K-module map $T(V) \to R$, and by construction it is an algebra homomorphism.

(ii) If $\theta : V \to W$ is a K-module homomorphism, then there is a unique K-algebra homomorphism $T(\theta) : T(V) \to T(W)$ with $T(\theta)(v) = \theta(v)$ for $v \in V$.

This follows from (i) with the composition $V \to W \to T(W)$, or from property (ii) of tensor powers with the mappings $T^d(\theta) : T^d(V) \to T^d(W)$.

(iii) If V has basis $(b_1, \dots, b_n)$, then there is an algebra isomorphism

$$K\langle X_1, \dots X_n\rangle \to T(V)$$

sending X_i corresponding to b_i.

The free algebra is the free K-module with basis the words $X_{i_1} \dots X_{i_d}$ with $d \geq 0$ and $i_1, \dots, i_d \in \{1, \dots, n\}$. We define the map by sending this basis element to $b_{i1} \otimes \dots \otimes b_{i_d}$. The elements of this form for fixed d give a basis of $T^d(V)$, so allowing d to vary, we get a basis of $T(V)$. Thus the mapping is an isomorphism of K-modules. Clearly it is also an algebra isomorphism.

Definition. Let V be a K-module. A **quadratic form** $q : V \to K$ is a mapping with $q(\lambda v) = \lambda^2 q(v)$ for all $\lambda \in K$ and $v \in V$ and such that the mapping

$$b : V \times V \to K, \quad b(u, v) = q(u + v) - q(u) - q(v)$$

is bilinear. If K is a field with $\operatorname{char} K \neq 2$, then this agrees with the definition in Linear Algebra II §11.1, since in this case $q(v) = \frac{1}{2}b(v, v)$.

If $q : V \to K$ is a quadratic form, the **Clifford algebra** is

$$C(V, q) = T(V)/(v^2 - q(v)1 : v \in V).$$

It is not in general graded.

If V has basis $(b_1, \dots, b_n)$ and $a = (a_1, \dots, a_n) \in K^n$, then there is a quadratic form

$$q_a : V \to K, \quad q_a(x_1 b_1 + \dots + x_n b_n) = a_1 x_1^2 + \dots + a_n x_n^2$$

for $x_1, \dots, x_n \in K$.

If K is a field with $\operatorname{char} K \neq 2$ then by Linear Algebra II §11.1, any f.d. vector space with a quadratic form has a basis with respect to which the quadratic form is equal to q_a, for some a. Over $\mathbb{R}$, we may assume all $a_i \in \{0, 1, -1\}$, and over $\mathbb{C}$ we may assume that all $a_i \in \{0, 1\}$.

We shall concentrate mainly on $C(V, q_a)$ where V has basis $(b_1, \dots, b_n)$ and $a \in K^n$. In particular we can take $V = K^n$ and $(b_1, \dots, b_n)$ the standard basis.

Properties (of Clifford algebras)**.**

(i) Suppose V and W are equipped with quadratic forms q and p, and suppose that $\theta : V \to W$ is a K-module homomorphism with $p(\theta(v)) = q(v)$ for all $v \in V$. Then there is a unique algebra homomorphism $C(\theta) : C(V, q) \to C(W, p)$ with $C(\theta)(v) = \theta(v)$ for $v \in V$. If θ is an isomorphism of modules, then $C(\theta)$ is an isomorphism of algebras.

Proof. The map θ induces a homomorphism of algebras

$$T(V) \xrightarrow{T(\theta)} T(W) \xrightarrow{\text{canonical map}} C(W, q)$$

For for $v \in V$ we have

$$T(\theta)(v^2 - q(v)1) = T(\theta)(v \otimes v - q(v)1) = \theta(v) \otimes \theta(v) - q(v)1$$
$$= \theta(v)^2 - p(\theta(v))1 \in (w^2 - p(w)1 : w \in W).$$

Thus this homomorphism induces a homomorphism $C(\theta)$. Clearly it is unique. Also, if θ has inverse ϕ, then $C(\phi)$ is an inverse for $C(\theta)$.

(ii) If V has basis $(b_1, \dots, b_n)$, then in $C(V, q_a)$ we have $b_i^2 = a_i 1$ and $b_i b_j = -b_j b_i = 0$ for $i \neq j$.

Namely, consider $v^2 - q_a(v)1$ for the elements $v = b_i$ and $v = b_i + b_j$.

(iii) We have an isomorphism $K\langle X_1, \dots X_n\rangle / I_a \to C(V, q_a)$ sending X_i to b_i, where I_a is the ideal generated by the elements $X_i^2 - a_i 1$ and $X_i X_j + X_j X_i$ for $i \neq j$.

Proof. We have an isomorphism $\phi : K\langle X_1, \dots, X_n\rangle \to T(V)$.

We want to show that $\phi(I_a) = (v^2 - q_a(v)1 : v \in V)$. By (ii), the generators of I_a are sent to 0 in R, so $\phi(I_a) \subseteq (v^2 - q_a(v)1 : v \in V)$. For the reverse inclusion, note that if $v = \sum_{i=1}^n x_i b_i$ with $x_i \in K$, then

$$v^2 - q_a(v)1 = (\sum_{i=1}^n x_i b_i)^2 - \sum_{i=1}^n a_i x_i^2 = \sum_{i=1}^n x_i^2 (b_i^2 - a_i 1) + \sum_{i<j} x_i x_j (b_i b_j + b_j b_i)$$
$$= \phi\left(\sum_{i=1}^n x_i^2 (X_i^2 - a_i 1) + \sum_{i<j} x_i x_j (X_i X_j + X_j X_i)\right) \in \phi(I_a).$$

Theorem (1)**.** *Let V have basis $(b_1, \dots, b_n)$ and let $a \in K^n$. Then $R = C(V, q_a)$ has K-basis the elements b_I for I a subset of $\{1, \dots, n\}$, defined by*

$$b_I = b_{i_1} b_{i_2} \dots b_{i_d}$$

where $I = \{i_1 < i_2 < \dots < i_d\}$.

Proof. R is generated as a K-module by the products of the b_i, and because of (ii) we can write any product as a linear combination of products of the form b_I.

We prove that the b_I are a basis by induction on n. Let V' be the K-submodule of V generated by $b_1, \dots, b_{n-1}$ and let $a' = (a_1, \dots, a_{n-1})$. By induction $R' = C(V', q_{a'})$ has basis the b_J with $J \subseteq \{1, \dots, n-1\}$.

We get a homomorphism $\theta : K\langle X_1, \dots, X_n\rangle \to M_2(R')$ with

$$\theta(X_i) = \begin{pmatrix} b_i & 0 \\ 0 & -b_i \end{pmatrix} \quad (i < n), \quad \theta(X_n) = \begin{pmatrix} 0 & a_n \\ 1 & 0 \end{pmatrix}.$$

This map sends the generators of I_a to 0, so it induces a homomorphism

$$\theta : R \to M_2(R').$$

Any relation between the b_I in R can be written in the form

$$\sum_J \lambda_J b_J + \mu_J b_J b_n = 0$$

with $\lambda_J, \mu_J \in K$. Applying θ, and using that $b_n b_J = (-1)^{|J|} b_J b_n$ we get

$$\begin{pmatrix} \sum_J \lambda_J b_J & \sum_J \mu_J a_n b_J \\ \sum_J (-1)^{|J|} \mu_J b_J & \sum_J \lambda_J (-1)^{|J|} b_J \end{pmatrix} = 0.$$

Thus $\sum_J \lambda_J b_J$ and $\sum_J (-1)^{|J|} \mu_J b_J$ are zero in R'. Thus by induction the λ_J and μ_J are all zero. □

Theorem (2)**.** *Suppose K is a field with $\operatorname{char} K \neq 2$, V has basis $(b_1, \dots, b_n)$ and $a = (a_1, \dots, a_n) \in K^n$. If all $a_i \neq 0$ (or equivalently q_a is 'non-degenerate'), then $R = C(V, q_a)$ is a semisimple algebra.*

Proof. Recall that R is semisimple if and only if every R-module M is semisimple, that is, every submodule N of M has a complement. It is equivalent that there is an R-module homomorphism $f : M \to N$ with $f(m) = m$ for all $m \in N$. Namely, if there is a complement C, we can take f to be the projection onto N. Conversely, if there is f, then $\operatorname{Ker}(f)$ is a complement, since clearly $N \cap \operatorname{Ker}(f) = 0$ and if $m \in M$, then $m = f(m) + (m - f(m)) \in N + \operatorname{Ker}(f)$ since $f(m - f(m)) = f(m) - f(m) = 0$.

We prove the theorem by induction on n.

The case $n = 0$ is clear, since in this case $R = K$, which is a field, so semisimple.

Thus suppose $n > 0$. By the last theorem, the subalgebra R' of R spanned by the b_J with $J \subseteq \{1, \dots, n-1\}$ is isomorphic to the Clifford algebra $C(V', q_{a'})$ where V' is the subspace of V spanned by $b_1, \dots, b_{n-1}$ and $a' = (a_1, \dots, a_{n-1}$. Thus by induction R' is semisimple.

Let M be an R-module and N an R-submodule. We can consider M as an R'-module, and N is a submodule, so there is an R'-module map $f' : M \to N$ with $f'(m) = m$ for $m \in N$. Define $f : M \to N$ by

$$f(m) = \frac{1}{2} f'(m) + \frac{1}{2a_n} b_n f'(b_n m).$$

Then $f(m) = m$ for $m \in N$. Also $f(b_i m) = b_i f(m)$ for $i < n$ and

$$f(b_n m) = \frac{1}{2} f'(b_n m) + \frac{1}{2a_n} b_n f'(b_n^2 m) = b_n f(m),$$

so f is an R-module map. Thus N has an R-module complement in M. Thus any R-module M is semisimple. □

Examples. (1) $C(\mathbb{R}^1, q_{(-1)}) \cong \mathbb{C}$. Since $i^2 = -1$ there is a homomorphism $C(\mathbb{R}^1, q_{(-1)}) \to \mathbb{C}$ of $\mathbb{R}$-algebras with $b_1 \mapsto i$. It sends the basis $1, b_1$ to the basis $1, i$, so it is an isomorphism.

(2) $C(\mathbb{R}^1, q_{(1)}) \cong \mathbb{R} \times \mathbb{R}$ with $b_1 \mapsto (1, -1)$.

(3) $C(\mathbb{R}^2, q_{(-1,-1)}) \cong \mathbb{H}$ with $b_1 \mapsto i$ and $b_2 \mapsto j$ (so $b_1 b_2 \mapsto ij = k$).

(4) $C(\mathbb{R}^2, q_{(1,1)}) \cong M_2(\mathbb{R})$ with $b_1 \mapsto \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ and $b_2 \mapsto \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

(5) The Clifford algebra for 3-dimensional Euclidean space is $R = C(\mathbb{R}^3, q_{(1,1,1)})$.

Any rotation θ of $\mathbb{R}^3$ about the origin preserves distance from the origin, so preserves q. Thus it induces an isomorphism $C(\theta) : R \to R$.

There is an isomorphism $R \to M_2(\mathbb{C})$ sending the b_i to the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

This is a homomorphism since the Pauli matrices have square 1 and anticommute. Now the basis $1, b_1, b_2, b_3, b_1b_2, b_1b_3, b_2b_3, b_1b_2b_3$ of $C(\mathbb{R}^3, q_{(1,1,1)})$ gets sent to

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}$$

which is a basis of $M_2(\mathbb{C})$ as a vector space over $\mathbb{R}$. Thus this homomorphism is an isomorphism.

Pauli (1927) used his matrices to formulate a version of the Schrödinger equation for spin $1/2$ particles.

Why does $\mathbb{H}$ appear when studying rotations? It is isomorphic to the subalgebra of R with basis the words of even length: 1, $b_1b_2 \leftrightarrow i$, $b_2b_3 \leftrightarrow j$ and $b_1b_3 \leftrightarrow k$.

(6) In special relativity one uses Minkowski space, which is $\mathbb{R}^4$ with the quadratic form

$$q(t, x, y, z) = t^2 - x^2 - y^2 - z^2$$

where (x, y, z) is position, t is time, and the units are chosen so that the speed of light is 1.

The relevant Clifford algebra is $C(\mathbb{R}^4, q_{(1,-1,-1,-1)})$. It is sometimes called the Space-Time Algebra.

Dirac (1928) used matrices in $M_4(\mathbb{C})$ which give a homomorphism from the Clfford algebra to $M_4(\mathbb{C})$ to formulate a relativistic version of Pauli's equation. Actually they give an isomorphism $C(\mathbb{C}^4, q_{(1,-1,-1,-1)}) \cong M_4(\mathbb{C})$.

Definition. Let V be a K-module. The **exterior algebra** or **Grassmann algebra** $\Lambda(V)$ on V is the Clifford algebra given by V with the zero quadratic form

$$\Lambda(V) = C(V, 0) = T(V)/(v \otimes v : v \in V).$$

We write the product in $\Lambda(V)$ as $x \wedge y$.

If $\theta : V \to W$ is a K-module map, then as for Clifford algebras, $T(\theta)$ induces an algebra homomorphism

$$\Lambda(\theta) : \Lambda(V) \to \Lambda(W), \quad (\Lambda\theta)(v_1 \wedge \cdots \wedge v_d) = \theta(v_1) \wedge \cdots \wedge \theta(v_d)$$

for $v_i \in V$.

Lemma. *In $\Lambda(V)$ we have the following identities*

(i) $v \wedge v = 0$ for all $v \in V$.

(ii) $v \wedge v' = -v' \wedge v$ for all $v, v' \in V$

(iii) If π is a permutation of $\{1, \ldots, d\}$ and $\epsilon(\pi)$ denotes its sign, then

$$v_{\pi(1)} \wedge \cdots \wedge v_{\pi(d)} = \epsilon(\pi) v_1 \wedge \cdots \wedge v_d$$

for $v_i \in V$.

(iv) $v_1 \wedge \cdots \wedge v_d = 0$ for $v_i \in V$ if two of the v_i are equal.

Proof. (i) This is the definition.

(ii) As for Clifford algebras, consider $(v + v') \wedge (v + v')$.

(iii) Part (ii) gives the result in case π is the transposition of the form $(i\ i+1)$. Now any permutation can be written as a composition of such transpositions.

(iv) By (iii) we can permute the v_i to make consecutive ones equal. Then the result follows from (i). □

Properties (of exterior algebras)**.**

(i) In $T(V)$ we have

$$(v \otimes v : v \in V) = J := \bigoplus_{d=0}^{\infty} J_d$$

where J_d is the K-submodule of $T^d(V)$ generated by elements of the form $v_1 \otimes \cdots \otimes v_d$ with $v_i \in V$ and two of the v_i equal.

Namely, J is clearly an ideal in $T(V)$ and it contains $v \otimes v$ for $v \in V$, so $(v \otimes v : v \in V) \subseteq J$. By the lemma J is contained in the kernel of the homomorphism $T(V) \to \Lambda(V)$, so $J \subseteq (v \otimes v : v \in V)$.

(ii) $\Lambda(V)$ is a graded algebra, with decomposition

$$\Lambda(V) = \bigoplus_{d=0}^{\infty} \Lambda^d(V)$$

where $\Lambda^d(V)$ is generated as a K-submodule by the elements $v_1 \wedge \cdots \wedge v_d$ with $v_i \in V$. The homogeneous pieces $\Lambda^d(V)$ are called the **exterior powers** of V. The map $T^d(V) \to \Lambda(V)$ gives an isomorphism of K-modules $T^d(V)/J_d \cong \Lambda^d(V)$.

