Modern Algebra
Math 6302/6303

Klaus H. Kaiser

Chapter 1

Algebraic Systems

1.1 Operations, Algebraic Systems

An operation f on the set A is a map $f : A^n \to A$ where n is a natural number $n \geq 0$. The number n is called the *arity* of f.
$n = 1 :$ $f : A \to A$ and f is just a map on A,
$n = 2 :$ $f : A^2 \to A$ and f is a binary operation on A.
The case $n = 0$ deserves special attention: For any set S one has that $A^S = \{\alpha | \alpha : S \to A\}$. In particular, for $n = \{0, 1, \ldots, n-1\}$,

$$A^n = \{\alpha | \alpha : \{0, 1, \ldots, n-1\} \to A\} = \{\alpha = (a_0, \ldots, a_{n-1}) | a_\nu \in A\}$$

and this is the set of all $n-tuples$ $(a_0, \ldots, a_{n-1})$ of elements in A.
An $n-ary$ operation assigns to any $n-tuple$ $(a_0, \ldots, a_{n-1})$ of elements in A an element $f(a_0, \ldots, a_{n-1}$ as operation value.
Now, $A^\emptyset = \{\alpha | \alpha : \emptyset \to A\} = \{\emptyset\}$. A *nullary operation* assigns to the *empty map* an element $a \in A$. A nullary operation is therefore called a *constant*.
For $A = \emptyset$ one has that $A^\emptyset = \{\emptyset\}$ and $A^n = \emptyset$ for $n > 0$. That is, only operations of *positive* arity exist.

An *algebraic system* consists of

1. a set A,
2. a family (f_t) of $n_t - ary$ operations on A.

We use the notation

$$\mathbf{A} = (A, (f_t)_{t \in T})$$

and $\Delta = (n_t)_{t \in T}$ is called the *similarity type* of $\mathbf{A}$.

Examples

1. $\mathbb{Z} = (Z, (f_1, f_2, f_3))$ where $f_1 : Z^2 \to Z,\ (n, m) \mapsto n + m,\ f_2 : Z \to Z,\ x \mapsto -x,\ f_3 : \{\emptyset\} \to Z,\ \emptyset \mapsto 0$. The binary operation f_1 is the *addition*, the unary operation f_2 is the *additive inverse* and f_3 is the *zero*.

   We use for binary operations most of the time the symmetric notation xfy instead of $f(x, y)$.

2. Let S be any set. Map(S) is the algebra of all maps of S into itself with composition $\circ$ as a binary operation and id_S as a nullary operation. Thus we get the algebraic structure

$$\text{Map(S)} = (\{\varphi|\varphi : S \to S\},\ \circ,\ id_S),\ \circ(\phi, \psi) = \phi \circ \psi$$

3. Similarly, as before, we define

$$\text{Bij}(S) = (\{\varphi|\varphi : S \xrightarrow{bij} S\},\ \circ,\ ^{-1},\ id_S\})$$

The familiar algebras of *integers*, *real numbers* and *complex numbers*

$$\mathbb{Z} = (Z, +, -, 0, \cdot, 1)$$

$$\mathbb{R} = (R, +, -, 0, \cdot, 1)$$

$$\mathbb{C} = (C, +, -, 0, \cdot, 1)$$

are algebras of the same type $\Delta = (2, 1, 0, 2, 0)$. They are *similar*.
Generalizations of the concept of an algebraic system are:

partial algebras: The operations are not everywhere defined. On the set of real numbers, $^{-1} : \mathbb{R}\backslash\{0\} \to R,\ x \mapsto {}^1\!/_x$ can be added as a partial unary operation.

infinitary algebras: These are algebras with operations of infinitary arity. Taking the limit of an infinite sequence may be considered as a partial infinitary operation, the operation is defined only for convergent sequences. The *projections* $p_{i_0} : A^I \to A,\ (a_i)_{i\in I} \mapsto a_{i_0}$, are infinitary operations if I is infinite.

multi-valued algebras: The operation values are subsets of A, i.e., $f : A^n \to \mathcal{P}(A),\ f(a_1, \ldots, a_n) \subseteq \mathcal{P}A$. One has that $card(f(a_1, \ldots, a_n)) = 1$ iff f is an operation and $card(f(a_1, \ldots, a_n)) \leq 1$ iff f is a partial operation.

relational algebras : A relational system is a set A together with a family $(f_t)_{t\in T}$ of $n_t - ary$ operations and a family $(R_s)_{(s\in S)}$ of $m_s - ary$ relations R_s where one has that each $R_s \subseteq A^{m_s}$. An $n - ary$ operation f corresponds uniquely to an $(n + 1) - ary$ relation R_f :

$$R_f = \{(a_1, \ldots, a_n, a_{n+1})|a_{n+1} = f(a_1, \ldots, a_n)\} = graph(f)$$

An Example of a relational systems is

$$\mathbb{Z} = (Z, +, -, 0, \cdot, 1, \leq)$$

Any algebraic system may be considered as a relational system where $S = \emptyset$. Also, finitary partial algebras are relational systems.
Note that the operations of the algebraic system as well as the relations of a relational system are *finitary* but that the number of operations of an algebra **A** is in general infinite.

Example. Let $\mathbb{V}$ be a vector space over the field of real numbers. Then multiplication of a vector v by a real number α may be perceived as a unary operation and we have for each $\alpha \in \mathbb{R}$ the unary operation $v \mapsto \alpha.v$. As an algebraic system such a vector space looks like

$$\mathbb{V} = (V, +, -, 0,\ (\alpha.)_{\alpha\in R})$$

1.2 Homomorphisms of Algebras, Subalgebras and Direct Products

Let $\mathbf{A}$ and $\mathbf{B}$ be algebras of the same type Δ:

$$\mathbf{A} = (A,\ (f_t)_{t\in T}),\quad \mathbf{B} = (B, (g_t)_{t\in T})$$

where the arity of f_t is equal to the arity of g_t, $t \in T$.
A map $\varphi : \mathbf{A} \to \mathbf{B}$ of the underlying set A of $\mathbf{A}$ to the underlying set B of $\mathbf{B}$ is called a *homomorphism* if for every $t \in T$ one has that:

$$\varphi(f_t(a_0, \ldots, a_{n-1})) = g_t(\varphi(a_0), \ldots, \varphi(a_{n-1}))$$

This means for a binary operation, e.g., multiplication, which is denoted in both algebras by " $\cdot$ ": $\varphi(a_0 \cdot a_1) = \varphi(a_0) \cdot \varphi(a_1)$
For a unary operation, e.g., a multiplicative inverse $^{-1}$, this reads as $\varphi(a^{-1}) = (\varphi(a))^{-1}$
For a nullary operation, say $a_t = a$, and $b_t = b$ this is: $\varphi(a) = b$

Let for a homomorphism $\varphi : \mathbf{A} \to \mathbf{B}$, E_φ be the equivalence that is induced by φ, that is,

$$a_1 \sim a_2\ mod(E_\varphi)\ if\ and\ only\ if\ \varphi(a_1) = \varphi(a_2)$$

An equivalence relation E on an algebra $\mathbf{A}$ is a *congruence relation* if

$$a_1 \sim a_1',\ a_2 \sim a_2', \ldots,\ a_{n_t} \sim a_{n_t}'\ yields\ f_t(a_1, \ldots, a_{n_t}) \sim f_t(a_1', \ldots, a_{n_t}')$$

Proposition 1.1 *Let $\varphi : \mathbf{A} \to \mathbf{B}$ be a homomorphism of algebras. Then the equivalence for φ, $E\varphi = \{(a_1, a_2) | \varphi(a_1) = \varphi(a_2)\}$, is a congruence relation on $\mathbf{A}$*

Proof. Assume $a_1 \sim a_1',\ a_2 \sim a_2', \ldots,\ a_{n_t} \sim a_{n_t}'$. This is $\varphi(a_1) = \varphi(a_1'), \ldots, \varphi(a_{n_t}) = \varphi(a_{n_t}')$. Therefore,

$$\varphi(f_t(a_1, \ldots a_{n_t})) = g_t(\varphi(a_1), \ldots, \varphi(a_{n_t})) = g_t(\varphi(a_1'), \ldots, \varphi(a_{n_t}')) = \varphi(f_t(a_1', \ldots, a_{n_t}')).$$

This is, $f_t(a_1, \ldots, a_{n_t}) \sim f_t(a_1', \ldots, a_{n_t}')$. □

Proposition 1.2 *Let E be a congruence relation on the algebra $\mathbf{A}$. Then there is exactly one algebraic structure $(\bar{f}_t)$ on the set A/E of equivalence classes for E such that the canonical projection $q_E : \mathbf{A} \to \mathbf{A}/\mathbf{E}$ becomes a homomorphism.*

Proof. We first show uniqueness. If $q_E : a \to \bar{a} = [a]_E$ is homomorphic then

$$\bar{f}_t(C_1, \ldots, C_{n_t}) = \bar{f}_t([a_1], \ldots, [a_{n_t}]) = \bar{f}_t(q_E(a_1), \ldots, q_E(a_{n_t})) = q_E(f_t(a_1, \ldots, a_{n_t})).$$

Therefore, if $C_1 = [a_1], \ldots, C_{n_t} = [a_{n_t}]$ then, necessarily,

$$\bar{f}_t(C_1, \ldots, C_{n_t}) = [f_t(a_1, \ldots, a_{n_t})].$$

Assume now for the congruence E on $\mathbf{A}$ that $\bar{f}_t(C_1, \ldots, C_{n_t}) = \bar{f}_t([a_1], \ldots, [a_{n_t}]) = [f_t(a_1, \ldots, a_{n_t})]$ then, because E is a congruence, the choice of $a_1 \in C_1, \ldots, a_{n_t} \in C_{n_t}$ does not matter, $\bar{f}_t$ is properly defined by *means of representatives*. The map $q_E : \mathbf{A} \to (A/E,\ (\bar{f}_t))$ is homomorphic:

$$q_E(f_t(a_1, \ldots, a_{n_t})) = \bar{f}_{n_t}(q_E(a_1), \ldots, q_E(a_{n_t}))$$

is true by the definition of $\bar{f}_t$. □

Proposition 1.3 *Let $\varphi : \mathbf{A} \to \mathbf{B}$ be a homomorphism between algebras with associated congruence $E = E_\varphi : a \sim a'$ iff $\varphi(a) = \varphi(a')$. Then there is a unique injective homomorphism $\dot{\varphi} : \mathbf{A}/\mathbf{E} \to \mathbf{B}$ such that $\dot{\varphi} \circ q_E = \varphi$.:*

$$\mathbf{A} \xrightarrow{q_E} \mathbf{A}/E \xrightarrow{\dot{\varphi}} \mathbf{B} = \mathbf{A} \xrightarrow{\varphi} \mathbf{B}$$

PROOF. We only have to show that $\dot{\varphi}$ is homomorphic. But:
$\dot{\varphi}(\bar{f}_t(C_1, \ldots, C_{n_t})) = \dot{\varphi}(\bar{f}_t(\bar{a}_1, \ldots, \bar{a}_{n_t})) = \dot{\varphi}(q_{E_\varphi}(f_t(a_1, \ldots, a_{n_t})) = \varphi(f_t(a_1, \ldots, a_{n_t}))$
$= g_t(\varphi(a_1), \ldots, \varphi(a_{n_t})) = g_t((\dot{\varphi} \circ q_{E_{)\varphi}})(a_1), \ldots, (\dot{\varphi} \circ q_{E_{)\varphi}})(a_{n_t})) = g_t(\dot{\varphi}(C_1), \ldots, \dot{\varphi}(C_{n_t}))$. □

A bijective homomorphism is called an *isomorphism*. Let $\varphi : \mathbf{A} \to \mathbf{B}$ be an isomorphism between algebras. Then

$$\varphi^{-1} : B \to A$$

is homomorphic:

$$\begin{aligned} \varphi^{-1}(g_t(b_1, \ldots, b_{n_t})) &= f_t(\varphi^{-1}(b_1), \ldots, \varphi^{-1}(b_{n_t})) \text{ iff} \\ \varphi(f_t(\varphi^{-1}(b_1), \ldots, \varphi^{-1}(b_{n_t})) &= g_t(b_1, \ldots, b_{n_t}) \text{ iff} \\ g_t(\varphi(\varphi^{-1}(b_1), \ldots, \varphi(\varphi^{-1}(b_{n_t}))) &= g_t(b_1, \ldots, b_{n_t}) \end{aligned}$$

Corollary 1.4 (Homomorphism Theorem) *Let $\varphi : \mathbf{A} \to \mathbf{B}$ be a surjective homomorphism between algebras. Then $\mathbf{B}$ is isomorphic to a factor algebra of $\mathbf{A}$.*

Note, that the composition of homomorphisms is a homomorphism and that the identity on an algebra is a homomorphism. Thus, the class of algebras of similarity type Δ is a *category* with the algebras as objects (dots) and the homomorphisms as morphisms (arrows).

It is very easy to see that for every algebra $\mathbf{A}$, the intersection of congruence relations is a congruence. Recall that an ordered set $(P, \leq)$ is a *complete lattice* if every subset S of P has a largest lower bound (or a smallest upper bound). Thus we have

Proposition 1.5 *For every algebra $\mathbf{A}$, the congruence relations on $\mathbf{A}$ form a complete lattice: Con($\mathbf{A}$).*

$\Delta = \{(a, a) | a \in A\}$ is the smallest congruence on $\mathbf{A}$ and $\mathbf{A}/\Delta \cong \mathbf{A}$. $\mathbf{A} \times \mathbf{A}$ is the largest congruence on $\mathbf{A}$ and one has that $\mathbf{A}/A \times A \cong 1_\Delta$, i.e., the one-element algebra of type Δ.

Let $\mathbf{A}$ be an algebra and let C be a subset of $\mathbf{A}$. C is said to be closed if

$$a_1, \ldots, a_n \in C \text{ implies that } f_t(a_1, \ldots, a_{n_t}) \in C$$

In particular, a closed subset contains all constants.

Proposition 1.6 *Let C be a closed subset of the algebra $\mathbf{A}$. Then there is exactly one algebraic structure (g_t) on C such that the inclusion*

$$i_{C,A} : \mathbf{C} \hookrightarrow \mathbf{A}$$

is homomorphic.

PROOF. We first prove uniqueness:

$$i(g_t(a_1,\ldots,a_{n_t})) = f_t(i(a_1),\ldots,i(a_{n_t})) = f_t(a_1,\ldots,a_{n_t}) = g_t(a_1,\ldots,a_{n_t})$$

That is,

$$g_t = f_t|C^{n_t}$$

On the other hand, for any closed subset C of $\mathbf{A}$ we may define

$$g_t(a_1,\ldots,a_{n_t}) = f_t(a_1,\ldots,a_{n_t}) \qquad (a_1,\ldots,a_{n_t}) \in C$$

and $i : \mathbf{C} \hookrightarrow \mathbf{A}$ is homomorphic. □

Any injective homomorphism is called an *embedding.* A closed subset becomes an algebra such that the inclusion is an embedding. Any closed subset C with the operations f_t restricted to C is a subalgebra. It is easy to see that for any algebra $\mathbf{A}$, the intersection of closed subsets is closed.

Proposition 1.7 *For every algebra* $\mathbf{A}$*, the subalgebras of* $\mathbf{A}$ *form a complete lattice:* $Sub(\mathbf{A})$.

$\mathbf{A}$ is the largest element of $Sub(\mathbf{A})$ and the smallest element of $\mathbf{A}$ is $\emptyset$ if there are no constants, otherwise it is the subalgebra $\mathbf{C}_0$ that is generated by all constants:

$$\mathbf{C}_0 = \bigcap_{\mathbf{C}\in Sub\mathbf{A})} \mathbf{C} = \bigcap_{\substack{\mathbf{C}\supset M_0 \\ \mathbf{C}\in Sub(\mathbf{A})}} \mathbf{C} \text{ where } M_0 = \{c_{t_0} = f_{t_0}|t_0 \in T \textit{ and } n_{t_0} = 0\}$$

Let M_0 be any subset of the algebra $\mathbf{A}$. Then define:

$$\mathcal{C}_{\mathbf{A}}(M_0) = \bigcap_{\substack{\mathbf{C}\in Sub(\mathbf{A}) \\ C\supseteq M_0}} C$$

Then $\mathcal{C}_{\mathbf{A}}(M_0)$ is the smallest subalgebra of $\mathbf{A}$ that contains M_0.

For any set M and n_t-ary operation f_t of $\mathbf{A}$ define:

$$f_t(M^{n_t}) = \{a|a \in A, \quad a = f_t(a_1,\ldots,a_{n_t}), \quad a_1,\ldots,a_{n_t} \in M\}$$

M_0, $M_1 = \bigcup_{t\in T} f_t(M_0^{n_t}) \cup M_0$, $M_2 = \bigcup_{t\in T} f_t(M_1^{n_t}) \cup M_1, \ldots$ yields

$$M_0 \subseteq M_1 \subseteq M_2 \subseteq \ldots$$

Then let $\bar{M} = \bigcup_{\nu} M_\nu$. We notice:

- $\bar{M}$ is closed.

  Let $a_1,\ldots a_{n_t} \in \bar{M}$. Then there is some k such that $a_1,\ldots a_{n_t} \in M_k$. This implies that $f_t(a_1,\ldots a_{n_t}) \in M_{k+1} \subseteq \bar{M}$

- $\bar{M}$ contains M_0.

- Let $\mathbf{C}$ be any subalgebra containing M_0. Then $\mathbf{C}$ contains $\bar{M}$.

  We have $\mathbf{C} \supseteq M_0$ and assume that $\mathbf{C} \supseteq M_k$. Let $a \in M_{k+1}$. Then $a \in M_k$ or one has that $a = f_t(a_1,\ldots,a_{n_t})$ with $a_1,\ldots,a_{n_t} \in M_k$. In both cases we get that $a \in \mathbf{C}$.

Hence, $\bar{M} = \mathcal{C}(M)$.