Proof. Let $p : T(V) \to \Lambda(V)$ be the canonical map and let $\Lambda^d(V) = p(T^d(V))$. Since $T^d(V)$ is generated as a K-module by the elements, $v_1 \otimes \cdots \otimes v_d$, it follows that $\Lambda^d(V)$ is generated as a K-module by the elements $v_1 \wedge \cdots \wedge v_d$.

Since p is surjective and $T(V) = \sum_{d=0}^{\infty} T^d(V)$, we have $\Lambda(V) = \sum_{d=0}^{\infty} \Lambda^d(V)$.

Say $x_d \in \Lambda^d(V)$, all but finitely many zero, and $\sum x_d = 0$. Then $x_d = p(y_d)$ for some $y_d \in T^d(V)$, all but finitely many zero. Then $p(\sum y_d) = 0$. Thus by (i) we have

$$\sum y_d \in \operatorname{Ker}(p) = J$$

by (i). Now $J = \bigoplus_{d=0}^{\infty} J_d$ so $\sum y_d = \sum j_d$ for elements $j_d \in J_d$, all but finitely many zero. But this is an equality in $T(V) = \bigoplus_{d=0}^{\infty} T^d(V)$, so $y_d = j_d$ for all d. Thus $y_d \in J$, so $x_d = p(y_d) = 0$. Thus $\Lambda(V) = \bigoplus_{d=0}^{\infty} \Lambda^d(V)$.

(iii) The mapping

$$V^d \to \Lambda^d(V), \quad (v_1, \ldots, v_d) \mapsto v_1 \wedge \cdots \wedge v_d$$

is K-multilinear and alternating, meaning that $v_1 \wedge \cdots \wedge v_d = 0$ if two of the v_i are equal. Moreover it is universal for this property. That is, if $f : V^d \to M$ is an alternating K-multilinear map to a K-module M, then there is a unique K-module map $\alpha : \Lambda^d(V) \to M$ with $f(v_1, \ldots, v_d) = \alpha(v_1 \wedge \cdots \wedge v_d)$.

This follows from the universal property of $T^d(V)$ and the fact that $\Lambda^d(V) \cong T^d(V)/J_d$.

(iv) If $\theta : V \to W$ is a K-module map, then $\Lambda(\theta)$ restricts to give a K-module homomorphism

$$\Lambda^d(\theta) : \Lambda^d(V) \to \Lambda^d(W), \quad \Lambda^d(\theta)(v_1 \wedge \cdots \wedge v_d) = \theta(v_1) \wedge \cdots \wedge \theta(v_d).$$

(v) If V has basis $(b_1, \ldots, b_n)$, then $\Lambda(V)$ has basis the elements

$$b_I = b_{i_1} \wedge b_{i_2} \wedge \cdots \wedge b_{i_d}, \quad I = \{i_1 < \cdots < i_d\} \subseteq \{1, \ldots, n\}, d \geq 0.$$

Thus $\Lambda^d(V)$ has basis the elements b_I with I a subset with d elements. This basis has $\binom{n}{d}$ elements.

Theorem. *If V is a free K-module having a basis with n elements, then $\Lambda^n(V) \cong K$ and $\Lambda^d(V) = 0$ for $d > n$. Moreover any other free basis for V has n elements. We call n the **rank** of V.*

Proof. The first statement follows from property (v), since $\binom{n}{n} = 1$. Since V has a finite basis, it is finitely generated. Then by §2.3 Theorem (iii), any other basis for V must be finite. But if it has m elements, then $n = m = \max\{d : \Lambda^d(V) \neq 0\}$. □

Definition. If V is a free K-module of rank n and $\theta \in \mathrm{End}_K(V)$, we define the **determinant** $\det(\theta) \in K$ as follows. The map $\Lambda^n(\theta) : \Lambda^n(V) \to \Lambda^n(V)$ is an endomorphism of a rank 1 free K-module, so is multiplication by some element of K. This is $\det(\theta)$. Thus

$$(\Lambda^n(\theta))(x) = \det(\theta)x$$

for all $x \in \Lambda^n(V)$.

Remark. The definitions and results in Linear Algebra I about determinants of matrices over fields, extend to commutative rings, and agree with the definition here. For example we have the following.

Theorem (Leibnitz Formula)**.** *Suppose that V has basis $(b_1, \dots, b_n)$ and that $\theta : V \to V$ is a linear map with matrix $A = (a_{ij}) \in M_n(K)$ with respect to this basis, so that*

$$\theta(b_j) = \sum_{i=1}^{n} a_{ij} b_i.$$

Then

$$\det(\theta) = \sum_{\sigma \in S_n}^{n} \epsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}$$

(which is the Leibnitz formula for $\det(A)$ *in Linear Algebra I).*

Proof. $\Lambda^n(V)$ has basis $b_1 \wedge \dots \wedge b_n$. Moreover

$$\Lambda^n(\theta)(b_1 \wedge \dots \wedge b_n) = \theta(b_1) \wedge \dots \theta(b_n)$$

$$= (\sum_{i_1=1}^{n} a_{i_1,1} b_{i_1}) \wedge (\sum_{i_2=1}^{n} a_{i_2,1} b_{i_2}) \wedge \dots \wedge (\sum_{i_n=1}^{n} a_{i_n,n} b_{i_n})$$

$$\sum_{i_1,\dots,i_n=1}^{n} a_{i_1,1} a_{i_2,2} \dots a_{i_n,n} b_{i_1} \wedge b_{i_2} \wedge \dots \wedge b_{i_n}.$$

Since any term with two i_j's equal is zero, we can write this as a sum over permutations π

$$\sum_{\pi \in S_n} a_{\pi(1),1} a_{\pi(2),2} \dots a_{\pi(n),n} b_{\pi(1)} \wedge b_{\pi(2)} \wedge \dots \wedge b_{\pi(n)}$$

$$\sum_{\pi \in S_n} \epsilon(\pi) a_{\pi(1),1} a_{\pi(2),2} \dots a_{\pi(n),n} b_1 \wedge b_2 \wedge \dots \wedge b_n$$

so

$$\det(\theta) = \sum_{\pi \in S_n} \epsilon(\pi) a_{\pi(1),1} a_{\pi(2),2} \dots a_{\pi(n),n}.$$

Since K is commutative, the product is equal to $a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}$ where $\sigma = \pi^{-1}$. Since $\epsilon(\pi) = \epsilon(\pi^{-1})$, we obtain

$$\det(\theta) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}.$$

□

3.4 More applications of tensor products

Definition. Suppose $R \to S$ is a ring homomorphism. Then S is naturally an S-R-bimodule. We sometimes denote it by ${}_S S_R$. If Y is an R-module, then $S \otimes_R Y$ is an S-module. We call it the module obtained from Y by **inducing from R to S**. It is sometimes denoted $\mathrm{Ind}^S Y$ or $\mathrm{Ind}_R^S Y$.

Properties. (i) If Z is a left S-module, then by restriction we have an R-module ${}_R Z$, and

$$\mathrm{Hom}_S(S \otimes_R Y, Z) \cong \mathrm{Hom}_R(Y, {}_R Z).$$

This says that induction and restriction form an adjoint pair of functors. This follows from Hom-Tensor adjointness, since if we consider S as an S-R-bimodule, then $\mathrm{Hom}_S(S, Z) \cong {}_R Z$.

(ii) If Y is a free R-module with basis (b_i), then $S \otimes_R Y$ is a free S-module with basis $(1 \otimes b_i)$.

(iii) In particular if L/K is a field extension and V is a K-vector space, then $L \otimes_K V$ is an L-vector space and $\dim_L(L \otimes_K V) = \dim_K V$.

Definition. If R and S are K-algebras, then $R \otimes_K S$ is naturally a K-algebra with a product satisfying

$$(r \otimes s)(r' \otimes s') = (rr') \otimes (ss').$$

Note that there are algebra homomorphisms

$$R \to R \otimes_K S, \quad r \mapsto r \otimes 1,$$

$$S \to R \otimes_K S, \quad s \mapsto 1 \otimes s$$

and their images commute, since $(r \otimes 1)(1 \otimes s) = r \otimes s = (1 \otimes s)(r \otimes 1)$.

As a special case, if R is a commutative K-algebra, then $R \otimes_K S$ is an R-algebra.

Examples. (i) $R \otimes_K M_n(K) \cong M_n(R)$ via the map sending $r \otimes A$ with $A = (A_{ij})$ to the matrix with (i, j) entries $ra_{ij} \in R$.

(ii) $M_n(K) \otimes_K M_m(K) \cong M_{nm}(K)$. Here we index the rows and columns of $M_{nm}(K)$ by the set $\{1, \dots, n\} \times \{1, \dots, m\}$ and we send $A \otimes B$ to the matrix with $(i,j)(k,\ell)$ entry $a_{ik}b_{j\ell}$.

Let E^{ij} denote the matrix which is 1 at position (i,j) and zero elsewhere. This map sends $E^{ij} \otimes E^{pq}$ to $E^{(i,p)(j,q)}$. It preserves the multiplication since

$$(E^{ij} \otimes E^{pq})(E^{i'j'} \otimes E^{p'q'}) = E^{ij}E^{i'j'} \otimes E^{pq}E^{p'q'} = \delta_{j,i'}\delta_{q,p'}E^{ij'} \otimes E^{pq'}$$

which matches the product

$$E^{(i,p)(j,q)}E^{(i',p')(j',q')} = \delta_{(j,q)(i',p)}E^{(i,p)(j',q)}.$$

(iii) If R is a commutative K-algebra, then $R \otimes_K K[X_1, \dots, X_n] \cong R[X_1, \dots, X_n]$. In particular

$$K[X_1, \dots, X_n] \otimes_K K[Y_1, \dots, Y_m] \cong K[X_1, \dots, X_n][Y_1, \dots, Y_m]$$

$$\cong K[X_1, \dots, X_n, Y_1, \dots, Y_m].$$

Namely, we have a ring homomorphism

$$R \otimes_K K[X_1, \dots, X_n] \to R[X_1, \dots, X_n]$$

sending $r \otimes p(X_1, \dots, X_n)$ to $rp(X_1, \dots, X_n)$, where we need to apply the ring homomorphism

$$K \to R, \quad \lambda \mapsto \lambda 1_R$$

to the coefficients of $p(X_1, \dots, X_n)$.

It is an isomorphism since $K[X_1, \dots, X_n]$ is a free K-module on the monomials $X_1^{m_1} \dots X_n^{m_n}$, so $R \otimes_K K[X_1, \dots, X_n]$ is a free R-module on the elements $1 \otimes X_1^{m_1} \dots X_n^{m_n}$, and these get sent to the monomials $X_1^{m_1} \dots X_n^{m_n}$, which are a free R-basis of $R[X_1, \dots, X_n]$.

(iv) Similarly, if R is a commutative K-algebra, then there is an isomorphism of R-algebras

$$\theta : R \otimes_K K\langle X_1, \dots, X_n \rangle \to R\langle X_1, \dots, X_n \rangle$$

sending $r \otimes a$ to ra if $r \in R$ and $a \in K\langle X_1, \dots, X_n \rangle$.

Lemma. *Suppose R, S are K-algebras and I is an ideal in S generated by a subset $H \subseteq S$. Let $i : I \to S$ be the inclusion, and consider the map $\mathrm{Id} \otimes i : R \otimes_K I \to R \otimes_K S$. Then $\mathrm{Im}(\mathrm{Id} \otimes i)$ is the ideal in $R \otimes_K S$ generated by the set $\{1 \otimes h : h \in H\}$, and*

$$R \otimes_K (S/I) \cong (R \otimes_K S)/\,\mathrm{Im}(\mathrm{Id} \otimes i).$$

Proof. Let J be the ideal in $R \otimes_K S$ generated by $\{1 \otimes h : h \in H\}$. Now $\mathrm{Im}(\mathrm{Id} \otimes i)$ is an ideal, since if $r, r', r'' \in R$, $s', s'' \in S$ and $x \in I$ then

$$(r' \otimes s')(r \otimes x)(r'' \otimes s'') = r'rr'' \otimes s'xs'' \in \mathrm{Im}(\mathrm{Id} \otimes i).$$

Moreover this ideal contains the elements $1 \otimes h$, so $J \subseteq \mathrm{Im}(\mathrm{Id} \otimes i)$.

Conversely if $s, s' \in S$, then

$$1 \otimes shs' = (1 \otimes s)(1 \otimes h)(1 \otimes s') \in J.$$

It follows that $1 \otimes x \in J$ for any $x \in I$. Then

$$r \otimes x = (r \otimes 1)(1 \otimes x) \in J$$

for all $r \in R$, $x \in I$. Thus $\mathrm{Im}(\mathrm{Id} \otimes i) \subseteq J$.

Now the exact sequence

$$0 \to I \xrightarrow{i} S \to S/I \to 0$$

stays exact on the right on tensoring, so gives

$$R \otimes_K I \xrightarrow{\mathrm{Id} \otimes i} R \otimes_K S \to R \otimes_K (S/I) \to 0$$

so $R \otimes_K (S/I) \cong (R \otimes_K S)/\mathrm{Im}(\mathrm{Id} \otimes i)$. □

Examples. (a) We have

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C} \otimes_{\mathbb{R}} (\mathbb{R}[X]/(X^2+1)) \cong (\mathbb{C} \otimes_{\mathbb{R}} \mathbb{R}[X])/(1 \otimes (X^2+1))$$

$$\cong \mathbb{C}[X]/(X^2+1) = \mathbb{C}[X]/((X+i)(X-i)).$$

Now in $\mathbb{C}[X]$ we have $(X+i) \cap (X-i) = ((X+i)(X-i))$ and $(X+i)+(X-i) = \mathbb{C}[X]$, since $X+i$ and $X-i$ are coprime. Explicitly $(X+i)+(X-i)$ contains $\frac{i}{2}(X-i) - \frac{i}{2}(X-i) = 1$. Thus by the Chinese Remainder Theorem for rings in §3.3 of Algebra I, we have

$$\mathbb{C}[X]/((X+i)(X-i)) \cong \mathbb{C}[X]/(X+i) \times \mathbb{C}[X]/(X-i).$$

Now there are ring isomorphisms $\mathbb{C}(X]/(X \pm i) \to \mathbb{C}$, $p(X) \mapsto p(\mp i)$. Thus

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{C} \cong \mathbb{C} \times \mathbb{C}.$$

(b) If R is a commutative K-algebra and $a \in K^n$, then

$$R \otimes_K C(K^n, q_a) \cong C(R^n, q_a),$$

where the second copy of a really means its image in R^n, since

$$R \otimes_K C(K^n, q_a) \cong R \otimes_K (K\langle b_1, \dots, b_n\rangle/(b_i^2 - a_i 1, b_i b_j + b_j b_i))$$

$$\cong (R \otimes_K K\langle b_1, \dots, b_n\rangle)/(1 \otimes (b_i^2 - a_i 1), 1 \otimes (b_i b_j + b_j b_i))$$
$$\cong R\langle b_1, \dots, b_n\rangle/(b_i^2 - a_i 1_R, b_i b_j + b_j b_i).$$

(c) $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{H} \cong M_2(\mathbb{C})$. We have

$$\mathbb{C} \otimes_{\mathbb{R}} \mathbb{H} \cong \mathbb{C} \otimes_{\mathbb{R}} C(\mathbb{R}^2, q_{(-1,-1)}) \cong C(\mathbb{C}^2, q_{(-1,-1)}).$$

Over $\mathbb{C}$ we can multiply the basis elements by i to get

$$\cong C(\mathbb{C}^2, q_{(1,1)}) \cong \mathbb{C} \otimes_{\mathbb{R}} C(\mathbb{R}^2, q_{(1,1)}) \cong \mathbb{C} \otimes_{\mathbb{R}} M_2(\mathbb{R}) \cong M_2(\mathbb{C}).$$

(d) $\mathbb{H} \otimes_{\mathbb{R}} \mathbb{H} \cong M_4(\mathbb{R})$.