Proposition 1.8 *$\mathcal{C} = \mathcal{C}_{\mathbf{A}}$ defines an algebraic closure for subsets M of the algebra $\mathbf{A}$. That is:*

- *$\mathcal{C}$ is extensive, i.e., $\mathcal{C}(M) \supseteq M$*
- *$\mathcal{C}$ is monotone, i.e., If $M_1 \subseteq M_2$ then $\mathcal{C}(M_1) \subseteq \mathcal{C}(M_2)$*
- *$\mathcal{C}$ is idempotent, i.e., $\mathcal{C}(\mathcal{C}(M)) = \mathcal{C}(M)$.*

Moreover, $a \in \mathcal{C}(M)$ if and only if there is a finite subset F of M such that $a \in \mathcal{C}(F)$:

$$\mathcal{C}(M) = \bigcup_{\substack{F \subseteq M \\ F \text{ finite}}} \mathcal{C}(F$$

Proof. It is quite obvious that $\mathcal{C}$ satisfies the properties of a closure operator. We only have to show the algebraicity of $\mathcal{C}$. We need to show that:

$$\bigcup_{\substack{F \subseteq M \\ F \text{ finite}}} \mathcal{C}(\mathcal{F})$$

is closed. Let $a_1, \ldots, a_{n_t} \in \bigcup_{\substack{F \subseteq M \\ F \text{ finite}}} \mathcal{C}(F)$. We then have that:

$$a_1 \in \mathcal{C}(F_1), \ldots, a_{n_t} \in \mathcal{C}(F_{n_t}) \text{ implies that } a_1, \ldots, a_{n_t} \in \mathcal{C}(F_1 \cup \ldots \cup F_{n_t}) = \mathcal{C}(F)$$

where, of course, F is finite. But then:

$$f_t(a_1, \ldots, a_{n_t}) \in \mathcal{C}(F) \subseteq \bigcup_{\substack{F \subseteq M \\ F \text{ finite}}} \mathcal{C}(F)$$

Now, $\bigcup \mathcal{C}(F)$ is a closed subset that contains M. This is clear because for any element $a \in M$ one has that $a \in \mathcal{C}\{a\} \subseteq \bigcup \mathcal{C}(F)$ and therefore

$$M \subseteq \bigcup \mathcal{C}(F) \text{ and therefore } \mathcal{C}(M) \subseteq \bigcup \mathcal{C}(F)$$

The converse inclusion is of course obvious. □

Proposition 1.9 *Let $(\mathbf{A}_i)_{i \in I} = (A_i, (f_i^t)_{t \in T})_{i \in I}$ be a family of similar algebraic systems. Let*

$$A = \prod_{i \in I} A_i = \{\alpha | \alpha : I \to \bigcup_{i \in I} A_i,\ \alpha(i) \in A_i\}$$

be the cartesian product of the carrier sets A_i of the algebras $\mathbf{A}_i$. Then there is exactly one algebraic structure $(f_t)_{t \in T}$ on the set A such that all projections $p_i : A \to A_i,\ \alpha \mapsto \alpha(i)$, are homomorphic.

Proof. We first have to show uniqueness. Let $\alpha_1, \ldots, \alpha_{n_t} \in A$ and let $\alpha = f_t(\alpha_1, \ldots, \alpha_{n_t})$. Then one has that:

$$\alpha(i) = p_i(\alpha) = p_i(f_t(\alpha_1, \ldots, \alpha_{n_t})) = f_t^i(p_i(\alpha_1), \ldots, p_i(\alpha_{n_t})) = f_t^i(\alpha_1(i), \ldots, \alpha_{n_t}(i))$$

and that is

$$\alpha = (\alpha(i))_{i \in I} = (f_t^i(\alpha_1(i), \ldots, \alpha_{n_t}(i))_{i \in I}$$

This yields,

$$f_t(\alpha_1, \ldots, \alpha_{n_t}) = (f_t^i(\alpha_1(i), \ldots, \alpha_{n_t}(i)))_{i \in I}$$

On the other hand, if we define f_t on the cartesian product A by this last formula, then

$$p_i(f_t(\alpha_1, \ldots, \alpha_{n_t})) = f_t^i(\alpha_1(i), \ldots, \alpha_{n_t}(i)) = f_t^i(p_i(\alpha_1), \ldots, p_i(\alpha_{n_t}))$$

shows that the projections p_i are homomorphisms. □

Note that we have in particular

$$c_t = (c_t^i)_{i \in I} \text{ for nullary operations } c_t$$

For any algebra $\mathbf{A}$, and any set S, $\mathbf{A}^S$ is called the *direct power* of $\mathbf{A}$. The plane $\mathbb{R}^2$ is a typical example of the second power of the vector space $\mathbb{R}$.

Proposition 1.10 *Let $\varphi_i : \mathbf{B} \to \mathbf{A}_i,\ i \in I$, be an initial family of homomorphisms. Then there is exactly one homomorphism $\varphi : \mathbf{B} \to \mathbf{A} = \prod_{i \in I} \mathbf{A}_i$ such that $p_i \circ \varphi = \varphi_i,\ i \in I$.*

Proof. One defines

$$\varphi = \prod_{i \in I} \varphi_i : \mathbf{B} \to \mathbf{A},\ \varphi(b) = (\varphi_i(b))_{i \in I}.\ \square$$

Definition 1 [G. Birkhoff] A class $\mathcal{P}$ of algebras is called *primitive* if it is closed

- under the formation of direct products;
- under taking homomorphic images;
- under taking subalgebras.

Groups and rings are examples of primitive classes. However, fields are not closed under direct products.

Definition 2 [A. I. Mal'cev] A class $\mathcal{Q}$ of algebras is *quasi-primitive* if it is closed under

- isomorphic copies;
- direct products.

Cancellation semigroups are an example of a quasi-primitive class that is not primitive.

Remarks on Ordered Sets and Lattices

A relational system $(O, \leq)$ is called an *ordered set* if the relation $\leq$ is (i) *reflexive*, (ii) *transitive* and if (iii) *anti-symmetry* holds. That is:

(i) $x \leq x$ holds for all $x \in O$.

(ii) If $x \leq y$ and $y \leq z$ then $x \leq z$.

(iii) If $x \leq y$ and $y \leq x$ then $x = y$.

For $x \neq y$ but $x \leq y$ one writes $x < y$. Also, $x \leq y$ means the same as $y \geq x$.
An ordered set is *totally ordered* if one also has *trichotomy*:

(iv) Either $x \leq y$ or $y \leq x$ or $x = y$.

Prominent examples for ordered sets are the natural numbers with divisibility: $(N, |)$, and the power set $(\mathcal{P}(S), \subseteq)$ for the set S. The set of real numbers with their ordinary ordering is the prototype of a totally ordered set. Very often, ordered sets are called *partially ordered*, and totally ordered sets are called *linearly ordered*. A relation that is reflexive and transitive is called a *quasi-ordering*. The integers with division are an example: $n|m$ and $m|n$ only yields $n = \pm m$.

If S is a subset of the ordered set O then u is called an *upper bound* for S if $u \geq s$ holds for all $s \in S$. An upper bound that actually belongs to S is called the *maximum* of S. A maximum, if it exists, is, of course, by anti-symmetry, unique. If u is an upper bound for S and if $v \geq u$ then v is an upper bound for S. That is, the upper bounds for a subset S for the ordered set O form an *upper end* of O. *Lower bounds* and *minima* are defined similarly. By "default", every element of O is an upper, as well a lower bound, of the empty set. If the set of upper bounds for S has a minimum, then this minimum is called the *supremum of* S or *the least upper bound of* S:

$$\sup(S) = \min\{u | u \text{ is an upper bound for } S\}$$

Similarly, the *infimum of* S, or, *largest lower bound*, is the maximum of all lower bonds for S. An ordered set is called a *complete lattice* if every subset S has a supremum as well an infimum. By the very definition, the infimum of the empty set must be, if it exists, the maximum of O. Similarly, the supremum of the empty set must be, if it exists, the minimum of O. Only for the empty set the infimum may be greater than the supremum. An ordered set is *bounded* if it has a maximum as well a minimum. The following proposition is an easy but useful fact. The proof is an easy exercise.

Proposition 1.11 *Assume that every subset S of the ordered set O has an infimum. Then every subset S of O has a supremum.*

An ordered set O is called a *complete lattice* if every subset S has an infimum (and then as well) a supremum. The power set $\mathcal{P}(S)$ of a set S is the prototype of a complete lattice. Here the infimum of a subset, i.e.,a collection $\mathcal{C}$ of subsets of S, is the intersection of $\mathcal{C}$ and the supremum is the union of $\mathcal{C}$. A collection $\mathcal{C}$ of subsets of S is called a *closure system* of S if the intersection of every sub-collection $\mathcal{S}$ of $\mathcal{C}$ belongs to $\mathcal{C}$. As a corollary to the last proposition we state:

Proposition 1.12 *Let $\mathcal{T}$ be a closure system of subsets of the set S. Then $(\mathcal{T}, \subseteq)$ is a complete lattice.*

An ordered set is a *lattice* if every finite non-empty subset has an infimum as well a supremum. The natural numbers with divisibility form an example of a complete lattice. The infimum of a set of natural numbers is the greatest common divisor and the supremum is the lowest common multiple. The number 0 is the maximum of $(N, |)$, and 1 is the minimum of $(N, |)$.
An ordered set is called *well ordered* if every subset has a minimum. The natural numbers are well ordered by $\leq$. It is an axiom of set theory that every set can be well ordered. This statement is equivalent to the **Axiom of Choice**. The axiom of choice is also equivalent to

Zorn's Lemma: *Assume that every totally ordered subset C of the ordered set $(O, \leq)$ has an upper bound in O. Then O has a maximal element.*

An element m of O is *maximal* if $s \geq m$ implies that $s = m$.

Finally, we remark that if R is an ordering the dual relation aR^*b iff bRa is also an ordering. Instead of saying that a certain map $\varphi : A \to B$ is an order reversing isomorphism from A to B we say that it is an order isomorphism from A to B^*. Note, that φ is an order isomorphism if it is bijective and φ as well as φ^{-1} are order preserving.

Chapter 2

Basic Facts about groups, rings and fields. Modules and vector spaces.

2.1 Groups

A *groupoid* is an algebraic system with just one binary operation: $(A, \cdot)$.

A groupoid becomes a *semigroup* if the operation is associative, i.e., one has for all $x, y, z \in A$:

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z$$

If $(A, \cdot)$ is a groupoid, then $e \in A$ is called a *unit* for " $\cdot$ " if one has for all $x \in A$

$$x \cdot e = e \cdot x = x$$

A groupoid can have at most one unit. Assume that e and e' are units. We then have $e' \cdot e = e$ because e' is a *left-unit* and $e' \cdot e = e$ because e is a *right-unit*. Hence, $e = e'$.
A semigroup with unit is called a *monoid*: $(S, \cdot, e)$ where we consider the unit e as a nullary operation.

Let $(S, \cdot, e)$ be a monoid. Then $x \in S$ is called *invertible* if there is an $x' \in S$ such that

$$x \cdot x' = x' \cdot x = e$$

An element x of a monoid can have at most one inverse. Assume that x' is a *left-inverse* and that x" is a *right-inverse* of x, i.e., $x' \cdot x = e$ and $x \cdot x$" $= e$. But then:

$$x' = x' \cdot e = x' \cdot (x \cdot x") = (x' \cdot x) \cdot x" = e \cdot x" = x"$$

The inverse of an element x in a monoid, if it exists, is denoted as x^{-1}.

A *group* is a monoid where every element has an inverse.

Therefore, the algebraic system $\mathbf{G} = (G, \cdot, {}^{-1}, e)$ is a group if

- the binary operation $\cdot$ is associative:

$$(x \cdot y) \cdot z = x \cdot (y \cdot z)$$

- The constant e is a unit for $\cdot$:
$$x \cdot e = e \cdot x = x$$
- The unary operation $^{-1}$ associates for every $x \in G$ the inverse:
$$x \cdot x^{-1} = x^{-1} \cdot x = e$$
- A group is *abelian* if the binary operation $\cdot$ is commutative:
$$x \cdot y = y \cdot x$$

A commutative group is also called a *module*. The binary operation then is denoted as $+$ and called *addition*. The unit is called *zero* and $-$ is the operation that takes the additive inverse.

Examples

1. $\mathbb{Z} = (Z, +, -, 0)$ is a module.
2. $\mathbb{R}^* = (R \setminus \{0\}, \cdot, ^{-1}, 1)$ is a group, i.e., the *multiplicative* group of the reals.
3. $\mathbb{N} = (N, +, 0)$ and $\mathbb{N} = (N, \cdot, 1)$ are monoids.

Theorem 2.1 *The class of all semigroups is a primitive class of algebras. The same is true for the class of all monoids, the class of all groups and the class of abelian groups.*

Proof. Let C be a closed subset of the semigroup $(S, \cdot)$. Then $\mathbf{C} = (C, \cdot)$ is obviously a semigroup.

Let $\varphi : (C, \cdot) \to (A, \cdot)$ be a surjective homomorphism from a semigroup to a groupoid. Let $a, b, c \in A$. Then $a = \varphi(x),\ b = \varphi(y),\ c = \varphi(z)$ and

$$\begin{aligned}(a \cdot b) \cdot c &= (\varphi(x) \cdot \varphi(y)) \cdot \varphi(z) = \varphi(x \cdot y) \cdot \varphi(z) = \varphi((x \cdot y) \cdot z) = \varphi(x \cdot (y \cdot z)) \\ &= \varphi(x) \cdot \varphi(y \cdot z) = \varphi(x) \cdot (\varphi(y) \cdot \varphi(z)) \\ &= a \cdot (b \cdot c)\end{aligned}$$

Let $(A_i, \cdot)_{i \in I}$ be a system of groupoids. All operations are denoted by $\cdot$. Let $\alpha, \beta \in \prod_{i \in I} A_i$. Recall that $\alpha : I \to \bigcup_{i \in I} A_i,\ \alpha(i) \in A_i$. We also write $\alpha = (\alpha(i))_{i \in I}$ and consider the function α as an *I-tuple*where the i^{th} coordinate is in A_i:

$$\alpha = (\alpha(i))_{i \in I} = (a_i)_{i \in I}, \quad a_i = \alpha(i) = p_i(\alpha)$$

The operation on $\prod_{i \in I} A_i$ was defined by

$$p_i(\alpha \cdot \beta) = p_i(\alpha) \cdot p_i(\beta),\ \textit{i.e.},\ (\alpha \cdot \beta)(i) = \alpha(i) \cdot \beta(i),\ i \in I :\ \alpha \cdot \beta = (\alpha(i) \cdot \beta(i))_{i \in I},$$

that is:

$$(a_i)_{i \in I} \cdot (b_i)_{i \in I} = (a_i \cdot b_i)_{i \in I}$$

Thus, in case of associativity for all A_i:

$$\begin{aligned}((a_i)_{i\in I}\cdot(b_i)_{i\in I})\cdot(c_i)_{i\in I} &= (a_i\cdot b_i)_{i\in I}\cdot(c_i)_{i\in I}\\ &= ((a_i\cdot b_i)\cdot c_i)_{i\in I} = (a_i\cdot(b_i\cdot c_i))_{i\in I} = (a_i)_{i\in I}\cdot(b_i\cdot c_i)_{i\in I}\\ &= (a_i)_{i\in I}\cdot((b_i)_{i\in I}\cdot(c_i)_{i\in I})\end{aligned}$$

Therefore, semigroups form a primitive class.

If $(S_i,\cdot,e_i)$ are all monoids, then $e=(e_i)_{i\in I}$ is a nullary operation of $\prod S_i$ and:

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

shows that e is a right-unit and a similar argument shows that e is a left-unit.

If all $(G_i,\cdot,{}^{-1},e_i)$ are groups then

$$(a_i)_{i\in I}\cdot(a_i^{-1})_{i\in I} = (a_i\cdot a_i^{-1})_{i\in I} = (e_i)_{i\in I} = e$$

shows that $\mathbf{G}=\prod_{i\in I} G_i$ is a group.

We got that monoids and groups are closed under direct products. Closure under subalgebras is also easy to see. Notice, that a closed subset C must contain the unit e, and in a group the inverse of an element $c\in C$. We are going to show that monoids are closed under homomorphic images. Let $\varphi:(S,\cdot,e)\to(A,\cdot,e')$ be a surjective homomorphism from the semigroup $\mathbf{S}$ onto the groupoid $\mathbf{A}$. We already know that $\mathbf{A}$ is a semigroup. We need to show that e' is the unit of $\mathbf{A}$. But: $\varphi(e)=e'$ and $a\cdot e'=\varphi(x)\cdot\varphi(e)=\varphi(x\cdot e)=\varphi(x)=a$ shows that e' is a left-unit for every $a\in\mathbf{A}$ and similarly, e' is a right-unit for $\mathbf{A}$. Thus the class of monoids is primitive. If $\varphi:(G,\cdot,{}^{-1},e)\to(A,\cdot,f,e')$ is a surjective homomorphism from a group $\mathbf{G}$ onto a similar algebra $\mathbf{A}$, then we already know that $\mathbf{A}$ is a monoid. We wish to show that f is the inverse operation of $\mathbf{A}$. But if $a=\varphi(g)$ then:

$$a\cdot f(a)=\varphi(g)\cdot f(\varphi(g))=\varphi(g)\cdot\varphi(g^{-1})=\varphi(g\cdot g^{-1})=\varphi(e)=e'$$

shows that $\mathbf{A}$ is a group. Thus the class of all groups is primitive.

All proofs were very similar and showed that *equations* are preserved under taking subalgebras, homomorphic images and direct products. Thus, for example, abelian groups are a primitive class. □

Proposition 2.2 *Let $\mathbf{S}=(S,\cdot,e)$ be a monoid. Then the set of all $x\in S$ which have an inverse is a closed subset and therefore a group.*

PROOF. Let $Inv(\mathbf{S})=\{x|x \text{ is invertible in } S\}$. We then have

- $e\in Inv(\mathbf{S})$: $e\cdot e=e$
- If $x\in Inv(\mathbf{S})$ and $y\in Inv(\mathbf{S})$ then $x\cdot x'=x'\cdot x=e$ for a unique $x'\in S$ and $y\cdot y'=y'\cdot y=e$ for a unique $y'\in S$. But then (omitting the operation symbol $\cdot$):

  $$(xy)(y'x')=x(yy')x'=xex'=xx'=e$$

  Thus, $y'x'$ is a right-inverse of xy and a similar argument shows that it is a left-inverse. Hence, $x\cdot y\in Inv(\mathbf{S})$.
- If x has inverse x' then x' has inverse x.

Thus $Inv(\mathbf{S})$ is a group. □

Examples

1. Let S be any set. Then $Map(S) = (\{\varphi : S \to S\},\ \circ,\ id_S)$ is a monoid and $Inv(Map(S) = Bij(S)$ is the group of invertible maps on S.
2. Let $S = \{1, \ldots, n\}$. Then $Bij(S) = \mathbf{S}_n$ is called the *symmetric* group of *degree* n.

Let $\varphi : \mathbf{G} \to \mathbf{G}'$ be a homomorphism between groups. Then

$$\varphi(x_1) = \varphi(x_2) \text{ iff } \varphi(x_2)^{-1}\varphi(x_1) = e_{G'} \text{ iff } \varphi(x_2^{-1}x_1) = e_{G'} \text{ iff } x_2^{-1}x_1 \in \varphi^{-1}(e_{\mathbf{G}'}) = \mathbf{N}_\varphi$$

Therefore,

$$x_1 \sim x_2 \ mod(\varphi) \text{ iff } x_2^{-1}x_1 \in \mathbf{N}_\varphi$$

$\mathbf{N}_\varphi = \mathbf{N}$ is called the *kernel* of φ and has the properties of a *normal* subgroup of $\mathbf{G}$:

- $\mathbf{N}$ is a subgroup of $\mathbf{G}$.