By the theorem below we have

$$C(\mathbb{R}^4, q_{(1,1,-1,-1)}) \cong C(\mathbb{R}^2, q_{(1,1)}) \otimes_{\mathbb{R}} C(\mathbb{R}^2, q_{(1,1)}) \cong M_2(\mathbb{R}) \otimes_{\mathbb{R}} M_2(\mathbb{R}) \cong M_4(\mathbb{R}).$$

But by permuting the basis elements of $\mathbb{R}^4$, this is isomorphic to

$$C(\mathbb{R}^4, q_{(-1,-1,1,1)}) \cong C(\mathbb{R}^2, q_{(-1,-1)}) \otimes_{\mathbb{R}} C(\mathbb{R}^2, q_{(-1,-1)}) \cong \mathbb{H} \otimes_{\mathbb{R}} \mathbb{H}$$

where we have used the theorem below again.

Theorem. *If $a = (a_1, a_2) \in K^2$ with a_1, a_2 invertible in K and $a' = (a'_1, \dots, a'_n) \in K^n$, then*

$$C(K^2, q_a) \otimes_K C(K^n, q_{a'}) \cong C(K^{n+2}, q_{a''})$$

where $a'' = (a_1, a_2, -a_1 a_2 a'_1, \dots, -a_1 a_2 a'_n)$.

Proof. Let b_1, b_2 be the generators of $C(K^2, q_a)$, let $b'_1, \dots, b'_n$ be the generators of $C(K^n, q_{a'})$, and let $b''_1, \dots, b''_{n+2}$ be the generators of $C(K^{n+2}, q_{a''})$. The elements

$$c_1 = b_1 \otimes 1,\ c_2 = b_2 \otimes 1,\ c_3 = b_1 b_2 \otimes b'_1,\ \dots\ , c_{n+2} = b_1 b_2 \otimes b'_n$$

of $C(K^2, q_a) \otimes_K C(K^n, q_{a'})$ satisfy the relations for $C(K^{n+2}, q_{a''})$. For example

$$(c_3)^2 = (b_1 b_2 \otimes b'_1)(b_1 b_2 \otimes b'_1) = b_1 b_2 b_1 b_2 \otimes b'_1 b'_1 = -a_1 a_2 1 \otimes a'_1 1 = a''_3 (1 \otimes 1).$$

Thus we get an algebra homomorphism $C(K^{n+2}, q_{a''}) \to C(K^2, q_a) \otimes_K C(K^n, q_{a'})$ sending b''_i to c_i.

Now the usual K-bases b_I ($I \subseteq \{1, 2\}$) of $C(K^2, q_a)$ and b'_J ($J \subseteq \{1, \dots, n\}$) of $C(K^n, q_{a'})$ give a K-basis $b_I \otimes b'_J$ of the tensor product. The algebra homomorphism sends the usual basis of $C(K^{n+2}, q_{a''})$ to multiples (by invertible elements of K) of the basis elements $b_I \otimes b'_J$. For example $b''_1 b''_3 b''_4$ is sent to

$$c_1 c_3 c_4 = (b_1 \otimes 1)(b_1 b_2 \otimes b'_1)(b_1 b_2 \otimes b'_2) = (-a_1 a_2)(b_1 \otimes b'_1 b'_2).$$

This gives a bijective correspondence between the bases, so the homomorphism is an isomorphism. I omit the details. □

We use tensor products in the uniqueness part of the following theorem.

Theorem. *(Steinitz). Any field K has an algebraic closure, that is, there is an algebraic field extension L/K with L algebraically closed. Moreover the algebraic closure of K is unique up to isomorphism.*

Lemma. *Suppose L/K is an algebraic field extension. If every irreducible polynomial in $K[X]$ splits over L as a product of linear factors, then L is algebraically closed (so an algebraic closure of K).*

Proof. Suppose E/L is a field extension and $\alpha \in E$ is algebraic over L. Then $p(\alpha) = 0$ for some nonzero $p(X) \in L[X]$. Since L is algebraic over K, the coefficients of $p(X)$ all belong to some finite extension F of K. Then α is algebraic over F. Thus $F(\alpha)/F$ is a finite extension. Thus by the Tower Law, $F(\alpha)/K$ is a finite extension. Thus α is algebraic over K. By assumption its minimal polynomial $m_{\alpha/K}(X)$ splits as a product of linear factors over L. Since $m_{\alpha/K}(\alpha) = 0$, we get $\alpha \in L$. □

Proof of the Theorem. Let $\{p_i(X) : i \in I\}$ be the set of monic irreducible polynomials in $K[X]$ and let $n_i = \operatorname{Grad} p_i(X)$. Let $R = K[Y_{ik} : i \in I, 1 \leq k \leq n_i]$.

For $i \in I$,

$$p_i(X) - \prod_{k=1}^{n_i} (X - Y_{ik})$$

is a polynomial in $R[X]$ whose degree in X is $\leq n_i - 1$, so we can write it as

$$\sum_{k=0}^{n_i - 1} r_{ik} X^k$$

with $r_{ik} \in K[Y_{i1}, \dots, Y_{i,n_i}] \subseteq R$. Let $J = (r_{ik} : i \in I, 1 \leq k \leq n_i)$, the ideal in R generated by the r_{ik}. We want to show that $J \neq R$. Suppose $J = R$. Then

$$\sum_{ik} a_{ik} r_{ik} = 1$$

for some $a_{ik} \in R$, all but finitely many zero. Let $I' = \{i \in I : a_{ik} \neq 0 \text{ for some } k\}$, a finite subset of I. Let F/K be a splitting field for the polynomial $\prod_{i \in I'} p_i(X)$. We have a homomorphism

$$f : R \to F$$

sending the Y_{ik} for $i \in I'$ to the roots of $p_i(X)$ in F (in some order), and sending the other Y_{ik} to 0. Consider the induced homomorphism

$$R[X] \to F[X].$$

For $i \in I'$, it sends $p_i(X) - \prod_{k=1}^{n_i} (X - Y_{ik})$ to 0, so $f(r_{ik}) = 0$ for $i \in I'$. Thus f sends $\sum_{ik} a_{ik} r_{ik}$ to 0, which is impossible.

Thus J is a proper ideal in R, so it is contained in a maximal ideal m. Since K is a field, the ring homomorphism $K \to L = R/m$ must be injective, so we can consider L as a field extension of K. Then in $L[X]$ we have the factorization $p_i(X) = \prod_{k=1}^{n_i}(X - \overline{Y}_{ik})$ so each irreducible polynomial $p_i(X)$ splits.

Now L/K is an algebraic extension, since any element belongs to some extension of K obtained by adjoining finitely many of the $\overline{Y}_{ik}$, and they are algebraic over K, since they are roots of $p_i(X)$. Thus by the lemma, L is an algebraic closure of K.

For uniqueness, suppose that L/K and L'/K are algebraic closures. Then $L \otimes_K L'$ is a non-zero commutative ring, so has a maximal ideal m'. The factor ring $E = (L \otimes_K L')/m'$ is a field and has homomorphisms from L and L' such that the homomorphisms from K are equal.

We consider the field extension E/L. If $a \in L'$, then since L'/K is algebraic, the element $\overline{1 \otimes a}$ of E is algebraic over L. Thus since L is algebraically closed, $\overline{1 \otimes a} \in L$. It follows that $E = L$. More precisely, the map $L \to E$ is an isomorphism.

Similarly $L' \to E$ is an isomorphism. Thus $L \cong L'$. □

Examples. (1) The algebraic closure of $\mathbb{Q}$ is

$$L = \{a \in \mathbb{C} : a \text{ is algebraic over } \mathbb{Q}\}.$$

If a, b are algebraic over $\mathbb{Q}$, so are $a + b$, ab, $1/a$, so this is a subfield of $\mathbb{C}$, and it is algebraic over $\mathbb{Q}$.

Now any irreducible polynomial in $\mathbb{Q}[X]$ splits into linear factors over $\mathbb{C}$, and the roots are all in L, so it splits into linear factors over L. Thus by the lemma, L is a algebraic closure of $\mathbb{Q}$.

(2) Let p be a prime number. Recall that for each power q of p there is a unique field $\mathbb{F}_q$ with q elements. Moreover $\mathbb{F}_q \subseteq \mathbb{F}_{q'}$ if and only if q' is a power of q. Thus we have inclusions

$$\mathbb{F}_p \subseteq \mathbb{F}_{p^{2!}} \subseteq \mathbb{F}_{p^{3!}} \subseteq \dots$$

where $n! = n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$ is the factorial of n. The algebraic closure of $\mathbb{F}_p$ is the union L of these fields.

Namely, every element $a \in L$ is in $\mathbb{F}_q$ for some $q = p^{n!}$, and $[\mathbb{F}_q : \mathbb{F}_p] < \infty$, so a is algebraic over $\mathbb{F}_p$.

Now suppose $f(X)$ is an irreducible polynomial over $\mathbb{F}_p$. Let $E/\mathbb{F}_p$ be a splitting field. Then E is a finite extension of $\mathbb{F}_p$, so for some m we have

$$E \cong \mathbb{F}_{p^m} \subseteq \mathbb{F}_{p^{m!}} \subseteq L.$$

Thus $f(X)$ splits into linear factors over L, so by the lemma, L is an algebraic closure of $\mathbb{F}_p$.

4 Representations of finite groups

4.1 Representations and the group algebra

Let G be a group and let K be a field. We write G in multiplicative notation with neutral element 1.

Definition. A **(linear) representation** of the group G over the field K is given by a K-vector space V together with a group homomorphism

$$\rho : G \to \mathrm{GL}(V)$$

where $\mathrm{GL}(V)$ is the group of invertible linear maps $V \to V$.

We denote the representation by (V, ρ) or V or ρ. The **degree** of the representation is $\dim V$. A **real/complex representation** is one with $K = \mathbb{R}$ or $K = \mathbb{C}$.

We get a category with

$$\mathrm{Hom}((V,\rho),(W,\sigma)) = \{\theta \in \mathrm{Hom}_K(V,W) : \sigma(g)\theta = \theta\rho(g) \text{ for all } g \in G\}.$$

A **(matrix) representation** of G is a group homomorphism

$$A : G \to \mathrm{GL}_n(K)$$

where $\mathrm{GL}_n(K)$ is the group of $n \times n$ invertible matrices.

Two matrix representations $A, B : G \to \mathrm{GL}_n(K)$ of the same degree are said to be **equivalent** if there is an invertible matrix P such that $B(g) = PA(g)P^{-1}$ for all $g \in G$.

Lemma. *(i) If (V, ρ) is a linear representation of degree n and $(v_1, \dots, v_n)$ is a basis of V, then the map*

$$A : G \to \mathrm{GL}_n(K), \quad A(g) = \textit{matrix of } \rho(g) \textit{ with respect to this basis}$$

is a matrix representation.

(ii) If $A : G \to \mathrm{GL}_n(K)$ is a matrix representation, then K^n equipped with the map

$$\rho : G \to \mathrm{GL}(K^n), \quad \rho(g) = \textit{the map } K^n \to K^n \textit{ of left multiplication by } A(g)$$

is a linear representation of degree n.

(iii) These give inverse bijections between the isomorphism classes of representations of degree n and the equivalence classes of matrix representations of degree n.

Proof. Exercise. □

Examples. (1) The **trivial representation** of G is the representation $G \to \mathrm{GL}_1(K)$ with $\rho(g) = 1$ for all $g \in G$.

(2) Let $n > 0$. We denote by C_n a cyclic group of order n written multiplicatively, with generator g. Thus $C_n = \{1, g, g^2, \ldots, g^{n-1}\}$ and $g^n = 1$. If $\epsilon \in K$ is an nth root of 1, we get a representation $\rho : C_n \to \mathrm{GL}_1(K)$ with $\rho(g^r) = \epsilon^r$ for all r.

(3) The **sign representation** of the symmetric group S_n is the representation

$$\epsilon : S_n \to \mathrm{GL}_1(K)$$

where $\epsilon(\pi)$ is the sign of a permutation π.

(4) Suppose $\theta : G \to H$ is a group homomorphism and $\sigma : H \to \mathrm{GL}(V)$ is a representation of H. Then by composition we get a representation of G

$$G \to H \to \mathrm{GL}(V).$$

In particular, if N is a normal subgroup of G and $\sigma : G/N \to \mathrm{GL}(V)$ is a representation of the factor group, we get a representation of of G via

$$G \to G/N \to \mathrm{GL}(V).$$

(5) If L/K is a field extension and V is representation of G over K, then $L \otimes_K V$ is an L-vector space and it becomes a representation of G over L via

$$G \xrightarrow{\rho} \mathrm{GL}(V) \xrightarrow{\theta \mapsto \mathrm{Id} \otimes \theta} \mathrm{GL}(L \otimes_K V)$$

If $(v_1, \ldots, v_n)$ is a K-basis of V, then $(1 \otimes v_1, \ldots, 1 \otimes v_n)$ is an L-basis of $L \otimes_K V$, and the corresponding matrix representations are related by

$$G \to \mathrm{GL}_n(K) \xrightarrow{\text{inclusion}} \mathrm{GL}_n(L).$$

(6) Recall that the dihedral group D_n is the group of symmetries of a regular n-gon in the plane, say with one vertex on the x-axis. We have $D_n = \langle \sigma, \tau \rangle$ where σ is rotation by angle $2\pi/n$, τ is the reflection in the x-axis, $\sigma^n = 1$, $\tau\sigma = \sigma^{-1}\tau$. There is a corresponding natural representation

$$\rho : D_n \to \mathrm{GL}_2(\mathbb{R}), \quad \rho(\sigma) = \begin{pmatrix} \cos(2\pi/n) & -\sin(2\pi/n) \\ \sin(2\pi/n) & \cos(2\pi/n) \end{pmatrix}, \quad \rho(\tau) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

(7) Let G be the group of rotations of a Platonic solid. There is a natural representation

$$\rho : G \to \mathrm{GL}_3(\mathbb{R})$$

sending each rotation to the corresponding rotation matrix.

Remark. Given a group G, we would like to classify all representations of G. Then we would know all ways in which G can occur as a group of symmetries of an object.

Definition. The **group algebra** KG consists of the formal sums

$$\sum_{g\in G} a_g g$$

with coefficients $a_g \in K$, all but finitely many zero. Thus it is a vector space with basis G. It becomes an algebra with the multiplication coming from that in G, that is,

$$(\sum_{g\in G} a_g g)(\sum_{h\in G} b_h h) = \sum_{g,h\in G} a_g b_h gh = \sum_{k\in G} c_k k$$

where

$$c_k = \sum_{\substack{g,h\in G\\ gh=k}} a_g b_h = \sum_{g\in G} a_g b_{g^{-1}k}.$$

Example. Let $C_2 = \langle g\rangle$ with $g^2 = 1$. Then $KC_2 = \{a1 + bg : a, b \in K\}$ with $(a1 + bg)(a'1 + b'g) = (aa' + bb')1 + (ab' + ba')g$.

If $\operatorname{char} K \neq 2$, then $KC_2 \cong K \times K$, via the mapping $a1 + bg \mapsto (a + b, a - b)$.