  This is obvious because $\mathbf{N}$ is the counter image of the subgroup $e_{\mathbf{G}'}$ of $\mathbf{G}'$.

- $\mathbf{N}$ is *normal*. That is:

  $$\text{If } x \in \mathbf{N} \text{ and } y \in \mathbf{G} \text{ then } y^{-1}xy \in \mathbf{N}$$

  This is also obvious. If x is mapped to e' then $y^{-1}xy$ is mapped to e'.

Let $\mathbf{H}$ be any subgroup of $\mathbf{G}$. Then:

$$x_1 \sim_L x_2 \text{ iff } x_2^{-1}x_1 \in \mathbf{H}$$

defines an equivalence relation on the set G. It is called the *left-equivalence modulo* $\mathbf{H}$.

- We have that $x \sim x$ because $x^{-1}x = e \in H$.
- If $x \sim y$ then $y \sim x$: That $x \sim y$ means $y^{-1}x \in \mathbf{H}$ and $y \sim x$ is $x^{-1}y \in \mathbf{H}$. But $x^{-1}y = (y^{-1}x)^{-1} \in \mathbf{H}$.
- If $x \sim y$ and $y \sim z$ the $x \sim z$: We have $y^{-1}x \in \mathbf{H}$ and $z^{-1}y \in \mathbf{H}$; thus $(z^{-1}y)(y^{-1}x) = z^{-1}x \in \mathbf{H}$.

Similarly,

$$x_1 \sim_R x_2 \text{ iff } x_1x_2^{-1} \in \mathbf{H}$$

defines an equivalence on $\mathbf{G}$; it is called the *right-equivalence modulo* $\mathbf{H}$.

Let T_1 and T_2 be subsets of the groupoid $\mathbf{A}$. Then

$$T_1 \cdot T_2 = \{a | a = x_1 \cdot x_2 \text{ for some } x_1 \in T_1 \text{ and } x_2 \in T_2\}$$

is called the *complex product* of T_1 and T_2. We put:

$$xT = \{x\}T = \{xy | y \in T\}$$

We then have for the left-equivalence modulo $\mathbf{H}$:

$$[x] = x\mathbf{H} = \{y | y = xh \text{ for some } h \in \mathbf{H}\}$$

We have:

$x\mathbf{H} \subseteq [x]$: $\quad y = xh$ yields $h = x^{-1}y$ which is $y \sim_L x$

$[x] \subseteq x\mathbf{H}$: $\quad y \sim_L x$ iff $x^{-1}y \in \mathbf{H}$ iff $x^{-1}y = h$ for some h iff $y = xh$ for some h iff $y \in x\mathbf{H}$

The equivalence classes with respect to the left-equivalence modulo $\mathbf{H}$ are the *left-cosets*:

$$[x]^L_H = x\mathbf{H}$$

Similarly,

$$[x]^R_H = \mathbf{H}x$$

are the *right-cosets* modulo $\mathbf{H}$. $\mathbf{H}$ is a normal subgroup iff

$$\mathbf{H}x \subseteq x\mathbf{H}$$

holds for every x.
We have: $\mathbf{H}x \subseteq x\mathbf{H}$ for every x iff for every $h \in \mathbf{H}$ and $x \in \mathbf{G}$ one has some $h' \in \mathbf{H}$ such that $hx = xh'$ iff $x^{-1}hx = h' \in \mathbf{H}$ iff $x^{-1}\mathbf{H}x \subseteq \mathbf{H}$ iff $\mathbf{H}$ is normal.
Similarly, $\mathbf{H}$ is normal iff

$$x\mathbf{H} \subseteq \mathbf{H}x$$

holds for every x.

Thus $\mathbf{H}$ *is normal iff the left-equivalence modulo* $\mathbf{H}$ *is equal to the right-equivalence modulo* $\mathbf{H}$.

We call it the *congruence modulo* $\mathbf{H}$:

Let $\mathbf{N}$ *be a normal subgroup. Then the equivalence modulo* $\mathbf{N}$ *is a congruence relation.*

We need to show:

(i) If $x_1 \sim x_2$ *mod*$(\mathbf{N})$ and $y_1 \sim y_2$ *mod*$(\mathbf{N})$ then $x_1y_1 \sim x_2y_2$ *mod*$(\mathbf{N})$

(ii) If $x_1 \sim x_2$ *mod*$(\mathbf{N})$ then $x_1^{-1} \sim x_2^{-1}$ *mod*$(\mathbf{N})$

That is:

(i) If $x_2^{-1}x_1 \in \mathbf{N}$ and $y_2^{-1}y_1 \in \mathbf{N}$ then $(x_2y_2)^{-1}x_1y_1 \in \mathbf{N}$. But:
$(x_2y_2)^{-1}x_1y_1 = y_2^{-1}(x_2^{-1}x_1)y_1 = (y_2^{-1}y_1)(y_1^{-1}(x_2^{-1}x_1)y_1) \in N$

(ii) We need to show that if $x_2^{-1}x_1 \in \mathbf{N}$ then $(x_2^{-1})^{-1}x_1^{-1} \in \mathbf{N}$. But $x_2^{-1}x_1 \in \mathbf{N}$ implies that $x_2(x_2^{-1}x_1)x_2^{-1} \in \mathbf{N}$. Thus, $x_1x_2^{-1} \in \mathbf{N}$ but then also $(x_1x_2^{-1})^{-1} = x_2x_1^{-1} \in \mathbf{N}$

Assume that the left-equivalence $\sim_L$ *modulo* $\mathbf{H}$ *is a congruence. Then* $\mathbf{H}$ *is normal.*

Assume that $\sim_L$ is a congruence. Then one has that:

$$x_1 \sim_L x_2 \text{ iff } x_1x_2^{-1} \sim_L e \text{ iff } x_1x_2^{-1} \in \mathbf{H} \text{ iff } x_1 \sim_R x_2$$

Hence, $\sim_L = \sim_R$. That is, $\sim_L$ is a congruence.

The next proposition summarizes our observations on groups and their homomorphisms.

Proposition 2.3 *Let* $\mathbf{H}$ *be a subgroup of the group* $\mathbf{G}$*. Then:*

$$\sim_L\colon\ x_1 \sim_L x_2 \text{ iff } x_2^{-1}x_1 \in H \text{ and } \sim_R\colon\ x_1 \sim_R x_2 \text{ iff } x_1x_2^{-1} \in \mathbf{H}$$

are equivalence relations on $\mathbf{G}$*. The equivalence classes are the cosets modulo* $\mathbf{H}$*:*

$$[x]_L = x\mathbf{H}\ ,\ [x]_R = \mathbf{H}x$$

The subgroup $\mathbf{H}$ *is called normal iff* $\sim_L = \sim_R$ *or, equivalently,* $x\mathbf{H} = \mathbf{H}x$ *holds for every* x*; this is the same as* $x^{-1}\mathbf{H}x \subseteq \mathbf{H}$ *where actually equality follows.*

In case of a normal subgroup $\mathbf{N}$*, the equivalence modulo* $\mathbf{N}$ *is a congruence and for the factor algebra, which is denoted as* $\mathbf{G}/\mathbf{N}$*, is, as a homomorphic image of* $\mathbf{G}$ *also a group:*

$$G/N = (\{Nx | x \in G\}, \cdot, {}^{-1}, e = \mathbf{N}\})$$

where the operations are defined by:

$$\mathbf{N}x \cdot \mathbf{N}y = \mathbf{N}(xy),\ (\mathbf{N}x)^{-1} = \mathbf{N}x^{-1}$$

The map $p_{\mathbf{N}} : \mathbf{G} \to \mathbf{G}/\mathbf{N},\ x \mapsto \mathbf{N}x$ *is homomorphic.*

Corollary 2.4 (Homomorphism Theorem for Groups) *Let* $\varphi : \mathbf{G} \to \mathbf{G}'$ *be a homomorphism between groups. Then* $\ker(\varphi)$ *is the equivalence defined by the normal subgroup* $\mathbf{N}_\varphi = \varphi^{-1}\{e_{\mathbf{G}'}\}$ *and* $im(\varphi) = \varphi(G) \cong \mathbf{G}/\mathbf{N}_\varphi$*.*

Proposition 2.5 (Correspondence Theorem for Subgroups) *Let* $\varphi : \mathbf{G} \to \mathbf{G}'$ *be a surjective homomorphism between groups. Then the lattice* $Sub(\mathbf{G})$ *of all subgroups of* $\mathbf{G}'$ *is order isomorphic to the interval* $I_\varphi = [\ker(\varphi), \mathbf{G}] \subseteq Sub(\mathbf{G})$ *of all subgroups that contain* $\ker(\varphi)$*:*

$$\varphi^{-1} : Sub(\varphi(G)) \to I_\varphi\ , C\ \mapsto\ \varphi^{-1}(C)$$

has inverse $\varphi_{|I_\varphi}$*.*

Proof. For any subgroup $\mathbf{C}$ of $\mathbf{G}'$ one has that $\varphi^{-1}(\mathbf{C})$ is a subgroup of $\mathbf{G}'$ and $\varphi^{-1}(\mathbf{C}) \supseteq \ker(\varphi)$. Also, for any subgroup $\mathbf{B}$ of $\mathbf{G}$ one has that $\varphi(\mathbf{B})$ is a subgroup of $\mathbf{G}'$, and the maps φ and φ' are both monotone. We have:

- $(\varphi \circ \varphi^{-1})(\mathbf{C}) = \mathbf{C}$ because φ is surjective. (Note: $\subseteq$ holds in general)
- $(\varphi^{-1} \circ \varphi)(\mathbf{B}) = \mathbf{B}$ if $\mathbf{B} \supseteq \ker(\varphi)$. (Note: $\supseteq$ holds in general)

  Let $x \in \varphi^{-1}\varphi(\mathbf{B})$. Then $\varphi(x) = \varphi(b)$ for some $b \in \mathbf{B}$; but this is the same as $\varphi(xb^{-1}) = e'$. Hence, $xb^{-1} \in \ker(\varphi) \subseteq \mathbf{B}$, i.e., $xb^{-1} = b'$ where $b' \in \mathbf{B}$. Thus, $x = b'b \in \mathbf{B}$. □

Proposition 2.6 (Correspondence Theorem) *Let* $\varphi : \mathbf{G} \to \mathbf{G}'$ *be a surjective homomorphism between groups. Then the lattice* $Norm(\mathbf{G})$ *of all normal subgroups of* $\mathbf{G}'$ *is order isomorphic to the interval* $I_\varphi = [\ker(\varphi), \mathbf{G}] \subseteq Norm(\mathbf{G})$ *of all normal subgroups that contain* $\ker(\varphi)$*:*

$$\varphi^{-1} : Norm(\varphi(G)) \to I_\varphi\ , C\ \mapsto\ \varphi^{-1}(C)$$

has inverse $\varphi_{|I_\varphi}$*.*

PROOF. Let $\mathbf{N}'$ be a normal subgroup of $\mathbf{G}'$. ($\mathbf{N}' \lhd \mathbf{G}'$) Then $\varphi^{-1}(\mathbf{N}') = \mathbf{N} \lhd \mathbf{G}$. Let $x \in \mathbf{N}$. Then one has that $\varphi(x) \in \mathbf{N}'$. Let $y \in \mathbf{G}$. But then:

$$\varphi(yxy^{-1}) = \varphi(y)\varphi(x)\varphi(y^{-1}) \in \mathbf{N}'$$

Let $\mathbf{N} \lhd \mathbf{G}$. Then $\varphi(\mathbf{N}) \lhd \varphi(\mathbf{G}) = \mathbf{G}'$. Let $u = \varphi(x)$ where $x \in \mathbf{N}$. If $v \in \mathbf{G}'$, then by surjectivity of φ one has that $v = \varphi(y)$ for some $y \in \mathbf{G}$. But then:

$$vuv^{-1} = \varphi(y)\varphi(x)\varphi(y)^{-1} = \varphi(yxy^{-1}) \in \varphi(\mathbf{N})$$

This proves the claim. □

For the lattice of subgroups of $\mathbb{Z}$ we have that:

- Every subgroup $\mathbf{A}$ is of the form $k\mathbb{Z} = \{kn | n \in Z\}$ for a unique $k \geq 0$; $k = \min\{n | n \in \mathbf{A}\}$
- $l\mathbb{Z} \supseteq k\mathbb{Z}$ iff $l|k$

Hence,

Proposition 2.7 *The lattice of subgroups of $\mathbb{Z}$ is order isomorphic to $(N, \mid)^*$:*

$$Sub(\mathbb{Z}) \to (N, \mid)^*, k\mathbb{Z} \ \mapsto \ k$$

Corollary 2.8 *For any positive number k one has that the lattice of subgroups of $\mathbb{Z}$ which contain $k\mathbb{Z}$, i.e., the interval $[k\mathbb{Z}, \mathbb{Z}]$ in $Sub(\mathbb{Z})$ is order isomorphic to the lattice of divisors of k:*

$$[k\mathbb{Z}, \mathbb{Z}] \cong ([1, k], \mid)$$

PROOF. We have by Proposition 2.7:

$$[k\mathbb{Z}, \mathbb{Z}] \cong ([1, k], \mid)^*$$

but

$$([1, k], \mid)^* \ \cong \ ([1, k], \mid), \ l \ \mapsto \ k/l \quad \square$$

Corollary 2.9 *For any $k > 0$ one has that the lattice of subgroups of $\mathbb{Z}_k = \mathbb{Z}/k\mathbb{Z}$ is order isomorphic to the lattice $[1, k]$ of divisors l of k.*

PROOF. $Sub(\mathbb{Z}_k) \ \cong \ [kZ, \mathbb{Z}] \ \cong \ [1, k]$. □

Corollary 2.10 *Let $k\mathbb{Z}$ and $l\mathbb{Z}$ be subgroups of $\mathbb{Z}$. Then:*

(a) $kZ \wedge lZ = \ k\mathbb{Z} \cap l\mathbb{Z} = \ l.c.m.(k, l)\mathbb{Z}$;

(b) $k\mathbb{Z} \vee l\mathbb{Z} = \ k\mathbb{Z} + l\mathbb{Z} = \ g.c.d.(k, l)\mathbb{Z}$.

In particular,

$$g.c.d.(k, l) = n_0k + m_0l \text{ for suitable } n_0, m_0 \in \mathbb{Z}.$$

The map $x\mathbf{H} \to \mathbf{H}x, \ xh \ \mapsto \ hx$ is a bijection from the left-coset that contains x to the right-coset that contains x. Hence, all cosets modulo $\mathbf{H}$ have the same cardinal, namely the cardinal of $\mathbf{H}$. The cardinal of the set of cosets is called the *index* of $\mathbf{H}$ in $\mathbf{G}$.

Theorem 2.11 (Lagrange's Theorem) *Let* $\mathbf{G}$ *be a finite group and let* $\mathbf{H}$ *be a subgroup of* $\mathbf{G}$*. Then:*

$$card(\mathbf{G}) = card(\mathbf{H}) \cdot index(\mathbf{H})$$

In particular, $card(\mathbf{H})|card(\mathbf{G})$ *and* $index(\mathbf{H})|card(\mathbf{G})$*.*

Corollary 2.12 *Let* $\mathbf{G}$ *be a finite group. If* $card(\mathbf{G}) = p$*, where* p *is a prime, then* G *is generated by a single element. That is,* G *is cyclic.*

Corollary 2.13 *Let* G *be a finite group, where* $card(\mathbf{G}) = n$ *and let* x *be any element of* G*. Then, if the cyclic group that is generated by* x *in* $\mathbf{G}$ *has* m *elements, then* $m|n$*.*

Theorem 2.14 (Main Theorem for Cyclic Groups) *Let* G *be a cyclic group, i.e.,* $\mathbf{G} =< x >$ *for some* $x \in \mathbf{G}$*. Then either* $\mathbf{G}$ *is infinite, and in this case one has that* $\mathbf{G} \cong \mathbb{Z}$*, or, in case that* $\mathbf{G}$ *is finite, one has that* $\mathbf{G} \cong \mathbb{Z}/(n)$*, where* $n = card(\mathbf{G})$*.*

PROOF. $\mathbb{Z} \to \mathbf{G},\ n \mapsto x^n$, is a surjective homomorphism. Therefore, $\mathbb{Z}/n\mathbb{Z} \cong \mathbf{G},\ n \geq 0$. □

One has in every group,

$$\begin{aligned} \text{the left translations } \lambda_x &: \mathbf{G} \to \mathbf{G},\ y \mapsto xy \\ \text{the right translations } \rho_x &: \mathbf{G} \to \mathbf{G},\ y \mapsto yx \\ \text{the inner automorphisms } \sigma_x &: \mathbf{G} \to \mathbf{G},\ y \mapsto xyx^{-1}. \end{aligned}$$

All λ_x, ρ_x and σ_x are bijective; the σ_x are automorphisms.

Theorem 2.15 (Cayley's Theorem) *For any group* $\mathbf{G}$*, the map*

$$\lambda : G = (G, \cdot, ^{-1}, e) \ \to\ (Bij(G), \circ, ^{-1}, id_{\mathbf{G}}),\ x \mapsto \lambda_x$$

is an injective homomorphism, i.e., $\mathbf{G}$ *is isomorphic to its group of left translations.*

2.2 Rings, Domains and Fields

An algebraic structure $\mathbf{A} = (A, +, -, 0, \cdot, 1)$ is a ring if

(a) The *additive reduct* $(A, +, -, 0)$ is an abelian group.

(b) The *multiplicative reduct* is a monoid.

(z) Addition $+$ and multiplication $\cdot$ are related by distributivity:

$$\text{D1: } x \cdot (y + z) = x \cdot y + x \cdot z \qquad \text{D2: } (y + z) \cdot x = y \cdot x + z \cdot x$$

The following very basic facts hold in every ring. The proofs are simple and are based on distributivity.

- $a \cdot 0 = 0 \cdot a = 0$
- $(-a) \cdot b = -(a \cdot b),\ a(-b) = -(a \cdot b),\ (-a) \cdot (-b) = a \cdot b.$
- $(-1) \cdot a = a \cdot (-1) = -a$

One also has the rules of working with the sum sign. If one defines:

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n$$

then

$$(\sum_{i=1}^{n} a_i) \cdot (\sum_{j=1}^{m} b_j) = a_1 \cdot (\sum_{j=1}^{m} b_j) + \ldots + a_n \cdot (\sum_{j=1}^{m} b_j) = \sum_{j=1}^{m} a_1 \cdot b_j + \ldots + \sum_{j=1}^{m} a_n \cdot b_j \ = \ \sum_{i,j=1}^{n,m} a_i b_j$$

A ring is commutative if the multiplication is commutative. In a commutative ring one has the classical binomial theorem, and a great deal of the theory of determinants still holds. We will always assume that a ring has a unit 1 for multiplication. Rings without unit have been called by Jacobson [2] "rngs". Obviously, he doesn't want to deal with objects nobody can pronounce properly. However, Hungerford [1] develops rings without unit. For us, rings will always have a unit. If a ring doesn't have a unit, one can in a canonical fashion add one.