If $\operatorname{char} K = 2$, then $KC_2 \cong K[X]/(X^2)$, via the mapping $K[X] \to KC_2$, $X \mapsto 1+g$, since $(1 + g)^2 = 2 + 2g = 0$.

Lemma. *The formulas $gv = \rho(g)(v)$ and $(\sum_{g\in G} a_g g)v = \sum_{g\in G} a_g(gv)$ give bijections between*

(a) representations (V, ρ) of G;

(b) actions $G \times V \to V$, $(g, v) \mapsto gv$ of G on a vector space V, which are linear, meaning that for each $g \in G$, the map $v \mapsto gv$ is a linear map; and

(c) KG-modules V.

Moreover homomorphisms of representations correspond to KG-module homomorphisms.

Proof. Recall that an action of G on a set V is a mapping $G \times V \to V$, $(g, v) \mapsto gv$ with $g(g'v) = (gg')v$ and $1v = v$ for all $g, g' \in G$ and $v \in V$. A representation gives a linear action via the formula $gv = \rho(g)(v)$.

Given a linear action, the same formula gives a map $\rho : G \to \operatorname{End}(V)$ with $\rho(gg') = \rho(g)\rho(g')$ and $\rho(1) = \mathrm{Id}_V$. Since $\rho(g^{-1})\rho(g) = \rho(1) = \mathrm{Id}_V = \rho(g)\rho(g^{-1})$ it is a map to $\mathrm{GL}(V)$, so it gives a representation.

Given a linear action, V becomes a KG-module via

$$(\sum_{g\in G} a_g g)v = \sum_{g\in G} a_g(gv).$$

Conversely a KG-module structure on V gives a linear action by restriction.

For the last part observe that if $\theta \in \mathrm{Hom}_K(V,W)$, then

θ is a homomorphism of representations

$\Leftrightarrow \sigma(g)\theta = \theta\rho(g)$ for all $g \in G$

$\Leftrightarrow \sigma(g)(\theta(v)) = \theta(\rho(g)(v))$ for all $g \in G$ and $v \in V$

$\Leftrightarrow g\theta(v) = \theta(gv)$ for all $g \in G$ and $v \in V$

$\Leftrightarrow x\theta(v) = \theta(xv)$ for all $x \in KG$ and $v \in V$

$\Leftrightarrow \theta$ is a homomorphism of KG-modules. □

Definition. Let (V,ρ) be a representation of G.

A **subrepresentation** of V is submodule U of the corresponding KG-module, so a subspace of V with $\rho(g)(u) \subseteq U$ for all $g \in G$ and $u \in U$.

There is a **quotient representation** $\overline{\rho} : G \to \mathrm{GL}(V/U)$ corresponding to the quotient module, so $\overline{\rho}(g)(U+v) = U + \rho(g)(v)$.

The representation is **simple** or **irreducible** if the corresponding module is simple, so has exactly two subrepresentations 0 and V.

The **direct sum** of representations V and W is given by the direct sum of the corresponding KG-modules, so $V \oplus W$ with the action $g(v,w) = (gv, gw)$.

A representation is **semisimple** or **completely reducible** if the corresponding module is semisimple, so every subrepresentation has a complement.

Theorem (Maschke)**.** *If G is a finite group and either $\mathrm{char}\, K = 0$ or $\mathrm{char}\, K$ is a prime number which does not divide $|G|$, then every representation of G is semisimple, so the group algebra KG is semisimple.*

Proof. The assumption on the characteristic of K ensures that $|G|$ is nonzero as an element of K.

Let M be a KG-module and N a submodule. As for Clfford algebras, to show that M is semisimple it suffices to show that there is a KG-module map $f : M \to N$ with $f(m) = m$ for all $m \in N$.

Now M is semisimple as a K-vector space, so there is a K-linear map $f' : M \to N$ with $f'(m) = m$ for all $m \in N$. Define $f : M \to N$ by

$$f(m) = \frac{1}{|G|}\sum_{g\in G} g^{-1} f'(gm).$$

If $h \in G$ we have

$$f(hm) = \frac{1}{|G|} \sum_{g \in G} g^{-1} f'(ghm).$$

Let $k = gh$, so $g^{-1} = hk^{-1}$. As g runs through the elements of G, so does k. Thus

$$f(hm) = \frac{1}{|G|} \sum_{k \in G} hk^{-1} f'(km) = hf(m).$$

Thus f is a KG-module homomorphism. Thus M is semisimple. $\square$

Remark. Since $\mathbb{C}$ is algebraically closed, the only f.d. division algebra over $\mathbb{C}$ is $\mathbb{C}$ itself. Thus by the Artin-Wedderburn Theorem, if G is finite, then

$$\mathbb{C}G \cong M_{m_1}(\mathbb{C}) \times \cdots \times M_{m_r}(\mathbb{C}).$$

For example:

$\mathbb{C}S_3$ has dimension 6 and it is not commutative, so it must be isomorphic to $\mathbb{C} \times \mathbb{C} \times M_2(\mathbb{C})$.

$\mathbb{C}C_n$ has dimension n and is commutative, so it must be isomorphic to $\mathbb{C} \times \cdots \times \mathbb{C}$.

4.2 Characters

From now on $K = \mathbb{C}$, the group G is finite, and we only consider representations which are finite dimensional vector spaces.

Definition. The **character** of a representation (V, ρ) with $\rho : G \to \mathrm{GL}(V)$ is the function $\chi_V : G \to \mathbb{C}$ given by $\chi_V(g) = \operatorname{tr} \rho(g)$, where tr is the trace.

A **character** is a function $\chi : G \to \mathbb{C}$ which arises as the character of some representation.

An **irreducible character** is the character of an irreducible representation.

A **class function** is a function $f : G \to \mathbb{C}$ which is constant on conjugacy classes, that is $f(h^{-1}gh) = f(g)$ for $g, h \in G$. The class functions form a subspace of the vector space of all functions $G \to \mathbb{C}$. The dimension is the number of conjugacy classes.

Properties. (i) Any character is a class function.

First observe that if $g, g' \in G$, then $\chi_V(gg') = \operatorname{tr}(\rho(gg')) = \operatorname{tr}(\rho(g)\rho(g')) = \operatorname{tr}(\rho(g')\rho(g)) = \chi_V(g'g)$.

Then $\chi_V(h^{-1}(gh)) = \chi_V((gh)h^{-1}) = \chi_V(g)$.

(ii) $\chi_V(1) = \operatorname{tr} \mathrm{Id}_V = \dim V$ is the degree of the representation V.

(iii) If V is a representation of degree 1, that is $\dim V = 1$, then χ_V is a group homomorphism $G \to \mathbb{C}^\times$.

The trace of a 1×1 matrix is identified with the matrix itself, so the character is identified with the corresponding matrix representation.

(iv) Isomorphic representations have the same character. [We will show later that the converse also holds.]

If (V, ρ) is isomorphic to (W, σ), then there is an isomorphism $f : V \to W$ with $f\rho(g) = \sigma(g)f$ for all $g \in G$. Then $\chi_V(g) = \operatorname{tr} \rho(g) = \operatorname{tr}(f^{-1}\sigma(g)f) = \operatorname{tr} \sigma(g) = \chi_W(g)$.

(v) The character of a direct sum of representations is the sum of their characters, $\chi_{V \oplus W}(g) = \chi_V(g) + \chi_W(g)$.

If (V, ρ) and (W, σ) are representations, their direct sum $V \oplus W$ becomes a representation $\tau : G \to \mathrm{GL}(V \oplus W)$ where $\tau(g)(v, w) = (\rho(v), \sigma(w))$. Combining bases of V and W gives a basis of $V \oplus W$, and the matrix of τ is block diagonal, so $\operatorname{tr} \tau(g) = \operatorname{tr} \rho(g) + \operatorname{tr} \sigma(g)$.

(vi) If χ is a character and $g \in G$ has order n, then $\chi(g)$ is a sum of n roots of 1, and $\chi(g^{-1}) = \overline{\chi(g)}$, the complex conjugate.

Proof. Since $\rho(g)^n = 1$, it is diagonalizable by example (c) at the end of Linear Algebra II §8.3. Thus there is a basis with repect to which the matrix $A(g)$ is diagonal. The diagonal entries must be nth roots of 1, so $\chi(g)$ is a sum of nth roots of 1. Moreover they have absolute value 1, so their inverses are their complex conjugates. Thus

$$\chi(g^{-1}) = \operatorname{tr} A(g^{-1}) = \operatorname{tr} A(g)^{-1} = \operatorname{tr} \overline{A(g)} = \overline{\operatorname{tr} A(g)} = \overline{\chi(g)}.$$

Examples. (i) The character of the trivial representation is $\chi : G \to \mathbb{C}$ with $\chi(g) = 1$ for all $g \in G$. It is called the **trivial character**.

(ii) Recall that the actions $G \times X \to X$, $(g, x) \mapsto gx$ of G on a set X correspond bijectively to group homomorphisms $G \to S_X$, where S_X is the symmetric group on X. Given such an action, we get a linear action of G on the vector space $\mathbb{C}X$ with basis X via

$$g(\sum_{xinX} a_x x) = \sum_{x \in X} a_x gx.$$

The corresponding representation $\rho : G \to \mathrm{GL}(\mathbb{C}X)$ is called a **permutation representation**.

Suppose that X is finite, say $X = \{x_1, \dots, x_n\}$. The matrix $A(g) = (a_{ij})$ of $\rho(g)$ with respect to the basis $(x_1, \dots, x_n)$ of $\mathbb{C}X$ satisfies

$$gx_j = \sum_{i=1}^{n} a_{ij} x_i.$$

so

$$a_{ij} = \begin{cases} 1 & (gx_j = x_i) \\ 0 & (gx_j \neq x_i) \end{cases}$$

Thus the corresponding character is

$$\chi(g) = \operatorname{tr} A(g) = \sum_{i=1}^{n} a_{ii} = |\{i : gx_i = x_i\}| = |\{x \in X : gx = x\}|.$$

(iii) The **regular representation** of G is the representation corresponding to the module ${}_{\mathbb{C}G}\mathbb{C}G$. This is the permutation representation corresponding to the action $G \times G \to G$ given by multiplication. The corresponding character is given by

$$\chi(g) = \begin{cases} |G| & (g = 1) \\ 0 & (g \neq 1). \end{cases}$$

(iv) If V is a representation of G, then the **dual representation** is given by the dual vector space $V^* = \operatorname{Hom}_{\mathbb{C}}(V, \mathbb{C})$ with the action of G given by

$$G \times V^* \to V^*, \quad (g, \xi) \mapsto g\xi, \quad (g\xi)(v) = \xi(g^{-1}v)$$

for $g \in G$, $\xi \in V^*$ and $v \in V$. We need to use the inverse to get the action property:

$$(g(g'\xi))(v) = (g'\xi)(g^{-1}v) = \xi((g')^{-1}g^{-1}v) = \xi((gg')^{-1}v) = ((gg')\xi)(v).$$

The character is

$$\chi_{V^*}(g) = \chi_V(g^{-1}) = \overline{\chi_V(g)}.$$

Proof. Suppose V has basis $v_1, \dots, v_n$, and the action of $g \in G$ has matrix $A(g) = (a_{ij}(g))$, then

$$gv_j = \sum_{i=1}^{n} a_{ij}(g) v_i$$

Let $\xi_1, \dots, \xi_n$ be the dual basis of V^*. Then

$$(g\xi_i)(v_j) = \xi_i(g^{-1}v_j) = a_{ij}(g^{-1})$$

so

$$g\xi_i = \sum_{j=1}^{n} a_{ij}(g^{-1})\xi_j$$

Thus the action of g on V^* has matrix $A(g^{-1})^{\mathrm{T}}$, so $\chi_{V^*}(g) = \operatorname{tr}(A(g^{-1})^{\mathrm{T}}) = \operatorname{tr}(A(g^{-1}) = \chi_V(g^{-1})$.

(v) If V and W are representations of V, the **tensor product representation** is $V \otimes_{\mathbb{C}} W$ with the action of G given by

$$G \times (V \otimes_{\mathbb{C}} W) \to V \otimes_{\mathbb{C}} W, \quad g(v \otimes w) = (gv) \otimes (gw)$$

for $g \in G$, $v \in V$, $w \in W$. The character is $\chi_{V \otimes_{\mathbb{C}} W}(g) = \chi_V(g)\chi_W(g)$.

Proof. If V has basis $v_1, \dots, v_n$ and W has basis $w_1, \dots, w_m$, then $V \otimes_{\mathbb{C}} W$ has basis $v_i \otimes w_j$. If the action of $g \in G$ on V has matrix $A(g) = (a_{ij})$ and on W has matrix $B(g) = (b_{ij})$, then

$$g(v_i \otimes w_j) = (gv_i) \otimes (gw_j) = \left(\sum_{p=1}^{n} a_{pi} v_p\right) \otimes \left(\sum_{q=1}^{m} b_{qj} w_q\right)$$

$$= \sum_{p=1}^{n} \sum_{q=1}^{m} a_{pi} b_{qj} v_p \otimes w_q$$

so the action of g on $V \otimes_{\mathbb{C}} W$ has matrix $C(g) = (c_{(p,q),(i,j)})$ with rows and columns indexed by pairs (i,j) and $c_{(p,q),(i,j)} = a_{pi} b_{qj}$. The character is

$$\chi_{V \otimes W}(g) = \operatorname{tr} C(g) = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{(i,j),(i,j)} = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ii} b_{jj}$$

$$= \left(\sum_{i=1}^{n} a_{ii}\right) \left(\sum_{j=1}^{m} b_{jj}\right) = (\operatorname{tr} A(g))(\operatorname{tr} B(g)) = \chi_V(g)\chi_W(g).$$

Here is another way to see this. Fix $g \in G$. We can choose a basis $(v_1, \dots, v_n)$ of V with respect to which the action of g is diagonal, say $gv_i = \lambda_i v_i$ with $\lambda_i \in \mathbb{C}$. Similarly, we can choose a basis $(w_1, \dots, w_m)$ of W with respect to which the action of g is diagonal, say $gw_j = \mu_j w_i$ with $\mu_j \in \mathbb{C}$. Then $(v_i \otimes w_j : 1 \leq i \leq n, 1 \leq j \leq m)$ is a basis of $V \otimes W$, and with respect to this basis the action of g is diagonal, with

$$g(v_i \otimes w_j) = (gv_i) \otimes (gw_j) = (\lambda_i v_i) \otimes (\mu_j w_j) = \lambda_i \mu_j (v_i \otimes w_j).$$

Then

$$\chi_{V \otimes W}(g) = \sum_{i,j} \lambda_i \mu_j = (\sum_i \lambda_i)(\sum_j \mu_j) = \chi_V(g)\chi_W(g).$$

Lemma. *Suppose V is a representation of G. The set of fixed points*

$$V^G = \{v \in V : gv = v \text{ for all } g \in G\}$$

is a subrepresentation of V, and

$$\dim V^G = \frac{1}{|G|} \sum_{g \in G} \chi_V(g)$$

Proof. The first statement is straightforward. We have a map

$$f : V \to V, \quad f(v) = \frac{1}{|G|} \sum_{g \in G} gv$$

Then f has image contained in V^G, and $f(v) = v$ for $v \in V^G$. Thus f is idempotent and $V^G = \operatorname{Im} f$. Thus $V = \operatorname{Im} f \oplus \operatorname{Ker} f$ by Linear Algebra II §7.4 Proposition. Combining bases of $\operatorname{Im} f$ and $\operatorname{Ker} f$ gives a basis of V, and with respect to this basis the matrix of f has block form

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

where I_r is an $r \times r$ identity matrix, where $r = \dim \operatorname{Im} f = \dim V^G$. Thus

$$r = \operatorname{tr} f = \frac{1}{|G|} \sum_{g \in G} \chi_V(g).$$

□

Definition. If $\phi : G \to \mathbb{C}$ and $\psi : G \to \mathbb{C}$ are mappings, we define

$$\langle \phi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \phi(g)\overline{\psi(g)}.$$

This defines a scalar product on the $\mathbb{C}$-vector space of all mappings $G \to \mathbb{C}$. By restriction it defines a scalar product on the subspace of class functions.