Examples

1. $\mathbb{Z} = (Z, +, -, 0, \cdot, 1)$ is a commutative ring.

2. It is easy to see that the congruence modulo $k\mathbb{Z}$ in $\mathbb{Z}$ is actually a ring congruence, i.e., if $x \equiv x'$ and $y \equiv y'$ then $xy \equiv x'y'$, hence $[x][y] = [xy]$ defines a multiplication on $\mathbb{Z}/k\mathbb{Z}$ and $\mathbb{Z} \to \mathbb{Z}/k\mathbb{Z}, x \ \mapsto \ [x]$ is a homomorphism. It is easy to see that the homomorphic image of a ring is a ring.

3. Let $\mathbf{A}$ be an abelian group. Then

   $$End(\mathbf{A}) = (\{\varphi | \varphi : \mathbf{A} \to \mathbf{A}, \ \varphi \text{ is a homomorphism}\}, +, -, 0_{\mathbf{A}}, \circ, id_{\mathbf{A}})$$

   is a ring. Here, $0_{\mathbf{A}}$ is the zero-map on $\mathbf{A}$. Of course, $+$ is the addition of

   functions, i.e., $(\varphi_1 + \varphi_2)(a) = \varphi_1(a) + \varphi_2(a)$. One needs the commutativity of addition in $\mathbf{A}$ in order to show that the sum of homomorphisms is homomorphic; of course, $\circ$ is the composition. In general, a homomorphism of an algebra to itself is called an endomorphism.

Because rings are defined by equations, we have as before

Proposition 2.16 *The class of all rings is a primitive class of algebras.*

Let $\varphi : \mathbf{A} \to \mathbf{A}'$ be a homomorphism between rings. Then:

$$x_1 \sim_\varphi x_2 \text{ iff } \varphi(x_1) = \varphi(x_2) \text{ iff } \varphi(x_1 - x_2) = 0 \text{ iff } x_1 - x_2 \in \mathbf{I}_\varphi = \{x | \varphi(x) = 0\}$$

is a congruence relation on $\mathbf{A}$. $\mathbf{I} = \mathbf{I}_\varphi$ is called the kernel of φ and is an *ideal* of $\mathbf{A}$. That is:

- $\mathbf{I}$ is closed under $+$, $-$, and contains 0.

  This is obvious because φ is a homomorphism between the additive module structures of the rings.

- If $x \in \mathbf{I}$ and $y \in \mathbf{A}$ then one has that

$$\text{LI: } yx \in \mathbf{I} \text{ and RI: } xy \in \mathbf{I}$$

Indeed, if $\varphi(x) = 0$ then $\varphi(xy) = \varphi(x)\varphi(y) = 0 = \varphi(y)\varphi(x)$

A non-empty subset of $\mathbf{A}$ that is closed under addition is a *left-ideal* if LI holds, and it is a *right-ideal* if RI holds; a subset that is a left-ideal as well a right-ideal is an ideal.

Let $\mathbf{I}$ be an ideal of $\mathbf{A}$. Then:

$$x_1 \sim x_2 \; mod(I) \text{ iff } x_1 - x_2 \in I$$

defines an equivalence on $\mathbf{A}$. This equivalence is compatible with the ring multiplication and therefore a congruence. That is:

$$\text{If } x_1 \sim y_1 \text{ and } x_2 \sim y_2 \text{ then } x_1x_2 \sim y_1y_2$$

We have $x_1 - y_1 \in \mathbf{I}$ and $x_2 - y_2 \in \mathbf{I}$. But then, $x_1x_2 - y_1x_2 \in \mathbf{I}$ and $y_1x_2 - y_1y_2 \in \mathbf{I}$, because $\mathbf{I}$ is a right as well a left-ideal. But then $(x_1x_2 - y_1x_2) + (y_1x_2 - y_1y_2) = x_1x_2 - y_1y_2 \in \mathbf{I}$. This is, $x_1x_2 \sim y_1y_2$. On the other hand, let $\mathbf{I}$ be a submodule of $\mathbf{A}$. Then $\mathbf{I}$ is an ideal if the equivalence modulo $\mathbf{I}$ is compatible with the ring multiplication. Indeed, we have that $x \in \mathbf{I}$ iff $x \sim_{\mathbf{I}} 0$. But then, $yx \sim 0$ and $xy \sim 0$.

Proposition 2.17 *Let* $\mathbf{I}$ *be an ideal of the ring* $\mathbf{A}$*. Then the equivalence modulo* $\mathbf{I}$ *is a ring congruence. The factor algebra* $\mathbf{A}/\mathbf{I}$ *is (as a homomorphic image of* $\mathbf{A}$*) a ring:*

$$\mathbf{A}/\mathbf{I} = (\{I + x | x \in A\}, +, -, 0 = I, \cdot, 1 = I + 1)$$

where,

$$(I + x) + (I + y) = I + (x + y),\ -(I + x) = I + (-x),\ (I + x) \cdot (I + y) = I + xy$$

and $p_{\mathbf{I}} : \mathbf{A} \to \mathbf{A}/\mathbf{I}, x \mapsto I + x$ *is homomorphic.*

Corollary 2.18 (Homomorphism Theorem for Rings) *Let* $\varphi : \mathbf{A} \to \mathbf{A}'$ *be a homomorphism between rings. Then* $\ker(\varphi)$ *is the equivalence defined by the ideal* $\mathbf{I}_\varphi = \varphi^{-1}\{0\}$ *and* $im(\varphi) = \varphi(A) \cong \mathbf{A}/\mathbf{I}_\varphi$.

Examples

1. $k\mathbb{Z}$ is an ideal in $\mathbb{Z}$ and therefore $\mathbb{Z}/k\mathbb{Z} = \mathbb{Z}_k$ is a ring, the ring of integers modulo k.
2. We define for any ring with multiplicative unit a "scalar multiplication":

$$0.1 = 0, (n + 1).1 = n.1 + 1 \text{ where n is a positive natural number}$$

This defines $n.1$ for every natural number. Furthermore, we define for any $n \in \mathbb{N}$,

$$(-n).1 = -(n.1)$$

Hence, $k.1$ is defined for every $k \in \mathbb{Z}$.

One easily verifies the following rules:

$$(k + l).1 = k.1 + l.1,\ (-k).1 = -k.1,\ 0.1 = 0,\ (k \cdot l).1 = (k.1) \cdot (l.1),\ 1_{\mathbb{Z}}.1_{\mathbf{A}} = 1_{\mathbf{A}}$$

This tells us that

$$\kappa = \kappa_{\mathbf{A}} : \mathbb{Z} \to \mathbf{A}, k \mapsto k.1$$

is a homomorphism. We have that $\ker(\kappa) = k\mathbb{Z}$ for some number $k \geq 0$. The natural number k is called the *characteristic* of $\mathbf{A}$. It follows from this definition, that the characteristic of a ring is either 0, or the smallest positive natural number k for which $k.1 = 0$, if there is such a k.

Let $\mathbf{A}$ be any ring. Any subring of $\mathbf{A}$ must contain 0 and 1, but then also $k.1$ for every $k \in \mathbf{Z}$. Obviously, $im(\kappa)$ is the smallest subring of $\mathbf{A}$. It is also called the *prime ring* of $\mathbf{A}$. Hence,

$$\mathbb{Z}/k\mathbb{Z} = \mathbb{Z}_k \ \cong \ im(\kappa) = \text{ prime ring of } \mathbf{A}, \text{ where } k = char(\mathbf{A})$$

Any ring $\mathbf{A}$ may be considered as an *extension* of $\mathbb{Z}_k$, where k is the characteristic of $\mathbf{A}$. In this sense we may identify the unit of a ring with 1 or 1 modulo k, if $k \geq 0$.

A ring is called a *domain* if it is without *zero divisors*. That is,

DO: If $x \neq 0$ and $y \neq 0$ then $xy \neq 0$, or, equivalently $xy = 0$ implies that $x = 0$ or $y = 0$

Examples

1. $\mathbb{Z}$ is a domain. Because of this example, domains are sometimes called *integral domains*.
2. The *zero-ring* $\mathbb{O}$ is a domain. It is sometimes excluded.
3. $\mathbb{Z}_n$, $n > 0$, is a domain iff n is a prime. Indeed, if n is not a prime, then $n = n_1 n_2$ where n_1 and n_2 are proper divisors of n. But then $[n_1] \cdot [n_2] = [n] = [0]$ shows that $\mathbb{Z}_n$ is not a domain. On the other hand, if p is a prime then $\mathbb{Z}_p$ is a domain. Indeed, if $[n_1] \cdot [n_2] = [0]$ then $p | n_1 n_2$. This means that $p | n_1$ or $p | n_2$. hence, $[n_1] = [0]$ or $[n_2] = [0]$.
4. $\mathbb{Z} \times \mathbb{Z}$ is not a domain: $(0, 1) \cdot (1, 0) = (0, 0)$
5. A subring of a domain is a domain. But the class of domains is not closed under direct products and not closed under homomorphic images.

Equivalent to DO is the restricted cancellation law:

DO': If $xy = xy'$ and $x \neq 0$ then $y = y'$. If $yx = y'x$ and $x \neq 0$ then $y = y'$

Assume DO, $x \neq 0$ and $xy = xy'$. We then have $x(y - y') = 0$ and therefore, by DO, $y - y' = 0$, i.e., $y = y'$. The other half of DO' follows similarly.
Assume DO' and $xy = 0$. Then if $x \neq 0$. one has that $xy = x0$ and therefore $y = 0$. Similarly, if $y \neq 0$ then $xy = 0y$ yields $x = 0$.

A ring $\mathbf{A}$ is called a *division ring* if DR holds: the non-zero elements form a group under multiplication. That is:

- $1 \neq 0$
- If $x \neq 0$ and $y \neq 0$ then $xy \neq 0$
- If $x \neq 0$ then there is some $y \neq 0$ such that $xy = yx = 1$

One has that DR is equivalent to the apparently weaker condition DR':

- $1 \neq 0$

- If $x \neq 0$ then there is some x' such that $xx' = 1$

If $x \neq 0$ and $y \neq 0$ then $(xy)(y'x') = 1$ yields $xy \neq 0$.
If $x \neq 0$ then $xx' = 1$ yields $x' \neq 0$. Therefore, $x'x'' = 1$ and $x = x(x'x'') = (xx')x'' = x''$. Thus, $x'x = 1$.

For a domain one has that it is either the zero-ring, i.e., $1 = 0$, or the elements different from zero form a monoid under multiplication. For a division ring this monoid is a group.

Proposition 2.19 *A finite domain $\mathbf{D} \neq \mathbb{O}$ is a division ring.*

PROOF. Let $x \neq 0$. Then $\lambda_x : \mathbf{D}/\{0\} \to \mathbf{D}/\{0\}, y \mapsto xy$ is by DO' injective. Because $\mathbf{D}$ is finite, λ_x is bijective. Therefore, $xy = 1$ for some y. This is DR'. □

Proposition 2.20 *A ring $\mathbf{A}$ is a division ring if and only if $\mathbf{A}$ has exactly two right ideals: $\mathbb{O}$ and $\mathbf{A}$.*

PROOF. Let I be a right-ideal of the division ring $\mathbf{D}$. Then either $I = \mathbb{O}$ or there is an $x \neq 0$ in I. But then $xx' = e \in I$ and one has for any $y \in \mathbf{D}$ that $1y = y \in I$. This is $I = \mathbf{D}$.
For any ring $\mathbf{A}$ and $x \in \mathbf{A}$ one has the right ideal that is generated by x: $(x)_r = x \cdot \mathbf{A} = \{xy | y \in \mathbf{A}\}$
Assume now that $\mathbf{A}$ has exactly two right-ideals. Then $\mathbf{A} \neq \mathbb{O}$ and therefore $1 \neq 0$. Let $x \in \mathbf{A}, x \neq 0$. Then $(x)_r = \mathbf{A}$ yields some y such that $xy = 1$. This is DR'. □

Corollary 2.21 *Let $\varphi : \mathbf{D} \to \mathbf{A}$ be a homomorphism from the division ring $\mathbf{D}$ into $\mathbf{A}$. Then either $im(\varphi) = \mathbb{O}$ or $im(\varphi) \cong \mathbf{D}$.*

A commutative division ring is called a *field*. $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields. $\mathbb{Z}_n$ is a field if and only if n is a prime number. One can show that any finite division ring is a field.

Let I be an ideal of the commutative ring $\mathbf{A}$; I is said to be a *prime ideal* iff the complement $\mathbf{A} - I$ is multiplicatively closed. That is, if $x \notin I$ and $y \notin I$ then $xy \notin I$. Equivalently, if $xy \in I$ then $x \in I$ or $y \in I$. The following proposition is only a reformulation of the definition in terms of factor algebras:

Proposition 2.22 *Let $\mathbf{A}$ be a ring and let I be an ideal. Then $\mathbf{A}/I$ is a domain if and only if I is a prime ideal.*

In $\mathbb{Z}$, every ideal I is principal, i.e., $I = k\mathbb{Z} = (k)$ for a unique $k \geq 0$ and one has that (k) is prime iff k is prime.

An ideal I of the commutative ring $\mathbf{A}$ is called *maximal* if and only if $I \neq \mathbf{A}$ and there is no ideal J strictly between I and $\mathbf{A}$. That is, I is maximal in the ordered set of all ideals different from $\mathbf{A}$.

Proposition 2.23 *Let $\mathbf{A}$ be a commutative ring and let I be an ideal. Then I is maximal if and only if $\mathbf{A}/I$ is a division ring.*

PROOF. Let M be maximal. Then if $x \notin M$ one has that the ideal that is generated by M and x is the whole ring. It is easy to see that the smallest ideal that contains the ideals I and J is just the sum $I + J = \{i + j | i \in I, j \in J\}$. The ideal which is generated by x is the principal ideal $(x) = x\mathbf{A}$. Hence, we have that

$$M + (x) = \{m + xy | m \in M, y \in \mathbf{A}\} = \mathbf{A}$$

and, therefore, $m + xy = 1$ for certain $m \in M$ and $y \in \mathbf{A}$. For the factor ring $\mathbf{A}/M$ this means that for every class $[x] \neq [0] = M$ one has some $[y]$ such that $[m] + [x][y] = [x][y] = [1]$. Thus, $\mathbf{A}/M$ is a field.
If I is not maximal then there is some ideal J such that $I \subset J \subset \mathbf{A}$. This yields a homomorphism $\mathbf{A}/I \to \mathbf{A}/J$, $[x]_I \mapsto [x]_J$ whose kernel is $\bar{J} = \{[x]_I | x \in J\}$ which is a proper ideal that is different from the zero ideal $\{[0] = M\}$ and the whole ring $\mathbf{A}$. Hence, according to Corollary 2.21, $\mathbf{A}/I$ is not a field. □

Proposition 2.24 (Krull's Theorem) *Let* $\mathbf{A}$ *be any ring different from the zero ring and let* I *be any ideal different from* $\mathbf{A}$*. Then* I *is contained in some maximal ideal of* $\mathbf{A}$*.*

PROOF. Let $\mathcal{S} = (\{J | J \supseteq I, J \neq \mathbf{A}, J \in Ideal(A)\}, \subseteq)$. Then $\mathcal{S}$ is an ordered set. Any chain in $\mathcal{S}$ has an upper bound because the union of a chain of proper ideals is a proper ideal. Here we use that the ring has a unit. Hence, by Zorn's lemma, $\mathcal{S}$ has a maximal element, i.e., a maximal ideal M that contains I. □

2.3 Modules over a Ring. Vector Spaces

Let $\mathbf{A}$ be a ring. An $\mathbf{A}$-*module* $\mathbf{M}$ is a structure of type $\mathbf{M} = (M, +, -, 0, (\alpha.)_{\alpha \in A})$ such that

(i) $\mathbf{M} = (M, +, -, 0)$ is a module.

(ii) The familiar vector space axioms hold:

$$\begin{aligned} \alpha.(a+b) &= \alpha.a + \alpha.b \\ (\alpha+\beta).a &= \alpha.a + \beta.a \\ (\alpha \cdot \beta).a &= \alpha.(\beta.a) \\ 1.a &= a \end{aligned}$$

We denote elements of the ring, the "scalars", by lower case Greek letters and the elements of the module $\mathbf{M}$ by lower case Latin letters. Each $\alpha \in \mathbf{A}$ leads to a unary operation $a \mapsto \alpha.a$. We do not make a notational distinction between the zero of $\mathbf{A}$ and the "zero vector", that is the he zero of $\mathbf{M}$. While not always required, we will assume from now on that $\mathbf{A}$ is **commutative**. Here are some elementary algebraic facts on $\mathbf{A}$-modules:

- $0.a = 0$
- $\alpha.0 = 0$
- $(-1).a = -a, (-\alpha).a = -(\alpha.a)$

Examples

1. Every module is a module over $\mathbb{Z}$, where $n.a$ has been defined previously, see page 18. Let $\mathbf{A}$ be a ring. Then $\mathbf{M_A} = (A, +, -, 0)$ is an $\mathbf{A}$-module by means of $\alpha.x = \alpha \cdot x$.

2. In particular, the rings $\mathbb{Z}_n$ are modules over $\mathbb{Z}$ as well as over $\mathbb{Z}_n$: $n.[m] = [nm] = [n][m] = [n].[m]$. For example, $2.[3]_6 = [2 \cdot 3]_6 = [0]$. Hence, it may happen that $\alpha.a = 0$ even if $\mathbf{A}$ is without zero-divisors.

3. If $\mathbf{A}$ is a division ring then
$$\alpha.a = 0 \text{ iff } \alpha = 0 \text{ or } a = 0$$
The proof uses all $\mathbf{A}$-module axioms: Assume $\alpha.a = 0$ and $\alpha \neq 0$. Then $\alpha^{-1} \cdot \alpha = 1$. But then, $\alpha^{-1}.(\alpha.a) = \alpha^{-1}.0 = 0 = 1.a = a$

4. Let S be any set and let $\mathbf{A}$ be a ring. Then the module $\mathbf{A}^S$ is an $\mathbf{A}$-module according to the pointwise defined scalar multiplication, e.g., $\alpha.f$ is the map $s \mapsto \alpha.f(s)$. For $\mathbf{A} = \mathbb{R}$ and $S = \{0, \cdots, n-1\} = n$ we get the familiar vector space $\mathbb{R}^n$. For $S = \{0\} = 1$, $\mathbf{A}^1$ is the ring $\mathbf{A}$ perceived as a module over itself. More generally, we have, similarly as before:

Proposition 2.25 *For any ring* $\mathbf{A}$*, the class of* $\mathbf{A}$*-modules is a primitive class of algebras.*

A homomorphism φ between $\mathbf{A}$-modules $\mathbf{M}$ and $\mathbf{N}$ is called a *linear* map. it is very easy to see that one only has to stipulate the following two conditions:

- φ is *additive*, i.e., $\varphi(a+b) = \varphi(a) + \varphi(b)$.
- φ is *homogeneous*, i.e., $\varphi(\alpha.a) = \alpha.\varphi(a)$.

Note, that multiplication by a, i.e., the map $a \mapsto a \cdot b$ is homogeneous only if the ring $\mathbf{A}$ is commutative.

A subset C of an $\mathbf{A}$-module is already closed if

- C is non-empty.
- If a and b are in C then $a+b$ is in C.
- If a is in C then $\alpha.a$ is in C, $\alpha \in \mathbf{A}$

Let $\mathbf{B}_{t\in T}$ be a family of $\mathbf{A}$-submodules of $\mathbf{M}$. Then one has in the complete lattice of $\mathbf{A}$-submodules of the $\mathbf{A}$-module $\mathbf{M}$:

$$\bigwedge_{t\in T} B_t = \bigcap_{t\in T} B_t \text{ and } \bigvee_{t\in T} B_t = \{a | a = b_{t_1} + \cdots + b_{t_k};\ k \geq 0, t_1, \cdots, t_k \in T, b_{t_i} \in B_{t_i}\} = \sum_{t\in T} \mathbf{B}_t$$

Let $\varphi : \mathbf{M} \to \mathbf{N}$ be linear. Then:

$$a_1 \sim a_2\ mod(\varphi) \text{ iff } \varphi(a_1) = \varphi(a_2) \text{ iff } \varphi(a_1 - a_2) = 0 \text{ iff } a_1 - a_2 \in \ker(\varphi) = \{a | \varphi(a) = 0\} = \mathbf{N}_\varphi$$

We already know that $\mathbf{N}_\varphi$ is a module. But it is also a $\mathbf{A}$-submodule. That is, $\mathbf{N}_\varphi$ is closed under multiplication by scalars: Assume that $a \in \mathbf{N}_\varphi$. Then $\varphi(a) = 0$ and $\varphi(\alpha.a) = \alpha.\varphi(a) = 0$. This is $\alpha.a \in \mathbf{N}_\varphi$

Let $\mathbf{N}$ be an $\mathbf{A}$-submodule of $\mathbf{M}$. Then: $a_1 \sim a_2$ iff $a_1 - a_2 \in \mathbf{N}$ is a congruence relation. Again, we only have to show that this equivalence is compatible with the unary operations α. . But this is trivial.

On the other hand, let $\mathbf{N}$ be a submodule of the $\mathbf{A}$-module $\mathbf{M}$. Then the induced equivalence $mod_{\mathbf{N}}$ is a congruence if and only if $\mathbf{N}$ is an $\mathbf{A}$-submodule.

Proposition 2.26 *Let* $\mathbf{N}$ *be an* $\mathbf{A}$*-submodule of the* $\mathbf{A}$*-module* $\mathbf{M}$*. Then* $a_1 \sim a_2$ *iff* $a_1 - a_2 \in \mathbf{N}$ *is an* $\mathbf{A}$*-module congruence. The factor algebra* $\mathbf{A}/\mathbf{N}$ *is as a homomorphic image of* $\mathbf{M}$ *an* $\mathbf{A}$*-module. Note that* $\alpha.(\mathbf{N} + a) = \mathbf{N} + \alpha.a$ *defines the multiplication by scalars* $\alpha \in \mathbf{A}$*. The canonical map* $p_{\mathbf{N}} : a \mapsto \mathbf{N} + a$ *is a linear map from* $\mathbf{M}$ *to* $\mathbf{M}/\mathbf{N}$ *with kernel* $\mathbf{N}$*.*

Corollary 2.27 (Homomorphism Theorem for A-modules) *Let* $\varphi : \mathbf{M_1} \to \mathbf{M_2}$ *be a linear map between* $\mathbf{A}$*-modules. Then* $\mathbf{N}_\varphi = \{a | \varphi(a) = 0\}$ *is an* $\mathbf{A}$*-submodule of* $\mathbf{M}$ *and one has that* $\mathbf{M}/\mathbf{N}_\varphi \cong im(\varphi)$*.*

A subset S of the $\mathbf{A}$-module $\mathbf{M}$ is *generating* if $\mathcal{C}(S) = \mathbf{M}$. That is, every element of $\mathbf{M}$ is a finite linear combination of elements in S. A subset S is *linearly independent* if only trivial linear combinations are zero. That is: If $F = \{s_1, \ldots, s_k\}$ are k elements of S then $\alpha_1.s_1 + \ldots + \alpha_k.s_k = 0$ only if $\alpha_1 = 0, \ldots, \alpha_k = 0$. A subset B that is generating and linearly independent is called a *basis* for $\mathbf{M}$. We recall from elementary linear algebra the following:

Theorem 2.28 *Let* $\mathbb{V}$ *be a vector space with a finite generating set. Then* $\mathbb{V}$ *has a basis and all bases have the same number of elements. This number is called the dimension of* $\mathbb{V}$*.*

The existence of a basis is quite trivial. Any minimal generating set is a basis. The uniqueness of the dimension is based on the fact that m linear combinations of n vectors are linearly dependent if one has that $m > n$. This is the same as saying that a linear system of m homogeneous linear equations in n unknowns has a non-trivial solution. But this is an immediate consequence of the row echelon form. The details of all of this are very easy to prove.

Vector spaces that are not finitely generated have a basis by Zorn's lemma. A maximal linearly independent set is a basis. That any two bases are equivalent follows from simple cardinal arithmetic.

Definition 3 An $\mathbf{A}$-module $\mathbf{M}$ is called the *free A-module, freely generated by B* if for every $\mathbf{A}$-module $\mathbf{N}$ and every map $\varphi : B \to \mathbf{N}$ one has a unique linear map $\hat{\varphi} : \mathbf{M} \to \mathbf{N}$ that extends φ.

One writes $\mathbf{M} = \mathbf{F}(\mathbf{B})$. It is easy to see that $\mathbf{M}$ is uniquely determined by the set B, up to a linear isomorphism that leaves B fixed.

Theorem 2.29 *The $\mathbf{A}$-module $\mathbf{M}$ is free, freely generated by the set B, if and only if, B is a basis for $\mathbf{M}$.*

Proof. Assume that $\mathbf{M}$ has a basis B. Because two homomorphisms that agree on a generating set must be equal, uniqueness of $\hat{\varphi}$ is obvious. Now, every element m of $\mathbf{M}$ is a unique linear combination of elements in B. Then, in an obvious way, a map $\hat{\varphi}(m)$ can be defined, and it is linear. The details of this *linear extension principle* are easy.

Next, we show that $\mathbf{F}(\mathbf{B})$ exists. Let

$$\mathbf{M} = \mathbf{A}^{(S)} = \{f | f : S \to A, f(s) = 0 \text{ for almost all s}\}$$

This is clearly an $\mathbf{A}$-submodule of the direct power $\mathbf{A}^S$ and the maps e_b form a basis where $e_b(b) = 1$, and $e_{b'} = 0$, if $b' \neq b$. The map $b \to e_b$ defines a bijection from B to the basis $\{e_b | b \in B\}$ of $\mathbf{M}$. Thus we may identify the set of unit vectors $e_b, b \in B$ with B.

Because free modules are uniquely determined up to isomorphisms, any free module over B has B as a basis. □

The representation of a linear map $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ by a matrix $\mathbf{A}$ is an important application of this theorem. An algebraic structure $\mathbf{M} = (M, +, -, 0, (\alpha.)_{\alpha \in A}, <, >)$ is an *algebra* over $\mathbf{A}$ if

- $(M, +, -, 0, (\alpha.)_{\alpha \in A})$ is an $\mathbf{A}$-module.
- The map $<, >: M \times M \to M, (m, m') \mapsto < a, b >$ is *bilinear*. That is:
  - (i) $< a + a', b > = < a, b > + < a', b >$; $< a, b + b' > = < a, b > + < a, b' >$
  - (ii) $< \alpha.a, b > = < a, \alpha.b > = \alpha. < a, b >$

A **commutative** ring is an algebra over itself. A familiar example of an algebra over the reals is the *cross product* on $\mathbb{R}^3$. Also, the field $\mathbb{C}$ of complex numbers is an algebra over the reals where the underlying module is the plane $\mathbb{R}^2$. For any field $\mathbf{F}$ the $n \times n$ matrices form an algebra $\mathfrak{gl}_n(\mathbf{F})$ with respect to the ordinary matrix multiplication " $\circ$ ". Unlike the cross product of $\mathbb{R}^3$, this algebra is associative. However, with the auxiliary operation, the *commutator*: $[S, T] \mapsto S \circ \mathbf{T} - \mathbf{T} \circ \mathbf{S}$ it becomes the prototype of a *Lie Algebra*, i.e., the *Jacobian* identity holds:

$$[a, [b, c]] + [b, [c, a]] + [c, [a, b]] = 0$$

Any finite dimensional Lie algebra is isomorphic to a subalgebra of a Lie algebra of matrices. This is a fairly deep theorem for which no elementary proofs are known.

The free module over B becomes an algebra if a multiplication has been defined on the basis. That is, given a multiplication for any pair of basis elements b_i, b_j

$$< b_i, b_j >= \sum_{b \in B} \alpha_{i,j}^b . b$$

one has a unique bilinear extension. Here we assume that the elements of B are indexed by a set I. Of course, for almost all b, the $\alpha_{i,j}^b$ are zero. Again, the proof is rather straight forward. The familiar construction of the complex numbers is an example of this construction principle. We remark that commutativity as well as associativity are carried over from the base to the whole algebra.

2.4 Factorial Monoids and Domains

In the ring $\mathbb{Z}$ of integers, any element $n > 0$ is a unique product of primes. Such a factorization theorem holds for many other rings and concerns mainly the multiplicative structure of the ring. Thus, we are first studying *factorial monoids*.

Let $\mathbf{M}$ be a commutative, *cancellation monoid*. That is:

$$ab = ac \text{ implies that } b = c$$

We then define:

$$a|b \text{ iff there is some } c \in \mathbf{M} \text{ such that } ac = b$$

and say that *a divides b* or that *b is a multiple of a*. The relation "|" is :

- reflexive: $a|a$.
- transitive: If $a|b$ and if $b|c$ then $a|c$.

That is, the relation | is a quasi-ordering. For any quasi-ordering on a set A, the relation *a and b are associated* is an equivalence where a and b are associated if they divide each other:

$$a \sim b \text{ iff } a|b \text{ and } b|a$$

On the set $A/\sim$ of equivalence classes an ordering is defined by

$$[a]|[b] \text{ iff } a|b$$

This is called the *contraction* of a quasi ordering.

What are in our case the equivalence classes? If $a \sim b$ then $ac = b$ and $bc' = a$. Hence, $acc' = bc' = a$. This is $cc' = 1$, where 1 is the unit of our monoid. This is, c and c' are inverse to each other. Now let,

$$\mathbf{U} = Inv(M, \cdot, 1)$$

be the group of invertible elements of $\mathbf{M}$. We have shown that $a \sim b$ implies that $b \in a\mathbf{U}$. On the other hand, if $b \in a\mathbf{U}$ then $b = au$ and $bu^{-1} = a$ show that $a \sim b$. Therefore,

$$[a] = a\mathbf{U},\ a \in \mathbf{M}$$

is the partition of $\mathbf{M}$ into equivalence classes of associated elements.

The equivalence relation of being associated is actually a congruence relation on $\mathbf{M}$. We need to show that if $a \sim a'$ and $b \sim b'$ then $ab \sim a'b'$. We have that $au = a'$, $bv = b'$ but then $ab(uv) = a'b'$.

Thus $\mathbf{M} \to \mathbf{M}/\sim = \overline{\mathbf{M}}, a \mapsto aU$, is a monoid homomorphism. Of course, $\mathbf{U}$ is as the class of 1 the unit of $\overline{\mathbf{M}}$.

$\overline{\mathbf{M}}$ is also a cancellation monoid: If $[a][b] = [a][c]$ then $[ab] = [ac]$, i.e., $abu = ac$. But then $bu = c$, i.e., $[b] = [c]$.

If $a|b$ where a is not a unit and where $b \not| a$ then a is called a *proper* factor of b. A unit cannot have a proper factor. If $a|u$ then $av = u$ and $a = uv^{-1}$ is a unit.

Let $a \in \mathbf{M}$ and let u be a unit. Then $u(u^{-1}a) = a$ shows that u divides a.

If a is a proper factor of b and $ac = b$ then c is a proper factor of b. Indeed, assume that b divides c, i.e., $bd = c$. But then $(ac)d = c$, i.e., $ad = 1$ and a is a unit and therefore not a proper factor of b.

If $ac = b$ where a is not a proper factor. Then one factor is associated to bb and the other one is a unit.

An element $q \in \mathbf{M}$ is called *irreducible* if

- q is not a unit.
- q does not have any proper factors. That is, if $q = ab$ then either a is a unit or a is associated to q.

In $(\mathbb{Z} \setminus \{0\}, \cdot, 1)$ exactly the prime numbers are irreducible. The group of units is $\mathbf{U} = \{1, -1\}$ and $n \sim m$ iff $n = \pm m$.

Definition 4 A factorization $a = q_1 \cdot \ldots \cdot q_s$ is an *essentially unique* factorization of a into irreducible elements q_i if for every other such factorization $a = q'_i \cdot \ldots \cdot q'_s$ one has that $t = s$ and $q_i \sim q_{i'}$ for a suitable permutation $i \mapsto i'$.

Definition 5 Let $\mathbf{M}$ be a commutative cancellation monoid. Then $\mathbf{M}$ is called *factorial* if every non-unit of $\mathbf{M}$ has an essentially unique factorization into irreducible elements. An integral domain $\mathbf{D}$ is factorial if its monoid $(\mathbf{D} \setminus \{0\}, \cdot, 1)$ is factorial.

Let $\mathbf{M}$ be a factorial monoid. Then either

- $a \in \mathbf{U}$, i.e., $a \sim 1$ and a is the empty product of irreducible elements. We define in this case $l(a) = 0$.
- a is the product of s-many irreducible elements where $s \geq 1$ and $s = l(a)$ is uniquely defined.

The natural number $l(a)$ is called the *length* of a.

Let $a = bc$ where b is a proper divisor of a. Then also c is a proper divisor. Now, the factorizations of a and b can be combined to a factorization of ab. We see that $l(a) = l(b) + l(c)$ and where $1 \leq l(b), l(c) < l(a)$. Hence, the length of a proper divisor b of a is less than the length of a.

Definition 6 A binary relation R is said to satisfy the *minimal condition* if there is no infinite chain $a_1 R a_2 R a_3 \ldots$.

The following Lemma is obvious.

Lemma 2.30 *Let $(M, \cdot, 1)$ be factorial. Then the relation aRb that says "b has a as proper divisor" satisfies the minimal condition. For this we say that $(M, \cdot, 1)$ satisfies the* ***divisor chain condition****.*

Definition 7 An element p of $\mathbf{M}$ is *prime* if

- p is not a unit.
- If $p|ab$ then $p|a$ or $p|b$.

A *prime p is always irreducible*: Assume that $ab = p$. Thus $a|p$ and $b|p$. But $p \cdot 1 = ab$ also shows that $p|ab$. Because p is prime, we have $p|a$ or $p|b$. Hence, $a \sim p$ and b is a unit, or vice versa. Thus p does not have proper divisors, i.e., p is irreducible.

Lemma 2.31 *Let $(M, \cdot, 1)$ be factorial. Then every irreducible element is prime. This is* ***Euclid's condition.***

PROOF. Assume that $q|ab$, i.e., $qc = ab$ where q is irreducible.

If a is a unit, then $q(ca^{-1}) = b$, and we have shown that $q|b$.

Thus we may assume that a and b are not units and therefore, $a = q'_1 \cdot \ldots \cdot q'_s$ and $b = q''_1 \cdot \ldots \cdot q''_t$. Thus $ab = q'_1 \cdot \ldots \cdot q'_s q''_1 \cdot \ldots \cdot q''_t$.

If c was a unit then $q = (c^{-1}q'_1) \cdot \ldots \cdot q''_t$ shows that q is not irreducible. Hence, $c = q_1 \cdot \cdots \cdot q_r$.

Thus, $q \cdot (q_1 \cdot \cdots \cdot q_r) = q'_1 \cdot \ldots \cdot q'_s \cdot q''_1 \cdot \ldots \cdot q''_t$ and because of uniqueness of the factorization we have that $q \sim q'_i$ or $q \sim q''_j$. □

Theorem 2.32 *Let $\mathbf{M} = (M, \cdot, 1)$ be a commutative cancellation monoid. Then $\mathbf{M}$ is factorial if and only if*

(i) $\mathbf{M} = (M, \cdot, 1)$ *satisfies the divisor chain condition.*

(ii) $\mathbf{M} = (M, \cdot, 1)$ *satisfies Euclid's condition.*

PROOF. Assume that $\mathbf{M}$ satisfies the divisor chain condition. We are going to show that every element a in $\mathbf{M}$ is a product of irreducibles.

Let $a \in \mathbf{M}$. If a is a unit or irreducible, we are done. Otherwise, we are going to show that a has a proper divisor q_1 that is irreducible.
We assume that a is not a unit and not irreducible. Hence, a has a proper factor a_1. If a_1 is not irreducible, then a_1 has a proper factor a_2 etc. Because of the divisor chain condition, the chain $aRa_1Ra_2 \ldots$ must terminate with some a_j that is an irreducible element q_1.
We now have $a = q_1 \cdot b_1$. If b_1 is not irreducible, b_1 has an irreducible divisor q_2 and $a = q_1 \cdot b_1 = q_1 \cdot q_2 \cdot b_2$. The chain $b_1Rb_2 \ldots$ must terminate for some at some point s and we get $a = q_1 \cdot \ldots \cdot q_s$.

Because of Euclid's condition the factorization of a into irreducibles is essentially unique. Assume $a = q_1 \cdot \ldots \cdot q_s = q'_1 \cdot \ldots \cdot q'_t$. Assume by induction that every element that has some factorization with less than s-many factors, has an essentially unique factorization. The case $s = 1$ is obvious. The element a then is irreducible and cannot have more than one factor. That is, $a = q$ is the only factorization. For the general case, we have $q_1|q'_1 \cdot \ldots \cdot q'_t$ and, according to Euclid's condition, one has $q_1|q'_{i'}$ for some i'. But $q'_{i'}$ is irreducible. Thus $q_1 \sim q'_{i'}$, i.e., $q'_{i'} = q_1 \cdot u_1$ for some unit u_1. We may rearrange the second factorization and get $a = q_1 \cdot \ldots \cdot q_s = q_1 \cdot (u_1 \cdot q'_2) \cdot \ldots \cdot q'_t$. But then we may cancel on both sides q_1, and we are done by induction. □

In every factorial monoid the contraction of the divisibility relation is a lattice ordering. That is, modulo equivalence, every two elements a and b have a *greatest common divisor*, *g.c.d.*(a, b) and a

lowest common multiple, $l.c.m.(a,b)$. Let $\mathcal{P}$ be a complete set of representatives for the set of irreducible elements of $\mathbf{M}$. Then for a and b in $\mathbf{M}$ one has

$$\begin{aligned} a &\sim \prod_{p\in\mathcal{P}} p^{e_a(p)} \\ b &\sim \prod_{p\in\mathcal{P}} p^{e_b(p)} \\ g.c.d.(a,b) &\sim \prod_{p\in\mathcal{P}} p^{min(e_a(p),e_b(p))} \\ l.c.m.(a,b) &\sim \prod_{p\in\mathcal{P}} p^{max(e_a(p),e_b(p))} \end{aligned}$$

Of course, one has for every a that for almost all p, $e_a(p) = 0$.