Theorem. *If V and W are representations, then*

$$\dim \operatorname{Hom}_{\mathbb{C}G}(V, W) = \langle \chi_W, \chi_V \rangle.$$

In particular, if V is an irreducible representation, then $\langle \chi_W, \chi_V \rangle$ is the multiplicity of V in the decomposition of W as a direct sum of irreducible representations. Thus any representation W is determined up to isomorphism by its character.

Proof. The space $\operatorname{Hom}_{\mathbb{C}}(V, W)$ becomes a representation of G with action

$$G \times \operatorname{Hom}_{\mathbb{C}}(V, W) \to \operatorname{Hom}_{\mathbb{C}}(V, W), \quad (g\theta)(v) = g\theta(g^{-1}v)$$

for $g \in G$, $\theta \in \operatorname{Hom}_{\mathbb{C}}(V, W)$ and $v \in V$. Moreover we have an isomorphism

$$V^* \otimes_{\mathbb{C}} W \cong \operatorname{Hom}_{\mathbb{C}}(V, W).$$

Thus

$$\dim \operatorname{Hom}_{\mathbb{C}G}(V, W) = \dim \operatorname{Hom}_{\mathbb{C}}(V, W)^G = \dim(V^* \otimes W)^G$$

$$= \frac{1}{|G|}\sum_{g\in G}\chi_{V^*\otimes W}(g) = \frac{1}{|G|}\sum_{g\in G}\overline{\chi_V(g)}\chi_W(g) = \langle \chi_W, \chi_V\rangle.$$

Now by semisimplicity we can write $W = W_1 \oplus \cdots \oplus W_m$ with the W_i irreducible, and by Schur's Lemma, for V irreducible we have

$$\mathrm{Hom}_{\mathbb{C}G}(V, W_i) \cong \begin{cases} \mathbb{C} & (V \cong W_i) \\ 0 & (V \ncong W_i). \end{cases}$$

Then

$$\mathrm{Hom}_{\mathbb{C}G}(V, W) = \mathrm{Hom}_{\mathbb{C}}(V, \bigoplus_{i=1}^m W_i) \cong \bigoplus_{i=1}^m \mathrm{Hom}_{\mathbb{C}}(V, W_i)$$

so

$$\langle \chi_W, \chi_V\rangle = \dim \mathrm{Hom}_{\mathbb{C}G}(V, W) = \sum_{i=1}^m \dim \mathrm{Hom}_{\mathbb{C}G}(V, W_i) = |\{i : W_i \cong V\}|.$$

Now if representations W and W' have the same character, they are isomorphic to directs sums of irreducible representations with the same multiplicities, and so $W \cong W'$. □

4.3 The character table

Still $K = \mathbb{C}$, the group G is finite, and we only consider f.d. representations.

Definition. The **character table** of G is the table with

- columns indexed by representatives $g_1, \dots, g_k$ of the conjugacy classes in G.

- rows indexed by the irreducible characters $\chi_1, \dots, \chi_r$ of G. Equivalently by the simple modules for $\mathbb{C}G$.

- entries $\chi_i(g_j)$.

Let $n_1, \dots, n_k$ be the sizes of the conjugacy classes, so $n_j = [G : C_G(g_j)]$. If $\phi, \psi : G \to \mathbb{C}$ are class functions, e.g. characters, then

$$\langle \phi, \psi\rangle = \frac{1}{|G|}\sum_{j=1}^k n_j \phi(g_j)\overline{\psi(g_j)}.$$

Properties. (i) The rows of the character table are orthonormal:

$$\langle \chi_i, \chi_j\rangle = \delta_{ij}.$$

(ii) If ϕ is a character, then $\phi = \sum_{i=1}^r c_i\chi_i$, where $c_i = \langle \phi, \chi_i\rangle \geq 0$. Then $\langle \phi, \phi\rangle = \sum_{i=1}^r c_i^2$. Thus ϕ is an irreducible character if and only if $\langle \phi, \phi\rangle = 1$.

(iii) Recall that the character ϕ of the regular representation is given by $\phi(1) = |G|$ and $\phi(g) = 0$ for $g \neq 1$. Thus $\langle \phi, \psi \rangle = \psi(1)$ if ψ is another character (so $\psi(1)$ is its degree, which is real). Thus

$$\phi = \sum_{i=1}^{r} \chi_i(1)\chi_i.$$

(iv) In particular

$$|G| = \sum_{i=1}^{r} \chi_i(1)^2,$$

the sum of the squares of the degrees of the irreducible characters.

Theorem. *The character table is square. That is, the number r of irreducible characters is equal to the number k of conjugacy classes in G. Thus the irreducible characters are an orthonormal basis of the vector space of class functions.*

Proof. Recall that if R is an algebra, then $Z(R)$ is its centre.

If $a = \sum_{g \in G} a_g g \in \mathbb{C}G$, then $a \in Z(\mathbb{C}G)$

$\Leftrightarrow ha = ah$ for all $h \in G$

$\Leftrightarrow a = hah^{-1}$ for all $h \in G$

$\Leftrightarrow \sum_{g \in G} a_g g = \sum_{g \in G} a_g hgh^{-1}$

$\Leftrightarrow \sum_{g \in G} a_g g = \sum_{x \in G} a_{h^{-1}xh} x$

$\Leftrightarrow a_g = a_{h^{-1}gh}$ for all $h \in G$.

$\Leftrightarrow$ the map $g \mapsto a_g$ is a class function.

Thus $\dim Z(\mathbb{C}G)$ is the dimension of the space of class functions, which is the number k of conjugacy classes.

Since $\mathbb{C}$ is algebraically closed, the only f.d. division algebra over $\mathbb{C}$ is $\mathbb{C}$ itself. Thus by the Artin-Wedderburn Theorem

$$\mathbb{C}G \cong M_{m_1}(\mathbb{C}) \times \cdots \times M_{m_r}(\mathbb{C}).$$

Each factor corresponds to a simple module $\mathbb{C}^{m_i}$, so there are r simple modules. Now if $R = R_1 \times \cdots \times R_r$, then $Z(R) = Z(R_1) \times \cdots \times Z(R_r)$, and it is easy to see that $Z(M_n(\mathbb{C})) = \mathbb{C}1$, so $\dim Z(\mathbb{C}G) = r$, so $r = k$. □

Example. If $G = C_n$ is a cyclic group, then representatives of the conjugacy classes are $1, g, \ldots, g^{n-1}$. The irreducible characters are $\chi_1, \ldots, \chi_n$ with $\chi_i(g^j) = \epsilon^{(i-1)j}$, where $\epsilon = e^{2\pi i/n}$ (where in this last formula, $i = \sqrt{-1}$). For example for $n = 2$

g_j	1	g
n_j	1	1
χ_1	1	1
χ_2	1	-1

For $n = 3$

g_j	1	g	g^2
n_j	1	1	1
χ_1	1	1	1
χ_2	1	ϵ	ϵ^2
χ_3	1	ϵ^2	ϵ

For example C_3 acts on $\mathbb{R}^2$ by rotations. This gives a representation

$$\rho : C_3 \to \mathrm{GL}_2(\mathbb{C}), \quad g^j \mapsto \begin{pmatrix} \cos(2\pi j/3) & -\sin(2\pi j/3) \\ \sin(2\pi j/3) & \cos(2\pi j/3) \end{pmatrix}.$$

The corresponding character is $\chi_2 + \chi_3$.

Example. Consider a product of two groups $G \times G'$.

Let g_i be representatives of the conjugacy classes in G, sizes n_i.

Let g'_j be representatives of the conjugacy classes in G', sizes n'_j.

Then (g_i, g'_j) are representatives of the conjugacy classes in $G \times G'$, sizes $n_i n'_j$.

Let χ_i be the irreducible characters of G.

Let χ'_j be the irreducible characters of G'.

The compositions

$$G \times G' \xrightarrow{p_1} G \xrightarrow{\chi_i} \mathbb{C}$$

$$G \times G' \xrightarrow{p_2} G' \xrightarrow{\chi'_j} \mathbb{C}$$

are characters of $G \times G'$ and their tensor product is the character

$$\chi_{ij} : G \times G' \to \mathbb{C}, \quad \chi_{ij}(g, g') = \chi_i(g)\chi'_j(g').$$

The degree is $\chi_{ij}(1) = \chi_i(1)\chi'_j(1)$.

Now

$$\langle \chi_{ij}, \chi_{ij} \rangle = \frac{1}{|G \times G'|} \sum_{a,b} n_a n'_b \chi_i(g_a)\chi_j(g'_b)\overline{\chi_i(g_a)\chi_j(g'_b)} = \langle \chi_i, \chi_i \rangle \langle \chi'_j, \chi'_j \rangle = 1.$$

Thus χ_{ij} is irreducible.

The number of such characters is the number of conjugacy classes in $G \times G'$, so they are all of the irreducible characters.

For example, for Klein's four group $V = C_2 \times C_2$, we get

g_j	$(1,1)$	$(g_1,1)$	$(1,g_2)$	(g_1,g_2)
n_j	1	1	1	1
χ_{11}	1	1	1	1
χ_{21}	1	-1	1	-1
χ_{12}	1	1	-1	-1
χ_{22}	1	-1	-1	1

Example. The group $G = S_3$ of order 6. The conjugacy classes are given by the cycle type.

g_j	1	(12)	(123)
n_j	1	3	2
χ_1	1	1	1
χ_2	1	-1	1
ϕ	3	1	0
χ_3	2	0	-1

χ_1 is the trivial character. Irreducible.

χ_2 is the sign character. Irreducible.

ϕ is the character of the natural permutation representation.

We have

$$\langle \phi, \chi_1 \rangle = \frac{1}{6}(1.1.3 + 3.1.1) = 1,$$

so $\phi = \chi_1 + \chi_3$ for some character χ_3, and

$$\langle \chi_3, \chi_3 \rangle = \frac{1}{6}(1.2^2 + 2.(-1)^2) = 1,$$

so χ_3 is irreducible.

Example. $G = A_4$ of order 12. The conjugacy classes are

$$\{1\}, \{(12)(34), (13)(24), (14)(23)\},$$

$$\{(123), (243), (134), (142)\}, \{(132), (234), (143), (124)\}.$$

g_j	1	(12)(34)	(123)	(132)
n_j	1	3	4	4
χ_1	1	1	1	1
χ_2	1	1	ϵ	ϵ^2
χ_3	1	1	ϵ^2	ϵ
ϕ	4	0	1	1
χ_4	3	-1	0	0

χ_1 is the trivial character.

$V = \{1, (12)(34), (13)(24), (14)(23)\}$ is a normal subgroup of A_4, and $A_4/V \cong C_3$. χ_2 and χ_3 are lifts of irreducible characters for C_3, where $\epsilon = e^{2\pi i/3}$.

ϕ is the character of the natural permutation representation. We have

$$\langle \phi, \chi_1 \rangle = \frac{1}{12}(4.1 + 3.0.1 + 4.1.1 + 4.1.1) = 1$$

Thus $\chi_4 = \phi - \chi_1$ is a character. It is an irreducible character since

$$\langle \chi_4, \chi_4 \rangle = \frac{1}{12}(3^2 + 3.(-1)^2) = 1.$$

Example. $G = A_5$ of order 60.

g_j	1	(12)(34)	(123)	(12345)	(12354)
n_j	1	15	20	12	12
χ_1	1	1	1	1	1
ϕ	5	1	2	0	0
χ_2	4	0	1	-1	-1
χ_3	5	1	-1	0	0
χ_4	3	-1	0	α	β
ψ	60	0	0	0	0
χ_5	3	-1	0	β	α

χ_1 is the trivial character.

ϕ is the natural permutation representation. Now

$$\langle \phi, \chi_1 \rangle = \frac{1}{60}(1.1.5 + 15.1.1 + 20.1.2) = 1,$$

so $\chi_2 = \phi - \chi_1$ is a character. It is irreducible since

$$\langle \chi_2, \chi_2 \rangle = \frac{1}{60}(1.4.4 + 20.1.1 + 12.(-1)^2 + 12.(-1)^2) = 1.$$

Let H be the subgroup of permutations fixing 5. So $H = A_4$. Index 5. Let $c = (12345)$, so

$$c = (12345), c^2 = (13524), c^3 = (14253), c^4 = (15432), c^5 = \mathrm{Id}.$$

Then for $1 \leq k \leq 5$ we have $c^k H = \{g \in G : g(5) = k\}$, so c, c^2, c^3, c^4, c^5 is a set of representatives of the left cosets of H in G.

Let V be the degree 1 representation of $H = A_4$ given by the row $1, 1, \epsilon, \epsilon^2$. Let $0 \neq v \in V$, so

$$(12)(34)v = v, \quad (123)v = \epsilon v, \quad (132)v = \epsilon^2 v.$$

We consider the induced representation

$$\operatorname{Ind}^{\mathbb{C}G} V = \mathbb{C}G \otimes_{\mathbb{C}H} V$$

Let χ_3 be its character. Since the c^k are representatives of the left cosets of H in G, they also give a basis of $\mathbb{C}G$ as a right $\mathbb{C}H$-module. Then $\operatorname{Ind}^{\mathbb{C}G} V \cong \bigoplus_{k=1}^{5} c^k \otimes V$, and since V is 1-dimensional with basis v, $\operatorname{Ind}^{\mathbb{C}G} V$ is a $\mathbb{C}$-vector space with basis the elements $c^k \otimes v$ for $1 \leq k \leq 5$. Thus $\chi_3(1) = 5$.

Now we need to compute $g_i(c^k \otimes v)$. For example what is $(12)(34)(c \otimes v)$.

$$(12)(34)(c \otimes v) = (12)(34)c \otimes v$$

Now the permutation $(12)(34)c$ sends 5 to 2, so it is in c^2H, and in fact it is equal to c^2h with $h = (143)$. Thus

$$(12)(34)(c\otimes v) = (12)(34)c\otimes v = c^2(143)\otimes v = c^2\otimes(143)v = c^2\otimes\epsilon^2 v = \epsilon^2(c^2\otimes v).$$

This does not contribute to the trace, since $c^2 \otimes v$ is not the same basis element as $c \otimes v$. In fact we only get a contribution when $g_i c^k \in c^k H$, which is when g_i fixes k. We have

$$(12)(34)c^5 \otimes v = c^5(12)(34) \otimes v = c^5 \otimes (12)(34)v = c^5 \otimes v.$$

Thus $\chi_3((12)(34)) = 1$.

Now the fixed points of (123) are 4 and 5 and we have

$$(123)c^4 \otimes v = c^4(234) \otimes v = c^4 \otimes (234)v = c^4 \otimes \epsilon^2 v,$$

$$(123)c^5 \otimes v = c^5(123) \otimes v = c^5 \otimes (123)v = c^5 \otimes \epsilon v.$$

Thus $\chi_3((123)) = \epsilon^2 + \epsilon = -1$.

Also (12345) and (12354) have no fixed points, so

$$\chi_3((12345)) = \chi_3((12354)) = 0.$$

Now

$$\langle \chi_3, \chi_3 \rangle = \frac{1}{60}(5^2 + 15.1^2 + 20.(-1)^2) = 1,$$

so χ_3 is irreducible.