Let $\mathbf{D}$ be an integral domain. $\mathbf{D}$ is called a *principal ideal domain* if every ideal I is generated by some element a, i.e., $I = (a) = \{b.a | b \in \mathbf{D}\}$. The generator a of I is unique, up to equivalence. By the very definition, $(\mathbf{D} \setminus \{0\}, \cdot, 1)$ is a commutative cancellation monoid. We extend the divisibility relation, $a|b$ iff $ac = b$ for some c, so that it includes zero. But then:

$$1|a \text{ and } a|0$$

that is, [0] is the maximum of the ordered set $(\mathbf{D}/\!\sim, |)$. Now,:

$$a|b \text{ iff } (a) \supseteq (b)$$

shows us that

$$(D/\!\sim, |) \ \cong (Ideal(D, \subseteq)^*$$

Of course, $(Ideal(D, \subseteq)$ is a complete lattice. Hence, we have

Proposition 2.33 *Let* $\mathbf{D}$ *be a principal ideal domain. Then every non-empty subset of* $\mathbf{D}$ *has a g.c.d. and a l.c.m..*

Proof. To be more explicit, let $S \subseteq \mathbf{D}$. We claim that the generator d for (S) is the $g.c.d.(S)$. Indeed, because of $(d) = (S)$ one has that $d|s$ for every $s \in S$. On the other hand, assume that $e|s$ for every $s \in S$. Then $(e) \supseteq (S) = (d)$. This is $e|d$. □

Lemma 2.34 *Let* $\mathbf{D}$ *be a principal ideal domain. Then* $d = g.c.d.(a,b)$ *is the only divisor of* a *and* b *that is of the form* $d = c_1a + c_2b$*, i.e., the only divisor that belongs to the ideal* (a,b)*.*

Proof. For any commutative ring $\mathbf{A}$ one has that $(a,b) = \{c_1a + c_2b | c_i \in \mathbf{A}\}$ is the ideal that is generated by a and b. Thus, if $\mathbf{D}$ is a principal ideal domain then one has for the generator d of this ideal: $d = c_1a + c_2b$. Assume now $d' = c_1a + c_2b$ is a divisor of a and b. Then,

(i) $(d') \supseteq (a,b) = (d)$ because $d'|a$ and $d'|b$.

(ii) $(d') \subseteq (d)$ because $d' \in (a,b)$.

Thus $(d') = (d)$, i.e., $d' = g.c.d.(a,b)$. □

Theorem 2.35 *Let $\mathbf{D} \neq \mathbb{O}$ be a principal ideal domain. Then $(\mathbf{D} - 0, \cdot, 1)$ is factorial, i.e., every element different from zero that is not a unit has an essentially unique factorization into irreducible elements.*

PROOF. We show that $\mathbf{D}$ satisfies the *ascending chain condition* for Ideals. That is any ascending chain $I_1 \subseteq I_2 \subseteq \ldots$ of ideals becomes stationary after some n, i.e., $I_n = I_{n+1} = \ldots$. To show this, one observes that $I = \bigcup_{n\in\mathbb{N}} I_n$ is an ideal and therefore $I = (a)$ for some $a \in I$. But then $a \in I_n$ for some $n \in \mathbb{N}$ and $I = I_n = \ldots$.

But this implies the divisor chain condition. That is, if $a_1, a_2, \ldots$ is any sequence of elements in $\mathbf{D}$ such that $a_{i+1}|a_i$ then $a_n = a_{n+1} = \ldots$.

Next we show that irreducible elements are prime. For the domain of integers, this is usually referred to as Euclid's lemma: Assume $p|ab$ where p is irreducible (one says prime). Then $p|a$ or $p|b$. Assume that $p \not| b$. As an irreducible element, p has only trivial divisors. Thus the g.c.d. of p and b is 1. By the previous lemma, $1 = cp + db$. But then $a = acp + adb$. Now, $p|ab$, by assumption, and $p|acp$. Thus, we have that $p|a$. □

Definition 8 An integral domain $\mathbf{E}$ is called *Euclidean* if there exists some map $\varphi : \mathbf{E} \to \mathbb{N}$ such that if $a, b \in \mathbf{E}$ where $b \neq 0$, then there is some $q \in \mathbf{E}$ such that

$$a = bq \quad \text{or } a = bq + r \text{ where } \varphi(r) < \varphi(b)$$

Examples

1. $\mathbb{Z}$ is Euclidean with $\varphi(n) = |n|$.

2. Let $\mathbf{E}$ be the set of *Gaussian integers*, i.e., the complex numbers α of the form $z = m+ni$ where m and n are integers. Then $\mathbf{E}$ is an integral domain. We define $\varphi : \mathbf{E} \to \mathbb{N}$ by $\varphi(m+ni) = m^2+n^2$. Of course, $\varphi(\alpha\beta) = \varphi(\alpha)\varphi(\beta)$. This formula holds for arbitrary complex numbers.

   For any complex number $\alpha = r_1 + r_2 i$ one can find $m, n \in \mathbb{Z}$ such that $|r_1 - m| \leq 1/2$ and $|r_2 - n| \leq 1/2$. This yields,

   $$\varphi(\alpha - (m + ni)) = \varphi((r_1 - m) - (r_2 - n)i) = |r_1 - m|^2 + |r_2 - n|^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

   Let now $\alpha = m + ni$ and $\beta = m' + n'i \neq 0$. Then choose some $\gamma \in \mathbf{E}$ such that $\varphi(\alpha\beta^{-1} - \gamma) < 1$. But then, $\varphi(\beta)\varphi(\alpha\beta^{-1} - \gamma) = \varphi(\alpha - \gamma\beta) < \varphi(\beta)$ proves our claim.

Theorem 2.36 *A Euclidean domain is a principal ideal domain.*

PROOF. Let I be an ideal of the Euclidean domain $\mathbf{E}$. If I is the zero-ideal then we are done. Otherwise, there is some $b \in I$ where $b \neq 0$. We choose some $a \in I$ for which $\varphi(a)$ is minimal. That is $\varphi(b) \geq \varphi(a)$ for all $b \in I$. We claim that $I = (a)$. Indeed, $I \supseteq (a)$ because $a \in I$ and I is an ideal. On the other hand, if $b \in I$ then b is a multiple of a. If not then we have $b = aq + r$ where $\varphi(r) < \varphi(a)$. But $r \in I$ and $\varphi(r) < \varphi(a)$ is a contradiction. □

Corollary 2.37 *Every Euclidean domain is factorial.*

We have $\mathbb{Z} \subseteq \mathbf{E}$. However, 5 is irreducible in $\mathbb{Z}$ but as $5 = (2+i)(2-i)$ shows, 5 is not irreducible in E.

In $\mathbb{Z}$ the g.c.d. of two integers can be obtained without going through the prime factorization first. This **Euclidean Algorithm** is based on the observation that, if $a = bq + r$, one has that $(a, b) = (b, r)$. Assume that $a > b > 0$. Then, in $a = bq + r$, $b = q_1 r + r_1$, $r = q_2 r_1 + r_2$ etc. we have that $a > b > r > r_1 > r_2 > \ldots$, $(a, b) = (b, r) = (r, r_1) = \ldots$, and eventually we must have $r_{k+1} = 0$ where $r_k > 0$. But then $(a, b) = (r_k, 0)$ shows that r_k is the g.c.d. of a and b.

For a commutative ring $\mathbf{A}$ we have that $I = (p)$ is a prime ideal if $ab \in I$ iff $p|a$ or $p|b$, that is, p is a prime element.

The ideal $I = (q)$ is maximal amongst all principal ideals iff $(a) \supseteq (q)$ implies that $a \sim 1$ or $a \sim q$. This is, $a|q$ implies that $a \sim 1$ or $a \sim q$. Hence, (q) is maximal amongst principal ideals iff q is irreducible. Hence, we have

Proposition 2.38 *Let* $\mathbf{D}$ *be a principal ideal domain. Then the following statements are equivalent.*

(i) a *is a prime element.*

(ii) (a) *is a prime ideal.*

(iii) (a) *is a maximal ideal.*

(iv) a *is irreducible.*

2.5 Polynomial Rings

In elementary courses on algebra, polynomials are defined as formal expressions $f(x) = a_0 + a_1 x + \ldots + a_n x^n$ and one learns that the sum and product of polynomials is a polynomial where these operations are defined in an obvious way. Furthermore, every polynomial stands for a certain function: $f(a)$ is what one gets by substituting a for x. But what are polynomials and what is x? They should be elements of a ring. But what ring? These questions were answered in the thirties by **Hans Hasse** who gave a rigorous definition of the somewhat mysterious concept of a polynomial and the *transcendental* element x. Even today there are still textbooks that take the concept of a polynomial ring for granted without going through its somewhat lengthy "Hasse construction". We will construct polynomials as certain universal objects built upon the commutative ground ring $\mathbf{A}$ and a fixed element x. The universal character of the polynomial ring means that every substitution of x by an element a leads to a homomorphism from the polynomial ring to $\mathbf{A}$.

Definition 9 A *pointed* ring is an ordered pair $(\mathbf{A}, a)$ consisting of

- a commutative ring $\mathbf{A}$;
- an element $a \in \mathbf{A}$.

Definition 10 A *morphism* $\varphi : (\mathbf{A}, a) \to (\mathbf{B}, b)$ *between pointed rings* is

- a homomorphism $\varphi : \mathbf{A} \to \mathbf{B}$ where
- $\varphi(a) = b$.

Theorem 2.39 *For every commutative ring* $\mathbf{A}$ *there is a pointed ring* $(\mathbf{A}', x)$ *together with a homomorphism* $\alpha : \mathbf{A} \to \mathbf{A}'$ *such that there is for every homomorphism*

$$\varphi : \mathbf{A} \to (\mathbf{B}, b)$$

from $\mathbf{A}$ *into a pointed ring* $(\mathbf{B}, b)$ *a unique morphism*

$$\hat{\varphi} : (\mathbf{A}', x) \to (\mathbf{B}, b)$$

between pointed rings where $\hat{\varphi} \circ \alpha = \varphi$. *The pointed ring* $(\mathbf{A}', x)$ *is uniquely determined up to isomorphisms between pointed rings.*

PROOF. We first show **Existence** of the "polynomial ring" over $\mathbf{A}$. We define

$$\mathbf{A}[x] = \{p | p : \mathbb{N} \to \mathbf{A}, p(n) = 0 \text{ for almost all n}\}$$

That is, $\mathbf{A}[x]$ is the free $\mathbf{A}$-module with base:

$$e_i(n) = \begin{cases} 1 & \text{when } i = n, \\ 0 & \text{when } i \neq n. \end{cases}$$

i.e., $e_i(n) = \delta_i^n$ where δ is the Kronecker symbol. We know already that any multiplication on the base makes $\mathbf{A}[x]$ to an algebra over $\mathbf{A}$ by linear extension. We define a multiplication on the base by:

$$e_i \cdot e_j = e_{i+j}$$

Then $(\mathbf{A}[x], +, -, 0, (a.)_{a \in \mathbf{A}}, \cdot, e_0)$ is an associative and commutative algebra with $e_0 = e$ as unit.

Let $p \in \mathbf{A}[x]$. Then $p = \sum_{i \in N}^{*} a_i.e_i$ where the asterisk at the sum sign indicates that almost all coefficients are zero.

Now, $e_n = e_1 \cdots e_n = (e_1)^n$. If we define e_1 as x then we have for every p in $\mathbf{A}[x]$ a unique representation as a "polynomial":

$$\sum\nolimits_{i \in N}^{*} a_i \, x^i = \sum_{i=1}^{n} a_i \, x^i \text{ where } n \text{ is the last coefficient different from zero.}$$

n is called the *degree* of p and a_n is called the *leading coefficient* of p. The empty sum is defined as zero and $p = 0$ does not have a degree. We now define:

$$\alpha : A \to A[x], \; a \mapsto a.e$$

Then α is an injective ring homomorphism. We have:

$$\alpha(a + b) = (a + b).e = a.e + b.e = \alpha(a) + \alpha(b)$$

and

$$\alpha(ab) = (ab).e = (a.e) \cdot (b.e) = \alpha(a) \cdot \alpha(b), \; \alpha(1) = 1.e = e$$

and injectivity of α is obvious, $\{e\}$ is linearly independent.

We are going to show that $\alpha : \mathbf{A} \to \; (\mathbf{A}[x], x)$ is *universal* for all $\varphi : \mathbf{A} \to (\mathbf{A}, b)$. In order to see this, we define the map $\hat{\varphi}$ by

$$\sum a_i \, x^i \; \mapsto \; \sum \varphi(a_i) \, b^i$$

and it is easy to see that this map has the desired properties, i.e., it is a ring homomorphism for which $\hat{\varphi}(\alpha(a)) = \hat{\varphi}(a.e) = \varphi(a)\, b^0 = \varphi(a)$ and $\hat{\varphi}(x) = b$. The multiplicativity of $\hat{\varphi}$ needs the commutativity of $\mathbf{A}$ and of $\mathbf{B}$.

The proof of **Uniqueness** up to isomorphism of the universal object in the category of pointed rings over $\mathbf{A}$ is a by now familiar exercise. □

For every fixed $a \in \mathbf{A}$ we have a homomorphism:

$$E_a : \mathbf{A}[x] \to \mathbf{A}$$

which is defined by:

$$A \xrightarrow{\alpha} (\mathbf{A}[x], x) \xrightarrow{E_a} (A, a) = A \hookrightarrow (A, a)$$

We have that $E_a(p) = \hat{p}(a)$. This is true because $E_a(x) = a$ and $E_a(\alpha(c)) = c = E_a(c.e)$

Now, $\mathbf{A}[x] \xrightarrow{E_a} \mathbf{A},\ a \in \mathbf{A}$, yields according to the definition of the direct product with projections q_a a homomorphism

$$\mathbf{A}[x] \xrightarrow{E} \mathbf{A}^A$$

such that

$$\mathbf{A}[x] \xrightarrow{E} \mathbf{A}^A \xrightarrow{q_a} \mathbf{A} = \mathbf{A}[x] \xrightarrow{E_a} \mathbf{A}$$

$E(p)$ is a map from A to $\mathbf{A}$ which has at a the value $q_a(E(p)) = E_a(p) = \hat{p}(a)$

If $p = \sum a_i\, x^i$ then $E(p)(a) = \sum a_i\, a^i = \hat{p}(a)$.

Theorem 2.40 *The evaluation map*

$$E : \mathbf{A}[x] \to \mathbf{A}^A,\ p = \sum a_i\, x^i \ \mapsto\ \hat{p} = (a \ \mapsto \sum a_i\, a^i,\ a \in A)$$

is a homomorphism between the ring of polynomials and the ring of polynomial functions.

Because the map $\alpha, a \ \mapsto a.e$, is an injective homomorphism we may identify a with $a.e$, and consider $\mathbf{A}$ as a subring of $\mathbf{A}[x]$. With this identification, we have that $a \cdot p = \alpha(a) \cdot p = (a.e) \cdot p = a.(e \cdot p) = a.p$. The degree of a polynomial p, $\deg(p)$, has been defined earlier. We then have the degree rules

$$\begin{aligned} \deg(p+q) &\leq \max(\deg(p), \deg(q)) \\ \deg(p \cdot q) &\leq \deg(p) + \deg(q) \end{aligned}$$

where equality in the product formula holds if the leading coefficient of one of the factors is not a zero divisor. If we define the degree of the zero polynomial to be $-\infty$ then these formulas are universally true. We conclude from the product formula:

Proposition 2.41 *For an integral domain* $\mathbf{D}$*, the polynomial ring* $\mathbf{D}[x]$ *is an integral domain.*

Theorem 2.42 (Division Algorithm) *Let* p *and* d *be polynomials in* $\mathbf{A}[x]$ *and assume that the leading coefficient of* d *is an invertible element in* $\mathbf{A}$*. Then there exist unique polynomials* q *and* r *such that:*

$$p = q \cdot d + r,\ \deg(r) < \deg(d).$$

PROOF. We first show **Uniqueness** of the *quotient* q and *remainder* r. Indeed, $f = q_1 d + r_1 = q_2 d + r_2$ yields $0 = (q_1 - q_2)d + (r_1 - r_2)$. This is, $(q_1 - q_2)d = (r_2 - r_1)$. Now, $\deg(r_1) < \deg(d)$ and $\deg(r_2) < \deg(d)$. But $\deg((q_1 - q_2)d) \geq \deg(d)$ or $-\infty$. Because d is not a zero divisor, the latter is only the case if $(q_1 - q_2) = 0$, i.e., if $q_1 = q_2$. But then also $r_1 = r_2$.

The part of **Existence** is just the familiar *long division*. First, if $\deg(p) < \deg(d)$ then $p = 0d + p$ is the desired decomposition. Thus, we may assume that $\deg(d) \leq \deg(p)$:

$$p = a_0 + a_1 x + \ldots + a_n x^n,\ d = b_0 + b_1 x + \ldots + b_m x^m \text{ where } m \leq n$$

The proof now goes by induction over $n = \deg(p)$.

If $n = 0$ then $m = 0$ and $p = a_0$, $d = b_0$ where b_0 has an inverse, and $a_0 = (a_0 \cdot b_0^{-1})b_0 + 0$.

Assume that $n > 0$. Then,

$$(a_0 + a_1 x + \ldots + a_n x^n) - (b_0 + b_1 x + \ldots + b_m x^m) \cdot (a_n b_m^{-1}) x^{n-m} = p - (a_n b_m^{-1}) x^{n-m} d = p_1$$

where $\deg(p_1) < \deg(p)$. By induction hypothesis,

$$p_1 = q_1 d + r_1 \text{ where } \deg(r_1) < \deg(d) \text{ or } r_1 = 0$$

Hence,

$$p = (q_1 d + r_1) + (a_n b_m^{-1}) x^{n-m} d = ((a_n b_m^{-1}) x^{n-m} + q_1)d + r_1 = q\, d + r_1$$

□

Let $\mathbf{F}$ be a field. Then every element different from zero has an invertible leading coefficient. Thus, for every $p \in \mathbf{F}[x]$ and $d \neq 0$ one has some $q \in \mathbf{F}[x]$ such that $p = q \cdot d$ or $\deg(p - q \cdot d) = \deg(r) < \deg(d)$. This shows that the degree map $\deg : \mathbf{F}[x] \setminus \{0\} \to \mathbb{N}$ makes $\mathbf{F}[x]$ to a Euclidean ring:

Theorem 2.43 *For any field* $\mathbf{F}$*, the polynomial ring* $\mathbf{F}[x]$ *is a Euclidean domain. One has that* p *is a unit, if and only if,* $\deg(p) = 0$*, if and only if,* $p \in \mathbf{F}[x] - 0$.

Corollary 2.44 $\mathbf{F}[x]$ *is a principal ideal domain.*

Corollary 2.45 $\mathbf{F}[x]$ *is factorial.*

As a corollary of Theorem 2.42 we note:

Proposition 2.46 *Let* $p \in \mathbf{A}[x]$, $a \in \mathbf{A}$*. Then* $p = q \cdot (x - a) + r$*, where* $r \in \mathbf{A}$ *and* $\hat{p}(a) = r$.

Corollary 2.47 *The element* $a \in \mathbf{A}$ *is a root of* p*, i.e.,* $\hat{p}(a) = 0$*, if and only if,* $(x - a) \mid p$.

Corollary 2.48 *Let* $\mathbf{D}$ *be an integral domain. Then a polynomial* $p \in \mathbf{D}[x]$ *has at most* n *roots in* $\mathbf{D}$.

PROOF. We prove this by induction on $n = \deg(p)$. If $n = 0$ then $p = a \neq 0$ and p has no root.

Let $n = \deg(p)$ and let a be a root of p. Then $p = (x - a)p_1$ and $\hat{p}(b) = 0$ if and only if $a = b$ or $\hat{p}_1(b) = 0$. Here we use that $\mathbf{D}$ is a domain. Now, $\deg(p_1) = n - 1$ proves the claim. □

Corollary 2.49 *For an infinite integral domain* $\mathbf{D}$*, the map*

$$E : \mathbf{D}[x] \to \mathbf{D}^D$$

is injective.

Proposition 2.50 *Let* $p \in \mathbf{A}[x]$ *and let* $a \in \mathbf{A}$*. Then the set* $\{l \mid (x-a)^l | p\}$ *has a maximum* m *and if* $p = (x-a)^m \cdot q$ *then* $\hat{q}(a) \neq 0$*. Conversely, each relation* $p = (x-a)^l \cdot r$ *where* $\hat{r}(a) \neq 0$ *implies that* $l = m$*. The natural number* m *is called the* ***multiplicity*** *of* a *for* p*.*

Proof. Assume that $p = (x-a)^l \cdot q$. Then $\deg(p) = n = l + \deg(q)$. Thus, $l \leq n$ and a maximum m must exist. If $p = (x-a)^m \cdot q$, then $\hat{q}(a) = 0$ would yield $(x-a)|q$ and therefore $(x-a)^{m+1}|p$, which is a contradiction to the choice of m.
Assume now that $p = (x-a)^l \cdot r$ where $\hat{r}(a) \neq 0$. Then $m \geq l$ and $(x-a)^l(r-(x-a)^{m-l} \cdot q) = p - p = 0$. Hence, because $(x-a)^l$ is not a zero divisor, we have that $r = (x-a)^{m-l} \cdot q$. Now, $m > l$ would yield $\hat{r}(a) = 0$, a contradiction. Thus, $m = l$. □

Definition 11 The *derivative* of the polynomial $p = a_0 + a_1\, x + \ldots + a_n\, x^n$ is defined as the polynomial $\delta(p) = a_1 + \ldots + n\, a_n\, x^{n-1}$. One then has the familiar rules, e.g., $\delta(p+q) = \delta(p) + \delta(q)$ and $\delta(p \cdot q) = \delta(p) \cdot q + p \cdot \delta(q)$

Proposition 2.51 *For any polynomial* p *of the ring* $\mathbf{A}[x]$ *one has that* $a \in \mathbf{A}$ *is a multiple root if and only if it is a root of* p *and of* $\delta(p)$*.*

Proof. Assume that a is a root of p. We then have $p = (x-a)^m \cdot q$ where $\hat{q}(a) \neq 0$.
Assume that $m > 1$, i.e., a is a multiple root. Then $\delta(p) = m\,(x-a)^{m-1}\, q + (x-a)^m\, \delta(q)$ yields $\delta(p)(a) = 0$.
If $m = 1$, i.e., a is a simple root, then $\delta(p)(a) = \hat{q}(a) \neq 0$. □

Chapter 3

Modules over a Principal Ideal Domain

3.1 Free Modules

Let $\mathbf{M}$ be a module over a principal ideal domain $\mathbf{D}$ and let $x \in \mathbf{M}$. Then

$$\gamma_x : \mathbf{D} \to \mathbf{M},\ d \mapsto d.x$$

is a homomorphism between $\mathbf{D}-$modules. Clearly,

$$\text{im}(\gamma_x) = \{d.x | d \in \mathbf{D}\} =< x >= \text{span of x}$$

is a $\mathbf{D}-$submodule of $\mathbf{M}$ and

$$\ker(\gamma_x) = \{d | d.x = 0\}$$

is an ideal of $\mathbf{D}$

A generating element of $\ker(\gamma_x)$ is called the *period*, $\text{per}(x)$, of x. Thus,

$$\mathbf{D}/(\text{per}(x)) \cong < x >$$

Example

Let $\mathbf{A}$ be an abelian group and let $a \in A$. Then $\mathbf{A}$ is a $\mathbb{Z}-$ module and

$$< a > \cong \mathbb{Z}/(per(a))$$

Now, $\text{per}(a) = \min\{n \mid n > 0, n.a = 0\}$ or $\text{per}(a) = 0$. Thus,

$$< a > \cong \mathbb{Z}_n \text{ or } < a > \cong \mathbb{Z}$$

A cyclic group $< a >$ is either finite and isomorphic to $\mathbb{Z}_n$ or isomorphic to $\mathbb{Z}$. More generally: *A cyclic module $< x >$ over a principal ideal domain $\mathbf{D}$ is isomorphic to a factor module $\mathbf{D}/(d_x)$ where $d_x = per(x)$.*

If $x \in M$ and if $d.x = 0$ then $\mathrm{per}(x)|d$. Let X be a generating set of $\mathbf{M}$, i.e., $\mathbf{M} =< X >$. Then one has:
$d.x = 0$ for all $x \in M$ iff $d.x = 0$ for all $x \in X$ iff $d \in (per(x))$ for all $x \in X$ iff $d \in \bigcap_{x\in X}(per(x))$ iff $d \in (l.c.m._{x\in X}(per(x))$. Thus,

$$per(\mathbf{M}) = \{d \mid d.x = 0 \text{ for all } x \in M\} = (l.c.m._{x\in X}(per(x))$$

Notice that: If $\mathrm{per}(x) = 1$ then $\mathbf{D}/(1) \cong < x >$, thus $\mathbf{0} \cong < x >$ which is $x = 0$. On the other hand one has that $\mathrm{per}(0) = 1$. Thus:

$$\mathrm{per}(x) = 1 \text{ iff } x = 0$$

If $\mathrm{per}(x) = 0$ then $\mathbf{D}/(0) \cong < x >$ which is $\mathbf{D} \cong < x >$. On the other hand, assume $\mathbf{D} \cong < x >$ and let $\varphi \to < x >$ be any isomorphism. Assume that $1 \mapsto a.x$. Then $\mathrm{per}(x) \mapsto (\mathrm{per}(x) \cdot a).x$ shows $\varphi(\mathrm{per}(x)) = 0$, thus $\mathrm{per}(x) = 0$. Hence,

$$\mathrm{per}(x) = 0 \text{ iff } < x > \text{ is free.}$$

Recall, that a module over the ring $\mathbf{A}$ is free, freely generated by the subset B if every map φ from B into any $\mathbf{A}-$module $mbfM$ has a unique linear extension. Such a module is isomorphic to $\mathbf{A}^{(B)}$.
Modules over fields are always free and all bases have the same cardinal.

Theorem 3.1 *Let $\mathbf{M}$ be a free module over the principal ideal domain $\mathbf{D}$. Then all bases B of $\mathbf{M}$ have the same cardinal which is called the **dimension** of $\mathbf{M}$ over $\mathbf{D}$.*

Proof. Let $p \in \mathbf{D}$ be a prime element. Then $\mathbf{D}/(p)$ is a field and $p.\mathbf{M}$ is a $\mathbf{D}$-submodule of $\mathbf{M}$. Thus we can form the factor module $\mathbf{M}/p.\mathbf{M}$. Let $d + (p) \in \mathbf{D}/(p)$ and $x + p.\mathbf{M} \in \mathbf{M}/p.\mathbf{M}$. Then:

$$d + (p).x + p.\mathbf{M} = d.x + p.\mathbf{M}$$

is a proper definition. That is, if $d_1 - d_2 \in (p)$ and $x_1 - x_2 \in p.\mathbf{M}$ then $d_1.x_1 - d_2.x_2 \in p.\mathbf{M}$. Thus $\mathbf{M}/p$ becomes a module over the field $\mathbf{D}/(p)$. Let X be a base of $\mathbf{M}$. Then $[X] = \{[x] \mid x \in X\}$ is a base of $\mathbf{M}/p.\mathbf{M}$. Of course, $[X]$ is generating for $\mathbf{M}/p.\mathbf{M}$. With respect to linear independence, assume that $\sum_{x\in X}[d_x][x] = [0]$, i.e., $\sum d_x.x = p.y = p.\sum_{x\in X} e_x.x = \sum_{x\in X} p.e_x.x$. Then one has that $d_x = p \cdot e_x$ holds for all $x \in X$. Hence, $[d_x] = [0]$ for all $x \in X$.
In particular, $x \mapsto [x]$ has to be injective, or $\mathrm{card}(X) = \mathrm{card}[X]$ where $\mathrm{card}(X) = \dim_{\mathbf{D}/(p)}(\mathbf{M}/p.\mathbf{M})$.
□

Theorem 3.2 *Let $\mathbf{M}$ be a free module over the principal ideal domain $\mathbf{D}$ and let $\mathbf{N}$ be a $\mathbf{D}-$submodule of $\mathbf{M}$. Then $\mathbf{N}$ is also free and $\dim_{\mathbf{D}}(\mathbf{N}) \leq \dim_{\mathbf{D}}(\mathbf{M})$.*

Proof. Let $x_1, \ldots, x_n$ be a base of $\mathbf{M}$ and let

$$\mathbf{N}_r = \mathbf{N} \cap < x_1, \ldots, x_r >, \text{r=1,...,n}$$

Clearly, $\mathbf{N}_n = \mathbf{N}$. We show successively that each $\mathbf{N}_r$ is free and that $\dim(\mathbf{N}_r) \leq r$.
For $n = 1$ we have that:

$$\mathbf{N}_1 = \mathbf{N} \cap < x_1 >= \{d.x_1 \mid d \in \mathbf{D}, d.x_1 \in \mathbf{N}\}$$

Now, $\{d \mid d.x_1 \in < x_1 > \cap \mathbf{N}\}$ is an ideal of $\mathbf{D}$, thus a principal ideal $I_1 = (d_1)$ and we have:

$$\mathbf{N}_1 = \mathbf{N} \cap < x_1 >=< d_1.x_1 >$$

The case $d_1 = 0$ is clear: $\mathbf{N}_1 = \mathbb{O}$, and $\dim(\mathbf{N}_1) = 0 < 1$.
Assume that $d_1 \neq 0$. Note, if $d.(d_1.x_1) = 0)$ one has that $\mathrm{per}(x_1)|d \cdot d_1$. But $\mathrm{per}(x_1) = 0$, thus $d \cdot d_1 = 0$ which yields $d = 0$. Hence, $\mathrm{per}(d_1.x_1) = 0$ and therefore, $\mathbf{N}_1 =< d_1.x_1 >$ is free and $\dim(\mathbf{N}_1) = 1 \leq 1$.
Now assume that $\mathbf{N}_r$ is free and that $\dim(\mathbf{N}_r) \leq r$. Similarly as before, we get an ideal I_{r+1} of $\mathbf{D}$:

$$\mathbf{N}_{r+1} = \mathbf{N} \cap < x_1, \ldots, x_r, x_{r+1} >= \{\sum_{i=1}^{r+1} e_i.x_i \mid e_i \in \mathbf{D}\} \cap \mathbf{N}$$

where

$$I_{r+1} = \{d \mid \exists_{e_1,\ldots,e_r}(e_1.x_1 + \ldots + e_r.x_r + d.x_{r+1}) \in \mathbf{N}_{r+1}\}$$

is an ideal of $\mathbf{D}$ and therefore $I_{r+1} = (d_{r+1})$.
If $d_{r+1} = 0$ then $\mathbf{N}_{r+1} = \mathbf{N}_r$ and $\mathbf{N}_{r+1}$ is free and $\dim(\mathbf{N}_{r+1}) \leq r < r + 1$.
Otherwise we have some $y \in \mathbf{N}_{r+1}$ where

$$y = e_1.x_1 + \ldots + e_r x_r + d_{r+1}.x_{r+1} \text{ where } d_{r+1} \neq 0$$

Let $y_1, \ldots, y_k$ be a base of $\mathbf{N}_r$, where $k \leq r$. Then our claim is that

$$\{y_1, \ldots, y_k, y\} \text{ is a base of } \mathbf{N}_{r+1}$$

We first show that$\{y_1, \ldots, y_k, y\}$ generates $\mathbf{N}_{r+1}$. Let $x \in \mathbf{N}_{r+1}$. Then $x = f_1.x_1 + \ldots + f_r.x_r + d.x_{r+1}$ where $d \in I_{r+1}$, thus $d = f.d_{r+1}$. We get

$$f.y = fe_1.x_1 + \ldots fe_r.x_r + fd_{r+1}x_{r+1}$$

Therefore, $fd_{r+1}.x_{r+1} = fy - fe_1.x_1 - \ldots - fe_r.x_r$ and

$$x = f_1.x_1 + \ldots + f_r.x_r + fd_{r+1}.x_{r+1} = f_1.x_1 + \ldots + f_r.x_r + (fy - fe_1.x_1 - \ldots - fe_r.x_r) = g_1.x_1 + \ldots + g_r.x_r + f.y$$

Hence, $x - f.y \in \mathbf{N}_{r+1} \subseteq \mathbf{N}$ but also $x - f.y \in < x_1, \ldots, x_r >$. This is, $x - f.y \in \mathbf{N}_r =< y_1, \ldots, y_k >$ and, finally, $x \in < y_1, \ldots, y_k, y >$.
We need to show that $\{y_1, \ldots, y_k, y\}$ is a linearly independent set. Assume that $f_1.y_1 + \ldots + f_k.y_k + f.y = 0$. Then:

$$\underbrace{f_1.y_1 + \ldots + f_k.y_k}_{\in <x_1,\ldots,x_r>} = -f.y = -f.\underbrace{(e_1.x_1 + \ldots + e_r.x_r + d_{r+1}.x_{r+1})}_{\in <x_1,\ldots,x_{r+1}>}$$

Thus, $-f.d_{r+1} = 0$. From this it follows that $f = 0$ and therefore $f_1.y_1 + \ldots + f_k.y_k = 0$ which is $f_1 = \ldots = f_k = 0$. □

Corollary 3.3 *For modules over principal ideal domains, submodules of finitely generated modules are finitely generated.*

Proof. Let $\mathbf{M} =< x_1, \ldots, x_k >$ be finitely generated by k elements and let $\mathbf{N}$ be a submodule of $\mathbf{M}$. Then let $\mathbf{D}^k$ be the free $\mathbf{D}-$module, which is freely generated by the k unit vectors e_i. We then have a unique surjective homomorphism φ which extends $e_i \mapsto x_i$. The counter image $\mathbf{L} = \varphi^{-1}(\mathbf{N})$ of $\mathbf{M}$ is a submodule of $\mathbf{D}^k$ and finitely generated by $l-$many elements, $\{y_1, \ldots, y_l\}$ where $l \leq k$. Of course, $\varphi(\mathbf{L}) = \mathbf{N}$ and $\mathbf{N} =< \varphi(y_1), \ldots, \varphi(y_l) >$ □

For a $\mathbf{D}-$module $\mathbf{M}$,

$$\mathrm{Tor}_{\mathbf{D}}(\mathbf{M}) = \{x \mid \mathrm{per}(x) \neq 0\}$$

is a submodule of $\mathbf{M}$. The module $\mathbf{M}$ is called a *torsion module* in case that $\mathrm{Tor}(\mathbf{M}) = \mathbf{M}$ and $\mathbf{M}$ is called *torsion free* in case that $\mathrm{Tor}(\mathbf{M}) = \mathbb{O}$.
A finite module is always a torsion module. The module $\mathbb{Z}_2^{\mathbb{N}}$ is an example of an infinite torsion module.

A product of torsion free modules is torsion free. The sum of torsion modules is a torsion module. For a domain $\mathbf{D}$*, the one dimensional* $\mathbf{D}-$*module* $\mathbf{D}$ *is torsion free. For a domain* $\mathbf{D}$*, any free* $\mathbf{D}-$*module is torsion free.*

Corollary 3.4 *A finitely generated torsion free module over a principal ideal domain is free.*

PROOF. Let $Y = \{y_1, \ldots, y_n\}$ be a set of generators for $\mathbf{M}$ where $\mathbf{M}$ is torsion free. We may assume that $Y' = \{y_1, \ldots, y_k\}$ is a maximal linearly independent subset of Y and the submodule $\mathbf{N}$ of $\mathbf{M}$ which is generated by Y' is of course free. However, one cannot conclude that Y' is a base for $\mathbf{M}$. For example, $\{2, 3\}$ is a linearly dependent set that generates $\mathbb{Z}$ but neither 2 or 3 is a base. We can only conclude that $\mathbf{M}$ is isomorphic to a submodule of the free module $\mathbf{N}$.
In order to show this, we notice that for $i > k$ one has some $d_i \neq 0$ such that

$$d_1.y_1 + \ldots + d_k.y_k + d_i.y_i = 0$$

Then let

$$d = \prod_{i=k+1}^{n} d_i$$

Clearly, $d.y_i \in < y_1, \ldots, y_k >$ for $i = k+1, \ldots n$. Then

$$\varphi : \mathbf{M} \to \mathbf{N}, m \mapsto d.m$$

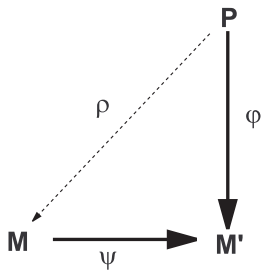
is linear and it is injective because $\mathbf{M}$ is torsion free. □

Corollary 3.5 *A finitely generated module over a principal ideal domain is free if and only if it is torsion free.*

3.2 The Decomposition into a Free and Torsion Part

A $\mathbf{D}-$module $\mathbf{A}$ is the *direct sum* of finitely many submodules $\mathbf{A_i}$ if every element a of $\mathbf{A}$ is a unique sum of elements $a_i \in \mathbf{A_i}$. This is the same as saying that $\mathbf{A}$ is the co-product of its submodules $\mathbf{A_i}$. We also note that for finitely many modules, product and co-product are the same. We use the "$\oplus$sign to indicate that a sum is direct. If $\mathbf{C} = \mathbf{A} + \mathbf{B}$ then the sum is direct if and only if $\mathbf{A} \cap \mathbf{B} = \mathbb{O}$. If $\mathbf{C} = \mathbf{A} \oplus \mathbf{B}$ then $\mathbf{C}/\mathbf{A} \cong \mathbf{B}$.