The group of rotations of a dodecahedron has 60 elements, since a given face can be rotated to any of the other faces, and it has 5 possible orientations. The dodecahedron has 5 inscribed cubes which are permuted by these rotations. This gives a homomorphism from the rotation group to S_5, but the image has order 2, so it is a subgroup of index 2 in S_5. Thus it must be A_5.

This gives a representation of A_5. Let χ_4 be the character. It has degree 3 so $\chi_4(1) = 3$.

The conjugacy class $(12)(34)$ corresponds to a rotation about an axis through edge midpoints by angle π, so with respect to a suitable basis the matrix is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

So $\chi_4((12)(34)) = -1$.

The conjugacy class (123) corresponds to a rotation about an axis through opposite vertices by angle $2\pi/3$, so with respect to a suitable basis the matrix is

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(2\pi/3) & -\sin(2\pi/3) \\ 0 & \sin(2\pi/3) & \cos(2\pi/3) \end{pmatrix}$$

So $\chi_4((123)) = 1 + 2\cos(2\pi/3) = 0$.

The conjugacy class (12345) corresponds to a rotation about an axis through face centres. If we number the inscribed cubes appropriately, then it corresponds to rotation by angle $2\pi/5$, so $\chi_4((12345)) = 1 + 2\cos(2\pi/5) = \alpha$.

Then $(12345)^2 = (13524)$ is in the same conjugacy class as (12354) since the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix} = (235)$$

is even, and it corresponds to rotation by angle $4\pi/5$ so $\chi_4((12354)) = 1 + 2\cos(4\pi/5) = \beta$.

Letting $\eta = e^{2\pi i/5}$, we have $1 + \eta + \eta^2 + \eta^3 + \eta^4 = 0$, and $\alpha = 1 + \eta + \eta^4$, and then $(2\alpha - 1)^2 = 5$, so $\alpha = (1 + \sqrt{5})/2$ and then $\beta = 1 + \eta^2 + \eta^3 = 1 - \alpha = (1 - \sqrt{5})/2$.

Now

$$\langle \chi_4, \chi_4 \rangle = \frac{1}{60}(3^2 + 15.(-1)^2 + 12.\alpha^2 + 12.\beta^2) = 1$$

so χ_4 is irreducible.

Now there is only one more irreducible character χ_5. The degrees satisfy

$$\begin{aligned} 60 = |G| &= \chi_1(1)^2 + \chi_2(1)^2 + +\chi_3(1)^2 + \chi_4(1)^2 + \chi_5(1)^2 \\ &= 1^2 + 4^2 + 5^2 + 3^2 + \chi_5(1)^2, \end{aligned}$$

so $\chi_5(1) = 3$.

Let ψ be the character of the regular representation. Then

$$\psi = \chi_1(1)\chi_1 + \chi_2(1)\chi_2 + \chi_3(1)\chi_3 + \chi_4(1)\chi_4 + \chi_5(1)\chi_5$$

so

$$\chi_5 = \frac{1}{3}(\psi - \chi_1 - 4\chi_2 - 5\chi_3 - 3\chi_4).$$

(Alternatively we could have obtained χ_5 by arguing that if we had numbered the inscribed cubes differently, then the rotation by angle $4\pi/5$ could have corresponded to the cycle (12345). Then we could have used the regular representation to find χ_3, so avoiding induced representations.)

5 Commutative algebra

All rings are now commutative, unless explicitly stated otherwise.

5.1 Localization and prime ideals

Let R be a (commutative!) ring.

Definition. Recall from Exercise Sheet 7, that a subset $S \subseteq R$ is **multiplicative** if $1 \in S$ and $st \in S$ for all $s, t \in S$. If so, the **localization** of R at S is

$$S^{-1}R = \{r/s : r \in R, s \in S\}$$

where $r/s = r'/s' \Leftrightarrow t(s'r - sr') = 0$ for some $t \in S$. It is a ring with the usual addition and multiplication of fractions. There is a ring homomorphism $R \to S^{-1}R$, $r \mapsto r/1$ with kernel $\{r \in R : sr = 0 \text{ for some } s \in S\}$.

Definition. Suppose $\theta : R \to R'$ is a homomorphism of rings.

If I is an ideal of R, its **extension** to R' is the ideal $I^e := (\theta(I))$ of R'.

If I' is an ideal of R', its **contraction** to R is the ideal $(I')^c := \theta^{-1}(I')$ of R.

It is easy to see that $I \subseteq I^{ec}$ and $(I')^{ce} \subseteq I'$.

If θ is the inclusion of a subring, then $I^e = (I)$ and $(I')^c = R \cap I'$.

Proposition. *Suppose S is a multiplicative set in R and let $\theta : R \to R' = S^{-1}R$ be the natural map.*

(i) If I' is an ideal in R', then $(I')^{ce} = I'$.

(ii) Suppose I is an ideal in R then:
(a) $I^e = \{a/s : a \in I, s \in S\}$.
(b) If $r \in R$ then $r/1 \in I^e \Leftrightarrow sr \in I$ for some $s \in S$.
(c) If $S \cap I = \emptyset$ and no element of S is a zero divisor in R/I, then $I = I^{ec}$.

Proof. (i) If $a/s \in I'$, then $a/1 = (s/1)(a/s) \in I'$ so $a \in (I')^c$ so $a/1 \in (I')^{ce}$ so $a/s = (1/s)(a/1) \in (I')^{ce}$.

(ii) (a) Any element of I^e has the form $\sum_i x_i(a_i/1)$ for some $x_i \in R'$ and $a_i \in I$. Writing the x_i over a common denominator as $x_i = r_i/s$, the element is a/s where $a = \sum r_i a_i \in I$.

(b) If $sr \in I$ then $r/1 = sr/s \in I^e$. Conversely if $r/1 \in I^e$ then $r/1 = a/s$ with $a \in I$ and $s \in S$, so $t(sr - a) = 0$ for some $t \in S$, so $(ts)r = ta \in I$.

(c) If $a \in I^{ec}$ then $a/1 \in I^e$, so $sa \in I$ for some $s \in S$. Then in R/I we have $\overline{s}\,\overline{a} = 0$, so $\overline{a} = 0$ so $a \in I$. □

Corollary. *If R is noetherian, then so is $S^{-1}R$.*

Definition. Recall that an ideal P in R is **prime** if R/P is an integral domain. Equivalently $P \neq R$ and $a, b \in R$ and $ab \in P$ implies $a \in P$ or $b \in P$. Any maximal ideal is prime.

$$\operatorname{Maxspec} R := \{\text{maximal ideals } M \text{ in } R\} \subseteq \operatorname{Spec} R := \{\text{prime ideals } P \text{ in } R\}$$

If $\theta : R \to R'$, then the map $P' \mapsto (P')^c$ gives a mapping $\operatorname{Spec} R' \to \operatorname{Spec} R$, since θ induces an injective homomorphism $R/(P')^c \to R'/P'$.

Corollary. *Extension and contraction give inverse bijections between the prime ideals of $S^{-1}R$ and the prime ideals of R which are disjoint from S.*

Definition. A ring R (commutative or not) is **local** if the set of non-invertible elements forms an ideal I. If so, then R/I is a division ring and I is the Jacobson radical of R (see Aufgabe 6.2).

A commutative ring is local if and only if it has a unique maximal ideal.

Proposition. *An ideal P in R is prime if and only if $S = R \setminus P$ is a multiplicative set. In this case $S^{-1}R$ is denoted R_P, and it is a local ring with maximal ideal P^e. Moreover R_P/P^e is isomorphic to the quotient field $\kappa(P)$ of R/P.*

Proof. The first statement is clear. The prime ideals in R_P are of the form p^e with $p \subseteq P$, so P^e is the unique maximal ideal. Now if $s \in S$ then $P + s$ is nonzero in R/P, so we get a homomorphism

$$R_P \to \kappa(P), \quad r/s \mapsto (P + r)/(P + s)$$

for $r \in R$ and $s \in R \setminus P$. Clearly it is surjective.

The kernel contains the elements $a/1$ with $a \in P$, so it also contains P^e. Thus we get a homomorphism $R_P/P^e \to \kappa(P)$. Since R_P/P^e is a field, this is homomorphism is injective, so an isomorphism. □

Definition. If I is an ideal in R, its **radical** is

$$\sqrt{I} = \{a \in R : a^n \in I \text{ for some } n > 0\}.$$

It is an ideal in R, since if $a^n = b^m = 0$, then

$$(a + b)^{n+m} = \sum_{i=0}^{n+m} \binom{n+m}{i} a^i b^{n+m-i} = 0.$$

Theorem. *$\sqrt{I}$ is the intersection of the prime ideals in R containing I.*

Proof. Let $a \in R$. If $I \subseteq P$ with P prime and $a \in \sqrt{I}$, then $a^n \in I \subseteq P$. Then $P + a$ is nilpotent in R/P. But this is a domain, so has no nilpotent elements. Thus $P + a = 0$, so $a \in P$.

Now suppose $a \notin \sqrt{I}$. Let $S = \{1, a, a^2, \dots\}$ and consider $\theta : R \to R' = S^{-1}R$. Then $1/1 \notin I^e$ by (ii)(b) of the proposition above, so I^e is a proper ideal in R', so it is contained in a maximal ideal M. Now $a/1$ is a unit in R', with inverse $1/a$, so $a/1 \notin M$, so $a \notin M^c$, and this is a prime ideal in R. Moreover $I^e \subseteq M$ implies $I \subseteq I^{ec} \subseteq M^c$. □

Definition. Let I be an ideal in R. A **minimal prime over** I is a prime ideal containing I which is minimal with this property.

Proposition. *Suppose I is an ideal in a ring R.*

(i) Any prime ideal P of R containing I contains a minimal prime over I.

(ii) $\sqrt{I}$ is the intersection of the minimal primes over I.

(iii) If R is noetherian, there are only finitely many minimal primes over I.

Proof. (i) Let X be the set of prime ideals p with $I \subseteq p \subseteq P$. We partially order X by the opposite of the inclusion ordering. Any intersection I of a chain of ideals in X is in X. Namely, suppose $ab \in I$ and $a, b \notin I$. Then $b \notin p$ and $a \notin p'$ for some primes $p, p' \in X$. Without loss of generality, $p \subseteq p'$. Then $a, b \notin p$ but $ab \in p$, a contradiction. Now Zorn's Lemma implies that X has a maximal element, which is a minimal prime over I.

(ii) Follows from (i).

(iii) Suppose false. Let J be maximal such that there are infinitely many minimal primes over J. Then J is not prime, so there are $a, b \notin J$ with $ab \in J$. If p is a minimal prime over J then $ab \in p$, so $a \in p$ or $b \in p$. Thus p is minimal over $J + (a)$ or $J + (b)$, but these have finitely many minimal primes. □

Example. Recall that if R is a UFD, then any $0 \neq a \in R$ which is not a unit can be written as a product of irreducible elements $a = b_1 b_2 \dots b_n$ and this decomposition is unique up to ordering and multiplication by units. Moreover the principal ideal (b_i) generated by an irreducible element is prime. The minimal primes over (a) are the prime ideals (b_i), for if $(a) \subseteq p \subseteq (b_i)$, with p a prime ideal, then $a = b_1 \dots b_n \in p$, so some $b_j \in p$, but then $b_j \in (b_i)$. It follows that b_j is a unit times b_i, so $(b_i) = (b_j) \subseteq p \subseteq (b_i)$.

5.2 Integral extensions

Definition. Suppose R is a subring of R' and $\alpha \in R'$.

We say that α **is integral over** R if there is a monic polynomial $f(X) \in R[X]$ with $f(\alpha) = 0$.

If every element of R' is integral over R, we say that R' is **integral over** R or that $R \subseteq R'$ is an **integral extension**.

Proposition. *Suppose $R \subseteq R'$.*

(i) If R' is a f.g. R-module, then it is integral over R.

(ii) $\alpha \in R'$ is integral over $R \Leftrightarrow R[\alpha]$ is a f.g. R-module.

Proof. (i) Let $\alpha \in R'$ and let $R' = \sum_{i=1}^n Rx_i$ with $x_1 = 1$. We can write $\alpha x_i = \sum_{j=1}^n a_{ij}x_j$ for some matrix $A = (a_{ij}) \in M_n(R)$.

Let $f(X) = \det(A - XI) \in R[X]$ be the characteristic polynomial of A. It suffices to prove that $f(\alpha) = 0$.

Let $T = A - \alpha I \in M_n(R')$. Let $x \in (R')^n$ be the column vector with components x_i. Then $Tx = 0$.

Now $\operatorname{adj}(T)T = \det(T)I$. (By Linear Algebra I §6.4 Satz 5 this holds for a matrix T with entries in a field, but here we need it for matrices over the ring R'. Working over the field $\mathbb{Q}(X_{ij} : 1 \le i, j \le n)$, for $n = 3$ we get an identity

$$\begin{pmatrix} X_{22}X_{33} - X_{23}X_{32} & -(X_{21}X_{33} - X_{23}X_{31}) & * \\ * & * & * \\ * & * & * \end{pmatrix} \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix}$$

$$= (X_{11}X_{22}X_{33} - X_{12}X_{21}X_{33} + \dots) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This involves matrices over $\mathbb{Z}[X_{ij}]$, and now by a ring homomorphism $\mathbb{Z}[X_{ij}] \to R'$ we can specialize the X_{ij} to elements of R'.)

Thus $\det(T)x = \operatorname{adj}(T)Tx = 0$. Thus since $x_1 = 1$ we get $\det(T) = 0$, so $f(\alpha) = 0$.

(ii) Suppose $f(X) \in R[X]$ is a monic polynomial of degree n with $f(\alpha) = 0$. Then $\alpha^n \in \sum_{i=1}^{n-1} R[X]\alpha^i$. By induction the sum contains α^m for all $m \ge n$. Thus the sum is $R[\alpha]$. The converse follows from (i). □

Properties. (i) Suppose $R \subseteq R'$ and $\alpha_1, \dots, \alpha_n \in R'$. If each α_i is integral over $R[\alpha_1, \dots, \alpha_{i-1}]$, then $R[\alpha_1, \dots, \alpha_n]$ is f.g. as an R-module, so it is integral over R.

Proof. By induction $R'' = R[\alpha_1, \dots, \alpha_{n-1}]$ is f.g. as an R-module, say $R'' = \sum_{i=1}^n Rx_i$. Since α_n is integral over R', we have

$$R[\alpha_1, \dots, \alpha_n] = R''[\alpha_n] = \sum_{j=1}^m R''y_j = \sum_{i,j} Rx_iy_j.$$

(ii) If $R \subseteq R' \subseteq R''$, then R'' is integral over R if and only if R'' is integral over R' and R' is integral over R.

Proof. It is clear that if R'' is integral over R then it is integral over R' and R' is integral over R.

For the converse, say $\alpha \in R''$. Then there is a monic polynomial $f(X) \in R'[X]$ with $f(\alpha) = 0$. Let $a_0, \ldots, a_n \in R'$ be its coefficients. Then by property (i), $R[a_0, \ldots, a_n, \alpha]$ is integral over R, so α is integral over R.

(iii) If $R \subseteq R'$ is an integral extension and $\theta : R' \to R''$ is a ring homomorphism, then $\theta(R')$ is integral over $\theta(R)$.

If $\alpha \in R'$ and $f(\alpha) = 0$ with $f(X) \in R[X]$ monic. Then $f'(X) = \theta(f(X)) \in \theta(R)[X]$ is monic and $f'(\theta(\alpha)) = \theta(f(\alpha)) = 0$, so $\theta(\alpha)$ is integral over $\theta(R)$.

(iv) If $R \subseteq R'$ is an integral extension and S is a multiplicatively closed subset of R then $S^{-1}R'$ is integral over $S^{-1}R$.

If is clear that $S^{-1}R$ is a subring of $S^{-1}R'$. Suppose $a \in R'$ is a root of the monic polynomial

$$X^n + r_{n-1}X^{n-1} + \cdots + r_1X + r_0,$$

then for $s \in S$ the element a/s is a root of the monic polynomial

$$X^n + (r_{n-1}/s)X^{n-1} + \cdots + (r_1/s^{n-1})X + (r_0/s^n).$$

(v) If $R \subseteq R'$ is an integral extension of integral domains, then R is a field if and only if R' is a field.

If R' is a field and $0 \neq r \in R$, then r^{-1} exists in R', so there is a polynomial with

$$(r^{-1})^n + r_{n-1}(r^{-1})^{n-1} + \cdots + r_0 = 0$$

Multiplying by r^{n-1} we get $r^{-1} \in R$.

If R is a field then $0 \neq \alpha \in R'$ satisfies a polynomial

$$\alpha^n + r_{n-1}\alpha^{n-1} + \cdots + r_1\alpha + r_0 = 0$$

and since R' is an integral domain we may take $r_0 \neq 0$. Then

$$\alpha^{-1} = -(r_0)^{-1}(\alpha^{n-1} + r_{n-1}\alpha^{-2} + \cdots + r_1).$$

(vi) If $R \subseteq R'$ is an integral extension, then the contraction of a maximal ideal m' in R' is maximal in R.

Let $\theta : R' \to R'/m'$ be the canonical map. By Property (iii), $\theta(R') = R'/m'$ is integral over $\theta(R) = (R + m')/m' \cong R/(R \cap m') = R/(m')^c$. Now use (v).

Theorem (Lying over). *If $R \subseteq R'$ is an integral extension, then every $P \in \operatorname{Spec} R$ is the contraction of some $P' \in \operatorname{Spec} R'$.*

Proof. Consider

$$\begin{tikzcd} R \arrow[r] \arrow[d] & R' \arrow[d] \\ R_P \arrow[r] & (R \setminus P)^{-1}R' \end{tikzcd}$$

Then R_P is a local ring with maximal ideal ideal P^e. So it is not zero. So $(R \setminus P)^{-1}R'$ is not zero. Take a maximal ideal m in it.

Since $(R \setminus P)^{-1}R'$ is integral over R_P, the contraction of m to R_P is maximal, so equal to P^e. Thus the contraction of m to R is P. On the other hand the contraction of m to R' is a prime P', and its contraction to R is P. □

Definition. Let R be a subring of R'. The **integral closure of R in R'** is

$$\overline{R}^{R'} = \{\alpha \in R' : \alpha \text{ is integral over } R\}.$$

It is a subring of R', for if α_1, α_2 are integral over R, then α_2 is also integral over $R[\alpha_1]$, so $R[\alpha_1, \alpha_2] \subseteq \overline{R}^{R'}$ by property (i).

If $\overline{R}^{R'} = R$, we say that R is **integrally closed in R'**.

Note that $\overline{R}^{R'}$ is integrally closed in R', since if $\alpha \in R'$ is integral over $\overline{R}^{R'}$, then $\overline{R}^{R'}[\alpha]$ is integral over $\overline{R}^{R'}$, so by property (ii) it is integral over R, so α is integral over R, so $\alpha \in \overline{R}^{R'}$.

Theorem. *Suppose R is a UFD with field of fractions K.*

(i) $\overline{R}^{K} = R$, so R is integrally closed in K.

(ii) If L/K is a field extension, then

$$\overline{R}^{L} = \{\alpha \in L : \alpha \textit{ algebraic over } K \textit{ and } m_{\alpha/K}(X) \in R[X]\}.$$

Proof. (i) Rational root test, Algebra I §4.5.

(ii) Say $\alpha \in \overline{R}^{L}$. Take $p(X) \in R[X]$ monic of least degree with $p(\alpha) = 0$. Clearly $p(X)$ is irreducible in $R[X]$. Then by Algebra I, §4.4 Satz (i), $p(X)$ is irreducible in $K[X]$, so it is the minimal polynomial of α. □

Examples. A **number field** is a finite extension field L of $\mathbb{Q}$. Its **ring of integers** is $\mathcal{O}_L = \overline{\mathbb{Z}}^{L}$.

If $\alpha = a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ with $a, b \in \mathbb{Q}$ and $b \neq 0$, then

$$m_{\alpha/\mathbb{Q}}(X) = (X - a)^2 - 2b^2 = X^2 - 2aX + a^2 - 2b^2.$$

This is in $\mathbb{Z}[X] \Leftrightarrow 2a, a^2 - 2b^2 \in \mathbb{Z} \Leftrightarrow a, b \in \mathbb{Z}$, so $\mathcal{O}_{\mathbb{Q}(\sqrt{2})} = \mathbb{Z}[\sqrt{2}]$.

Similarly $\mathcal{O}_{\mathbb{Q}(i)} = \mathbb{Z}[i]$.

If $\alpha = a + b\sqrt{5} \in \mathbb{Q}(\sqrt{5})$ with $a, b \in \mathbb{Q}$ and $b \neq 0$, then

$$m_{\alpha/\mathbb{Q}}(X) = (X - a)^2 - 5b^2 = X^2 - 2aX + a^2 - 5b^2.$$

This is in $\mathbb{Z}[X] \Leftrightarrow 2a, a^2 - 5b^2 \in \mathbb{Z} \Leftrightarrow a, b \in \mathbb{Z}$ or $a, b \in \mathbb{Z} + \frac{1}{2}$. For example if $\alpha = \frac{1}{2}(1 + \sqrt{5})$ then $\alpha^2 = \alpha + 1$. It follows that $\mathcal{O}_{\mathbb{Q}(\sqrt{5})} = \mathbb{Z}[\frac{1+\sqrt{5}}{2}]$.

The ring of **algebraic integers** is $\overline{\mathbb{Z}}^{\mathbb{C}}$. We use algebraic integers to prove the following theorem.

Theorem. *The degrees of the irreducible complex representations of a finite group divide the order of the group.*

Proof. Let $g_1, \dots, g_k$ be representatives of the conjugacy classes. Let

$$c_j = \sum_{g \in [g_j]} g \in \mathbb{C}G$$

be the **class sum**, the sum of the elements conjugate to g_j.

The elements c_j are a basis for $Z(\mathbb{C}G)$, and clearly

$$c_i c_j = \sum_k a_{ijk} c_k$$

for some $a_{ijk} \in \mathbb{N}$.

By Schur's Lemma, c_j acts on an irreducible representation V as multiplication by a scalar $\omega_j \in \mathbb{C}$. Then $\omega_i \omega_j = \sum_k a_{ijk} \omega_k$. Thus $R = \sum_j \mathbb{Z}\omega_j$ is a subring of $\mathbb{C}$. Thus the ω_j are algebraic integers.

Considering the trace of the action of c_j on V, we get

$$n_j \chi_V(g_j) = (\dim V)\omega_j$$

where $n_j = |[g_j]|$. Then

$$1 = \langle \chi_V, \chi_V \rangle = \frac{1}{|G|} \sum_j n_j \chi_V(g_j) \chi_V(g_j^{-1}) = \frac{1}{|G|} \sum_j (\dim V)\omega_j \, \chi_V(g_j^{-1}).$$

Thus

$$\frac{|G|}{\dim V} = \sum_j \omega_j \chi_V(g_j^{-1}).$$

Now $\chi_V(g_j^{-1})$ is a sum of roots of 1, so an algebraic integer. Thus the right hand side is an algebraic integer, but rational, so an integer. □

5.3 The Nullstellensatz

Lemma (1). *Let K be a field and $n > 0$. By the substitution*

$$Y_i := X_i - X_1^{r^{i-1}} \quad (i = 2, \dots, n)$$

with $r > 0$, we can identify

$$K[X_1, \dots, X_n] = K[X_1, Y_2, \dots, Y_n] = R[X_1]$$

where $R = K[Y_2, \dots, Y_n]$,

Given $0 \neq f(X_1, \dots, X_n) \in K[X_1, \dots, X_n]$ we can choose r such that $f(X_1, \dots, X_n)$ corresponds to a scalar multiple of a monic polynomial in $R[X_1]$.

Proof. Any polynomial in $X_1, Y_2, \dots, Y_n$ can be written as a polynomial in $X_1, X_2, \dots, X_n$ via the substitution, and this is reversible by the substitution

$$X_i = Y_i + X_1^{r^{i-1}} \quad (i = 2, \dots, n).$$

Any monomial $X_1^{d_1} X_2^{d_2} \dots X_n^{d_n}$ involved in $f(X_1, \dots, X_n)$ becomes

$$X_1^{d_1}(Y_2 + X_1^r)^{d_2} \dots (Y_n + X_1^{r^{n-1}})^{d_n}$$

which is a monic polynomial in $R[X_1]$ of degree

$$d_1 + d_2 r + \dots + d_n r^{n-1}.$$

The finitely many monomials involved in $f(X_1, \dots, X_n)$ give a finite number of polynomials $d_1 + d_2 T + \dots + d_n T^{n-1}$, so we can choose r such that their evaluations at r are all different. Then

$$f(X_1, \dots, X_n) = f(X_1, Y_2 + X_1^r, \dots, Y_n + X_1^{r^{n-1}}) = \lambda \left(X_1^m + \sum_{j=0}^{m-1} g_j(Y_2, \dots, Y_n) X_1^j \right)$$

where m is the maximal value of $d_1 + d_2 r + \dots + d_n r^{n-1}$ for a monomial involved in $f(X_1, \dots, X_n)$, λ is the coefficient of that monomial and $g_j(Y_2, \dots, Y_n) \in R$. □

Lemma (2). *If $f(X) \in R[X]$ is a monic polynomial of degree $n > 0$, then the natural map $R \to R[X]/(f(X))$ is injective, so we can identify R as a subring of $R[X]/(f(X))$. Moreover $R[X]/(f(X))$ is integral over R.*

Proof. If $g(X) \in R[X]$ has leading term $a_m X^m$ then $f(X)g(X)$ has leading term $a_m X^{n+m}$, so it cannot be a nonzero element of R. Thus the map $R \to R[X]/(f(X))$ is injective. Let $\overline{X}$ be the image of X in $R[X]/(f(X))$. Then $f(\overline{X}) = \overline{f(X)} = 0$, and $R[X]/(f(X)) = R[\overline{X}]$, so it is integral over R by property (i) of integral extensions. □

Theorem (Noether Normalization). *Let K be a field. If R is a f.g. K-algebra, then it contains a subalgebra S which is isomorphic to a polynomial algebra $K[X_1, \ldots, X_n]$, and such that R is integral over R, and hence a f.g. S-module.*

Proof. Let $\theta : K[X_1, \ldots, X_n] \to R$ be a ring homomorphism with R integral over $\mathrm{Im}(\theta)$ and n minimal with this property. It exists since R is f.g. as a K-algebra, so there is even a surjective homomorphism.

If θ is not injective, then the kernel contains an non-zero element f. By Lemma 1 we may assume that f is monic in $R_0[X_1]$, where $R_0 = K[X_2, \ldots, X_n]$. Then we get

$$R_0 \xrightarrow{\phi} K[X_1, \ldots, X_n]/(f) \xrightarrow{\overline{\theta}} R.$$

Now R is integral over $\mathrm{Im}\,\overline{\theta}$ by assumption. Also $\mathrm{Im}\,\overline{\theta}$ is integral over $\mathrm{Im}\,\overline{\theta}\phi$ by Lemma 2 and property (iii) of integral extensions. Thus R is integral over $\mathrm{Im}\,\overline{\theta}\phi$ by property (ii) of integral extensions, contradicting the minimality of n.

Now R is f.g. as an $\mathrm{Im}(\theta)$-module by property (i) of integral extensions, since it is f.g. as a K-algebra. □

Theorem (Weak Nullstellensatz). *If L/K is a field extension, and L is f.g. as a K-algebra, then L/K is a finite field extension.*

Proof. By Noether normalization L is f.g. as a module over a subring $S \cong K[X_1, \ldots, X_n]$. Now by property (v) of integral extensions S is a field, so $n = 0$. Thus L/K is a finite field extension. □

Lemma. *(i) If K is a field, the K-algebra homomorphisms $\theta : K[X_1, \ldots, X_n] \to K$ are exactly the maps*

$$f(X_1, \ldots, X_n) \mapsto f(a) := f(a_1, \ldots, a_n)$$

for some $a = (a_1, \ldots, a_n) \in K^n$. Moreover $\mathrm{Ker}\,\theta$ is equal to

$$m_a = (X_1 - a_1, \ldots, X_n - a_n)$$

and it is a maximal ideal in $K[X_1, \ldots, X_n]$.

(ii) If K is an algebraically closed field, then every maximal ideal of $K[X_1, \ldots, X_n]$ is of the form m_a for some $a \in K^n$

Proof. (i) Given θ, set $a_i = \theta(X_i)$. We have $X_i - a_i \in \mathrm{Ker}\,\theta$ so $m_a \subseteq \mathrm{Ker}\,\theta$. Conversely suppose that $f \in \mathrm{Ker}\,\theta$. Let

$$g(Y_1, \ldots, Y_n) = f(Y_1 + a_1, \ldots, Y_n + a_n) \in K[Y_1, \ldots, Y_n].$$

Then $g(0, \ldots, 0) = f(a_1, \ldots, a_n) = 0$. Thus g has constant term zero, so we we can write

$$g = \sum_{i=1}^{n} h_i Y_i$$

with $h_i \in K[Y_1, \ldots, Y_n]$. Then

$$f(X_1, \ldots, X_n) = g(X_1 - a_1, \ldots, X_n - a_n) = \sum_{i=1}^{n} h_i(X_1 - a_1, \ldots, X_n - a_n)(X_i - a_i) \in m_a.$$

Since θ is surjective, $K[X_1, \ldots, X_n]/m_a \cong K$ is a field, so m_a is a maximal ideal.

(ii) Suppose M is a maximal ideal. It is the kernel of the canonical homomorphism $K[X_1, \ldots, X_n] \to K[X_1, \ldots, X_n]/M$. Now $K[X_1, \ldots, X_n]/M$ is a field extension of K, finitely generated as a K-algebra, so a finite extension of K by the weak Nullstellensatz. By algebraic closure, it is K. □

Definition. Let K be an algebraically closed field.

If V is a subset of K^n, we define

$$\begin{aligned}\mathbb{I}(V) &= \{f \in K[X_1, \ldots, X_n] : f(a) = 0 \text{ for all } a \in V\} \\ &= \{f \in K[X_1, \ldots, X_n] : f \in m_a \text{ for all } a \in V\}\end{aligned}$$

It is an ideal in $K[X_1, \ldots, X_n]$.

If S is a subset of $K[X_1, \ldots, X_n]$, we define

$$\mathbb{V}(S) = \{a \in K^n : f(a) = 0 \text{ for all } f \in S\} = \{a \in K^n : S \subseteq m_a\}.$$

Clearly $\mathbb{V}(S) = \mathbb{V}(I)$ where I is the ideal generated by S.

Theorem (Hilbert's Nullstellensatz)**.** *Let K be an algebraically closed field. If I is an ideal in $K[X_1, \ldots, X_n]$ then $\mathbb{I}(\mathbb{V}(I)) = \sqrt{I}$. Equivalently, $\sqrt{I}$ is the intersection of all maximal ideals containing I.*

Proof. If $f \in \sqrt{I}$, then $f^m \in I$ for some $m > 0$. For $a \in \mathbb{V}(I)$, we have $f(a)^n = 0$. Thus $f(a) = 0$. Thus $f \in \mathbb{I}(\mathbb{V}(I))$.

Now suppose $f \in \mathbb{I}(\mathbb{V}(I))$. Consider the ideal

$$J = (I, Xf - 1) \subseteq K[X_1, \ldots, X_n, X].$$

If J is a proper ideal in $K[X_1, \ldots, X_n, X]$, then it is contained in some maximal ideal

$$\begin{aligned} m_{(a_1,\ldots,a_n,b)} &= (X_1 - a_1, \ldots, X_n - a_n, X - b) \\ &= \operatorname{Ker} K[X_1, \ldots, X_n, X] \mapsto K, \quad g(X_1, \ldots, X_n, X) \mapsto g(a_1, \ldots, a_n, b).\end{aligned}$$

If $h \in I$, then $h \in J \subseteq m_{(a_1,\ldots,a_n,b)}$, so $h(a_1, \ldots, a_n) = 0$. Thus $a = (a_1, \ldots, a_n) \in \mathbb{V}(I)$. Also $Xf - 1 \in J \subseteq m_{(a_1,\ldots,a_n,b)}$, so $bf(a_1, \ldots, a_n) - 1 = 0$, so $f(a_1, \ldots, a_n) \neq 0$, contradicting the assumption that $f \in \mathbb{I}(\mathbb{V}(I))$.

Thus J is not a proper ideal. Thus we have

$$1 = \sum_{i=1}^{k} h_i g_i + h(Xf - 1)$$

with the g_i in I and $h_i, h \in K[X_1, \ldots, X_n, X]$. Applying the homomorphism

$$K[X_1, \ldots, X_n, X] \mapsto K(X_1, \ldots, X_n), \quad X_i \mapsto X_i, X \mapsto 1/f,$$

we get

$$1 = \sum_{i=1}^{k} h_i(X_1, \ldots, X_n, 1/f) g_i$$

in $K(X_1, \ldots, X_n)$. But we can write

$$h_i(X_1, \ldots, X_n, 1/f) = k_i / f^{d_i}$$

for some $k_i \in K[X_1, \ldots, X_n]$ and $d_i \geq 0$, and we may assume that all $d_i = d$. Then $1 = \ell / f^d$ where $\ell = \sum_{i=1}^{k} k_i g_i \in I$. Thus $f^d \in I$, so $f \in \sqrt{I}$. □

Definition. We say that an ideal I in a ring is **radical** if $I = \sqrt{I}$. Note that if I is any ideal, then $\sqrt{\sqrt{I}} = \sqrt{I}$, so $\sqrt{I}$ is radical.

We say that a subset $V \subseteq K^n$ is **Zariski-closed** if it is of the form $\mathbb{V}(S)$ for some subset S of $K[X_1, \ldots, X_n]$.

Corollary. *If K is an algebraically closed field, we have inverse bijections*

$$\{\textit{Radical ideals } I \textit{ in } K[X_1, \ldots, X_n]\} \underset{\mathbb{I}}{\overset{\mathbb{V}}{\rightleftarrows}} \{\textit{Zariski-closed subsets } V \textit{ of } K^n\}.$$

The bijections reverse inclusions.

Examples.

Poly ring	Ideal I	$\mathbb{V}(I) \cap \mathbb{R}^n$	Min primes over I
$\mathbb{C}[X, Y]$	$(Y - X^2)$	parabola $y = x^2$	I
$\mathbb{C}[X, Y]$	$(XY - 1)$	hyperbola $xy = 1$	I
$\mathbb{C}[X, Y]$	$((Y - X^2)(XY - 1))$	parabola ∪ hyperbola	$(Y - X^2), (XY - 1)$
$\mathbb{C}[X, Y, Z]$	(XY, XZ)	(plane $x = 0$) ∪ (line $y = z = 0$)	$(X), (Y, Z)$

We have a decomposition

$$\mathbb{V}(I) = \bigcup_{P} \mathbb{V}(P)$$

where P runs through the minimal primes over I. (The real picture of the hyperbola seems to have two connected components, but in $\mathbb{C}^2$ the pieces are joined together. Real pictures can be misleading.)

Remarks. (a) The Zariski-closed subsets are the closed subsets of a topology on K^n, the **Zariski topology**, since

(i) $K^n = \mathbb{V}(\emptyset)$ and $\emptyset = \mathbb{V}(K[X_1, \dots, X_n])$,

(ii) $\bigcap_i \mathbb{V}(S_i) = \mathbb{V}(\bigcup_i S_i)$,

(iii) $\mathbb{V}(S) \cup \mathbb{V}(T) = \mathbb{V}(\{st : s \in S, t \in T\})$,

(b) An **affine variety** is a Zariski-closed subset V of K^n together with its **coordinate ring**

$$K[V] := K[X_1, \dots, X_n]/\mathbb{I}(V).$$

We have a bijection

$$V \to \operatorname{Maxspec} K[V], \quad a \mapsto \overline{m_a}.$$

5.4 Krull dimension

Definition. If p is a prime ideal in R then the **height** of p is

$$\mathrm{ht}(p) = \sup\{n : \exists\ p_0 \subset p_1 \subset \cdots \subset p_n = p \text{ with } p_i \text{ distinct prime ideals}\}.$$

and the **Krull dimension** of R is

$$\dim R = \sup\{\mathrm{ht}(p) : p \text{ prime ideal in } R\}$$

$$= \sup\{n : \exists\ p_0 \subset p_1 \subset \cdots \subset p_n \text{ with } p_i \text{ distinct prime ideals}\}.$$

We will see that every prime ideal in a noetherian ring has finite height. There are examples by Nagata of noetherian rings with infinite Krull dimension, but we will see that f.g. algebras over a field have finite Krull dimension.

Example. (1) If K is a field, it has $\dim K = 0$.

(2) A principal ideal domain R which is not a field has $\dim R = 1$, since the prime ideals are either 0 or maximal, of the form (a) with a an irreducible element.

(3) If R is a UFD then the height 1 primes are exactly the ideals (b) with b an irreducible element.

Proof. If p is a height 1 prime, and a is a non-zero element of p, then p is a minimal prime over (a), so $p = (b)$ for some irreducible factor of a by the example at the end of §5.1. On the other hand if p is a prime ideal contained in (b) and a is a nonzero element of p, then (b) is a minimal prime over (a), so $p = (b)$, so (b) has height 1.

(4) A **Dedekind domain** is an integral domain of Krull dimension ≤ 1 which is integrally closed in its field of fractions. Any principal ideal domain is a Dedekind domain. By the next result, the ring of integers $\mathcal{O}_L$ of a number field L is a Dedekind domain.

Theorem. *If $R \subseteq R'$ is an integral extension, then R and R' have the same Krull dimension.*

Proof. Given a chain of prime ideals $p_0 \subset p_1 \subset \cdots \subset p_n$ in R we construct a chain $p'_0 \subset p'_1 \subset \cdots \subset p'_n$ in R' inductively such that p'_i contracts to p_i for all i. First, by Lying over, p_0 is the contraction of some prime ideal p'_0 in R'. If $p'_0, \dots, p'_{i-1}$ have been constructed, then R'/p'_{i-1} contains $(R+p'_{i-1})/p'_{i-1} \cong R/(R \cap p'_{i-1}) \cong R/p_{i-1}$, and it is integral over it by property (iii) of integral extensions. Thus by Lying-over, there is a prime ideal in R'/p'_{i-1} which contracts to p_i/p_{i-1} in R/p_{i-1}. We can write it in the form p'_i/p'_{i-1} for a suitable prime ideal p'_i in R', and then p'_i contracts to p_i.

Conversely, given a chain in R', the contractions give a chain in R, and they are all different. Namely, suppose $(p'_i)^c = (p'_{i-1})^c$. By assumption there is some $a \in p'_i \setminus p'_{i-1}$. Let

$$f(X) = X^n + r_{n-1}X^{n-1} + \dots r_1 X + r_0$$

be a monic polynomial in $R[X]$ of minimal degree with $f(a) \in p'_{i-1}$. It exists by integrality. Then

$$r_0 = f(a) - a^n - r_{n-1}a^{n-1} - \cdots - r_1 a \in p'_i$$

so

$$r_0 \in R \cap p'_i = (p'_i)^c = (p'_{i-1})^c \subseteq p'_{i-1}$$

so

$$(a^{n-1} + r_{n-1}a^{n-2} + \cdots + r_1)a = f(a) - r_0 \in p'_{i-1}.$$

Now by minimality the first factor is not in p'_{i-1}, so $a \in p'_{i-1}$. Contradiction □

Theorem. *If K is a field, then $K[X_1, \dots, X_n]$ has Krull dimension n. Thus any f.g. K-algebra has finite Krull dimension.*

Proof. Induction on n. Say m is a maximal ideal. Then m contains a height one prime (f). By §5.3 Lemma 1, we may suppose f is monic in X_1 over $K[X_2, \dots, X_n]$. Then by §5.3 Lemma 2, $K[X_1, \dots, X_n]/(f)$ is integral over $K[X_2, \dots, X_n]$, so has Krull dimension $n-1$. □

Proposition. *Let R be a noetherian ring. The following are equivalent:*

(i) R has Krull dimension 0, that is, every prime ideal in R is maximal.

(ii) $R/\sqrt{0}$ is a semisimple ring.

(iii) There is an ideal I in R with R/I semisimple and I nilpotent.

(iv) R is artinian, that is, it has the DCC on ideals.

Proof. (i)⇒(ii) The ideal $\sqrt{0}$ is the intersection of the minimal primes over 0, and since R is noetherian, there are only finitely many of them. These primes are maximal, say $m_1, \dots, m_n$. Then

$$R/\sqrt{0} \hookrightarrow (R/m_1) \oplus \dots \oplus (R/m_n).$$

This is a submodule of a semisimple module, so semisimple.

(ii)⇒(iii) If I and J are nilpotent ideals, then so is $I + J$, for if $I^n = J^m = 0$, then

$$(I + J)^{n+m} \subseteq \sum_{i=0}^{n+m} I^i J^{n+m-i} = 0.$$

Now $\sqrt{0}$ is finitely generated, so equal to $(x_1, \dots, x_r)$ for some r, and the ideals (x_i) are nilpotent, hence so is $\sqrt{0}$.

(iii)⇒(iv) Each I^n/I^{n+1} is a f.g. module for R/I, so semisimple. Thus it is artinian. Now use Aufgabe 5.3(ii).

(iv)⇒(i) Replacing R by R/P, with P a prime ideal, we may suppose that R is an integral domain, and need to show it is a field. Let $0 \neq x \in R$. The chain of ideals

$$(x) \supseteq (x^2) \subseteq (x^3) \supseteq \dots$$

stabilizes with $(x^n) = (x^{n+1})$ for some n. Then $x^n = ax^{n+1}$ for some $a \in R$. Then $1 = ax$, so x is invertible with inverse a. □

Lemma (Nakayama's Lemma). *If R is a local ring with maximal ideal m and M is a f.g. module with $mM = M$, then $M = 0$. More generally, if R is a ring, not necessarily commutative, J is its Jacobson radical and M is a f.g. R-module with $JM = M$, then $M = 0$.*

Proof. Suppose $M \neq 0$. Since it is f.g., it has a maximal proper submodule N. Then M/N is a simple module, so $J(M/N) = 0$. Thus $JM \subseteq N$. Contradiction. □

Theorem (Krull's Hauptidealsatz). *If R is a noetherian ring, then any minimal prime ideal over a principal ideal has height ≤ 1.*

Proof. Suppose otherwise. Then there is a minimal prime P over a principal ideal (x), and a chain of distinct prime ideals $p \subset q \subset P$. By passing to R/p and localizing at P we may suppose that R is a local integral domain with maximal ideal m minimal over (x) and containing a nonzero prime ideal q.

Then the only prime ideal in $R/(x)$ is $m/(x)$, so it is artinian. Consider $\theta : R \to R' = R_q$. Since R is an integral domain, so is R', for if $(a/s)(b/t) = 0$ with $a, b \in R$ and $t \in R \setminus q$, then $tab = 0$, and since $t \neq 0$, we have $a = 0$ or $b = 0$.

Let $Q = q^e$ the max ideal in R' and let $I_n = (Q^n)^c$. By the descending chain condition we have $I_n + (x) = I_{n+1} + (x)$ for some n. Suppose $a \in I_n$. We can write $a = b + rx$ with $b \in I_{n+1}$ and $r \in R$. Then in R' we have $\theta(rx) \in (I_n)^e = Q^n$. Now $\theta(x) \notin Q$, for otherwise $x \in Q^c = q$, contrary to the assumption. Since R' is local, with maximal ideal Q, it follows that $\theta(x)$ is invertible. Thus $\theta(r) \in Q^n$. Thus $r \in I_n$. It follows that $I_n = I_{n+1} + I_n x$.

Now $I_n/I_{n+1} = xI_n/I_{n+1} \subseteq mI_n/I_{n+1}$. Thus by Nakayama's Lemma $I_n/I_{n+1} = 0$. Thus $I_n = I_{n+1}$. Thus $Q^n = (I_n)^e = (I_{n+1})^e = Q^{n+1}$. Thus $Q^n = 0$ by Nakayama's Lemma. Thus R' is artinian. But it is an integral domain, so the ideal 0 is a prime ideal, so $Q = 0$. Thus $q \subseteq q^{ec} = Q^c = 0$. Contradiction. □

Corollary (Krull's height theorem)**.** *If R is a noetherian ring, then any minimal prime ideal over an ideal generated by n elements has height $\leq n$. Thus every prime ideal in a noetherian ring has finite height.*

Proof. Induction on n. Let P be a minimal prime over $(x_1, \dots, x_n)$. Replacing R by R_P, we may suppose that P is a maximal ideal. It follows that $P = \sqrt{(x_1, \dots, x_n)}$.

It suffices to show that any element of the set $X = \{p \in \operatorname{Spec} R : p \subset P, p \neq P\}$ has height $< n$. Since R is noetherian, it suffices to show that any maximal element Q of X has $\operatorname{ht}(Q) < n$.

Since P is minimal over $(x_1, \dots, x_n)$, without loss of generality $x_1 \notin Q$. Then P is minimal over $Q + (x_1)$, so $P = \sqrt{Q + (x_1)}$. Thus for $i = 2, \dots, n$, we have $x_i^{n_i} = q_i + r_i x_1$ with $n_i > 0$, $q_i \in Q$ and $r_i \in R$.

Consider the ring $\overline{R} = R/(q_2, \dots, q_n)$. The prime ideals $Q \subset P$ in R give primes $\overline{Q} \subset \overline{P}$ in $\overline{R}$. Now any prime ideal in $\overline{R}$ is of the form $\overline{p}$ for some prime ideal p in R, and if $\overline{p}$ is minimal over $(\overline{x}_1)$, then p contains x_1 and the q_i, so it contains all $x_i^{n_i}$, so it contains all x_i, so $p = P$. Then $\operatorname{ht}(\overline{P}) \leq 1$ since $\overline{P}$ is a minimal prime over $(\overline{x}_1)$, so $\operatorname{ht}(\overline{Q}) = 0$, so Q is a minimal prime over $(q_2, \dots, q_n)$, so by induction $\operatorname{ht}(Q) < n$. □