Definition 12 Let $\mathbf{A}$ be a ring. A module $\mathbf{P}$ is called *projective* if for any homomorphism $\varphi : \mathbf{P} \to \mathbf{M}'$ and any surjective homomorphism $\psi : \mathbf{M} \twoheadrightarrow \mathbf{M}'$, there is a homomorphism $\rho : \mathbf{P} \to \mathbf{M}$ such that $\psi \circ \rho = \varphi$.



Proposition 3.6 *A free module over a ring* $\mathbf{A}$ *is projective.*

PROOF. With the notation of the previous definition, let X be a base of $\mathbf{P}$ where $\mathbf{P}$ is free. Then let $m_x \in \psi^{-1}(\varphi(x))$. There is a unique homomorphism $\rho : \mathbf{P} \to \mathbf{M}, x \mapsto m_x$. Clearly, $\psi(\rho(x)) = \varphi(x)$ and therefore, $\psi \circ \rho = \varphi$ □

Proposition 3.7 *Let* $\psi : \mathbf{M} \twoheadrightarrow \mathbf{P}$ *be a surjective homomorphism from the module* $\mathbf{M}$ *onto the projective module* $\mathbf{P}$. *Then* $\ker(\psi)$ *is a direct summand. That is,* $\mathbf{M} = \ker(\psi) \oplus \mathbf{P}'$ *where, of course,* $\mathbf{P}' \cong \mathbf{P}$.

PROOF. We put $\mathbf{M}' = \mathbf{P}$ and let $\phi = id$. Then there is some ρ such that $\psi \circ \rho = id$ and we set:

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

Then $\psi(m - \rho(\psi(m)) = \psi(m) - \psi(m) = 0$. For any $m' \in \ker(\psi) \cap \rho(\psi((M))$ one has that $m' = \rho(\psi(m))$ for some m. Therefore, $\psi(m') = \psi(\rho(\psi(m)) = \psi(m)$. But then, $\psi(m') = 0 = \psi(m)$, thus $m' = \rho(\psi(m)) = 0$. This is

$$\ker(\psi) \cap \rho(\psi((M))) = \mathbb{O}$$

and $\mathbf{M}/\ker(\psi) \cong \mathbf{P} \cong \mathbf{P}'$ where $\mathbf{P}' = \rho(\psi M)$. □

Let $\mathbf{M}$ be a module over a principal ideal domain $\mathbf{D}$. Then the factor $\mathbf{M}/\mathrm{Tor}(\mathbf{M})$ is torsion free. Assume that $d.(x + \mathrm{Tor}(\mathbf{M})) = \mathrm{Tor}(\mathbf{M})$. Then $d.x \in \mathrm{Tor}(\mathbf{M})$ for some $d \neq 0$. But this is $\mathrm{per}(d.x) \neq 0$ and therefore $e.d.x = 0$ for some $e \neq 0$. Thus, $\mathrm{per}(x) \neq 0$, i.e., $x \in \mathrm{Tor}(\mathbf{M})$.

Theorem 3.8 *Let* $\mathbf{M}$ *be a finitely generated module over a principal ideal domain* $\mathbf{D}$. *Then* $\mathbf{M}$ *is the direct sum of its torsion submodule and of a submodule which is free:*

$$\mathbf{M} = \mathit{Tor}(\mathbf{M}) \oplus \mathbf{N} \text{ where } \mathbf{N} \text{ is free.}$$

The decomposition is unique in the following sense. If

$$\mathbf{M} = \mathbf{M}' \oplus \mathbf{N}' \text{ where } \mathbf{M}' \text{ is a torsion module and } \mathbf{N}' \text{ is torsion free}$$

then $\mathbf{M}' = \mathit{Tor}(M)$ *and* $\mathbf{N}' \cong \mathbf{N}$.

PROOF. The module $\mathbf{M}/\mathrm{Tor}(\mathbf{M})$ is torsion free and as a homomorphic image of a finitely generated module also finitely generated. Thus, $\mathbf{M}/\mathrm{Tor}(\mathbf{M})$ is free. We apply the last proposition to

$$q : \mathbf{M} \twoheadrightarrow \mathbf{M}/\mathrm{Tor}(\mathbf{M}), \ker(q) = \mathrm{Tor}(\mathbf{M})$$

and get:

$$\mathbf{M} = \mathrm{Tor}(\mathbf{M}) \oplus \mathbf{N}$$

where $\mathbf{M}/\mathit{Tor}(\mathbf{M}) \cong \mathbf{N}$.

Let $\mathbf{M} = \mathbf{M}' \oplus \mathbf{N}'$ be a similar decomposition and assume that $m \in \mathrm{Tor}(\mathbf{M})$. We have $m = m' + n'$ and $d.m = 0$ for some $d \neq 0$ gives $d.m' = 0, d.n' = 0$. Because $\mathbf{N}'$ is torsion free, we get that $n' = 0$. Hence, $m = m' \in \mathbf{M}'$ and this is $\mathrm{Tor}(\mathbf{M}) \subseteq \mathbf{M}'$. But $\mathbf{M}' \subseteq \mathrm{Tor}(\mathbf{M})$ holds by assumption. Thus, $\mathrm{Tor}(\mathbf{M}) = \mathbf{M}'$ and $\mathbf{N}' \cong \mathbf{M}/\mathbf{M}' \cong \mathbf{N}$. □

3.3 The Primary Decomposition Theorem

Let $\mathbf{M}$ be a module over the principal ideal domain $\mathbf{D}$ and let $d \in \mathbf{D}$. Then multiplication by d:

$$\mu_d : \mathbf{M} \to \mathbf{M}, x \mapsto d.x$$

is linear and

$$\mathbf{M}(d) = \ker(\mu_d) = \{x \mid d.x = 0\} = \{x \mid \mathrm{per}(x)|d\}$$

is a $\mathbf{D}-$submodule of $\mathbf{M}$.

Lemma 3.9 *Let $(d_1, d_2) = 1$ and $d = d_1 \cdot d_2$, Then:*

$$\mathbf{M}(d) = \mathbf{M}(d_1) \oplus \mathbf{M}(d_2)$$

Proof. We have $1 = e_1 \cdot d_1 + e_2 \cdot d_2$ and therefore

$$(*)\ x = e_1 d_1.x + e_2 d_2.x$$

for every $x \in \mathbf{M}$. Let now $x \in \mathbf{M}(d)$, i.e., $d.x = 0$. Then one has that $d_1 e_2 d_2.x = 0$ and $d_2 e_1 d_1.x = 0$, i.e., $e_2 d_2.x \in \mathbf{M}(d_1), e_1 d_1.x \in \mathbf{M}(d_2)$ and, of course, $\mathbf{M}(d_1), \mathbf{M}(d_2)$ are submodules of $\mathbf{M}$. This is $\mathbf{M}(d) = \mathbf{M}(d_1) + \mathbf{M}(d_2)$. If $x \in \mathbf{M}(d_1) \cap \mathbf{M}(d_2)$ then $x = 0$ by $(*)$. □

Let now $\mathbf{M}$ be a finitely generated torsion module over the principal ideal domain $\mathbf{D}$.
If $\mathbf{M} =< x_1, \ldots, x_n >$ then it is an easy exercise to show that $\mathrm{per}(\mathbf{M}) = \text{l.c.m.}(\mathrm{per}(x_i)) = d \neq 0$. Then, if $d = p_1^{\nu_1} \ldots p_k^{\nu_k}$, one has that

$$\mathbf{M} = \mathbf{M}(p_1^{\nu_1}) \oplus \ldots \oplus \mathbf{M}(p_k^{\nu_k})$$

Definition 13 Let p be a prime of the principal ideal domain $\mathbf{D}$. A *$p-$module* is a $\mathbf{D}-$module where the period of every element is a power of p.

Clearly, for every module $\mathbf{M}$ over the principal ideal domain $\mathbf{D}$, and every prime $p \in \mathbf{D}$, one has that

$$T_P(\mathbf{M}) = \{x \mid \mathrm{per}(x) = p^r, r \geq 0\}$$

is the largest $p-$submodule of $\mathbf{M}$.

Lemma 3.10 *Let $\mathbf{M}$ be a $\mathbf{D}-$module of period p^r. Then $\mathbf{M}$ is a $p-$module and $\mathbf{M}$ contains an element x_0 such that per(x_0) = per$(\mathbf{M}) = x^r$.*

Proof. We have $\mathrm{per}(x) \mid \mathrm{per}(\mathbf{M})$ for every $x \in \mathbf{M}$ and therefore, $\mathrm{per}(x) = p^s$ where $s \leq r$. Thus $\mathbf{M}$ is a $p-$module. Let now x_0 be an element of $\mathbf{M}$ with $\mathrm{per}(x_0) = p^{s_0}$ where s_0 is maximal. Then one has that $p^s = \mathrm{per}(x) | \mathrm{per}(x_0) = p^{s_0}$ for every x. Thus, $p^{s_0}.x = 0$ for every x, i.e., $\mathrm{per}(\mathbf{M}) | p^{s_0}$. This is $p^r | p^{s_0}$ or $r \leq s_0$. Therefore, $r = s$ because we have already shown that $s \leq r$. □

Lemma 3.11 *Let $\mathbf{M}$ be a $\mathbf{D}-$module of period $d = p_1^{\nu_1} \ldots p_k^{\nu_k}$. Then one has that $\mathbf{M}(p_i^{\nu_i}) = T_{p_i}(\mathbf{M})$.*

Proof. We have already shown that $\mathbf{M} = \mathbf{M}(p_1^{\nu_1}) \oplus \ldots \oplus \mathbf{M}(p_k^{\nu_k})$ holds for any module of finite period $d = p_1^{\nu_1} \ldots p_k^{\nu_k}$. Here, finite means larger than 0. We have $\mathbf{M}(p_1^{\nu_1}) = \{x \mid p_1^{\nu_1}.x = 0\} = \{x \mid \mathrm{per}(x) | p_1^{\nu_1}\} \subseteq T_{p_1}(\mathbf{M})$. For the other direction, let $x \in T_{p_1}(\mathbf{M})$, i.e., $p^s.x = 0$ for some s. We have that $x = x_1 + \ldots + x_k$ and therefore, $p_1^s.x = p_1^s.x_1 + \ldots + p_1^s.x_k = 0$ It follows that $p_1^s.x_1 = \ldots = p_1^s.x_k = 0$. Now, $p_1^s.x_2 = 0$ yields $\mathrm{per}(x_2) | p_1^s$ but $\mathrm{per}(x_2)$ is a power p_2^r of p_2. But this is only possible if $r = 0$. Thus, $x_2 = 0$, etc, and we get $x = x_1 \in \mathbf{M}(p_1^{\nu_1})$. □

A *primary* module is a $p-$module for some $p \in \mathbf{D}$.

Theorem 3.12 *A finitely generated torsion module over a principal ideal domain* $\mathbf{D}$ *is a finite direct sum of primary modules. If* $per(\mathbf{M}) = p_1^{\nu_1} \dots p_k^{\nu_k}$ *then*

$$\mathbf{M} = T_{p_1}(\mathbf{M}) \oplus \dots \oplus T_{p_k}(\mathbf{M})$$

where each $T_{p_i}(\mathbf{M})$ *is the largest* p_i-*submodule of* $\mathbf{M}$ *and one has that* $per(T_{p_i}) = p_i^{\nu_i}$.

As an example, let $\mathbf{A}$ be a finitely generated abelian group. We may consider $\mathbf{A}$ as a module over the domain $\mathbb{Z}$ of integers. We then have:

$$\mathbf{A} = \mathbf{B} \oplus \mathrm{Tor}(\mathbf{A})$$

where $\mathbf{B} \cong \mathbb{Z}^r$ and where $\mathrm{Tor}(\mathbf{A})$ is the direct sum of primary groups $T_{p^k} = \{a \mid p^k.a = 0\}$.
We are going to show that each such T_{p^k} is a direct sum of cyclic groups, i.e., groups that are isomorphic to $\mathbb{Z}/(p^l)$.

3.4 The Decomposition according to Elementary Divisors and Invariant Factors

We are going to show that every $\mathbf{D}-$module $\mathbf{M}$ is a direct sum of cyclic primary modules. For example, all finite abelian groups of order 24 are given by the list: $\mathbb{Z}_3 \oplus \mathbb{Z}_8, \mathbb{Z}_3 \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_2, \mathbb{Z}_3 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$. This is a listing according to the "elementary divisors": $(2,2,2,3), (4,2,3), (8,3)$. A different listing of the same groups is $\mathbb{Z}_{24}, \mathbb{Z}_{12} \oplus \mathbb{Z}_2, \mathbb{Z}_6 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2$ according to the "invariant factors": $24, (12,2), (6,2,2)$. This will be generalized: every $\mathbf{D}-$module is a direct sum of cyclic primary modules $\mathbf{D}/(p^j)$ where the p^j are called elementary divisors. Every $\mathbf{D}-$module is also a direct sum $\mathbf{D}/(q_1) \oplus \mathbf{D}/(q_2) \oplus \dots \oplus \mathbf{D}/(q_k)$ where $q_k | q_{k-1} | \dots | q_1$ is the list of invariant factors.
The proof needs some preparations in form of several technical lemmas.

Lemma 3.13 *Assume that* $per(x) \neq 0$. *Then one has that :*

$$per(c.x) = \frac{per(x)}{(c, per(x))}$$

Proof. Let $d = (c, \mathrm{per}(x))$. Then $d \neq 0$, $d \mid c$ and

$$\frac{\mathrm{per}(x)}{d}.c.x = \frac{c}{d}.\mathrm{per}(x).x = 0$$

which is

$$\mathrm{per}(c.x) | \frac{\mathrm{per}(x)}{d}$$

Now assume $f.c.x = 0$. We have $d = \alpha.c + \beta.\mathrm{per}(x)$ and therefore $d.f.x = \alpha.c.f.x + \beta.\mathrm{per}(x).f.x = 0$ which is $\mathrm{per}(x) | d \cdot f$. Now, d is a divisor of $\mathrm{per}(x)$ and therefore :

$$\frac{\mathrm{per}(x)}{d} | f$$

This is:

$$\frac{\mathrm{per}(x)}{d} = \mathrm{per}(c.x)$$

□

Corollary 3.14 *If* $(c, per(x)) = 1$ *then* $per(c.x) = per(x)$. □

Corollary 3.15 *If* $per(x) = p^r$ *then*

$$per(p^s.x) = \begin{cases} 1, & \text{if } s \geq r; \\ p^{r-s}, & \text{if } s < r. \end{cases}$$

PROOF. The first case is obvious. If $s < r$ then $(p^s, per(x)) = (p^s, p^r) = p^s$ and

$$\text{per}(p^s.x) = \frac{p^r}{p^s} = p^{r-s}.$$

□

Lemma 3.16 *Let* $\mathbf{M}$ *be a* $p-$*module with* $per(\mathbf{M}) = p^r, r \geq 1$. *Let* $x_1 \in \mathbf{M}$ *such that* $per(x_1) = per(\mathbf{M}) = p^r$. *Then every coset* $\overline{y}$ *of* $\mathbf{M}/ < x_1 >= \overline{\mathbf{M}}$ *contains an element* y *such that* $per(y) = per(\overline{y})$.

PROOF. We first remark that for every factor module and every $y \in \overline{y}$ one has that $per(\overline{y})|\text{per}(y)$. This is obvious: $c.y = 0$ implies that $c.\overline{y} = \overline{0}$.
Let now $\text{per}(\overline{y}) = p^n, n \leq r$. Then $p^n.\overline{y} =< x_1 >$ which is $p^n.y \in < x_1 >$ or $p^n.y = a.x_1 = (p^s \cdot c).x_1$ where we assume that $(p, c) = 1$. We need to consider two cases:

$s \geq r$: We have that $p^n.y = c \cdot p^s.x_1 = 0$ and therefore $\text{per}(y)|p^n = \text{per}(\overline{y})$ and by the pervious remark, $\text{per}(\overline{y})|\text{per}(y)$. Thus, $\text{per}(y) = \text{per}(\overline{y}) = p^n$.

$s < r$: We have $\text{per}(c.x_1) = \text{per}(x_1) = p^r$ by the previous lemma, $(c, p^r) = (1)$. Thus:

$$\text{per}(p^n)y = \text{per}((p^s \cdot c).x_1) = \frac{\text{per}(c.x_1)}{(p^s, \text{per}(x_1))} = \frac{p^r}{(p^s, p^r)} = \frac{p^r}{p^s} = p^{r-s}$$

We have $\text{per}(y) = p^k$ where $k \geq n$ and therefore by the second corollary of the last lemma, $\text{per}(p^n.y) = p^{k-n}$. Thus, $p^{r-s} = p^{k-n}$ which is $k = n + (r - s)$. Now, $k \leq r$ gives $n + (r - s) \leq r$, i.e.,$n \leq s$. The element

$$y' = y - (p^{s-n} \cdot c).x_1$$

is the element we are looking for. We have

$$y' \equiv y \mod < x_1 > \text{ i.e., } y' \in \overline{y}$$

and $p^n.y' = p^n.y - (p^s \cdot c).x_1 = 0$ gives $\text{per}(y')|p^n = \text{per}(\overline{y})$. We trivially have $\text{per}(\overline{y})|\text{per}(y')$ and therefore $\text{per}(y') = \text{per}(\overline{y})$ □

Definition 14 $y_1, \ldots, y_m$ are *independent* if and only if

$$< y_1, \ldots, y_m >=< y_1 > \oplus \ldots \oplus < y_m > \text{ iff } (a_1.y_1 + \ldots + a_m.y_m = 0 \Rightarrow a_1.y_1 = 0, \ldots, a_m.y_m = 0)$$

Lemma 3.17 *Let* $\mathbf{M}$ *be a* $p-$*module where* $per(\mathbf{M}) = p^r, r \geq 1$ *and let* $x_1 \in \mathbf{M}$ *such that* $per(x_1) = p^r$. *Then if* $\overline{y_1}, \ldots, \overline{y_m}$ *are independent in* $\mathbf{M}/ < x_1 >$ *there are* $y_1 \in \overline{y_1}$ *with* $per(y_i) = per(\overline{y_i})$ *such that for every* $z \in < x_i >$ *one has that* $z, y_1, \ldots, y_m$ *are independent in* $\mathbf{M}$.

PROOF. We choose $y_1, \ldots, y_m$ according to the last lemma and assume $a.z + a_1.y_1 + \ldots + a_m.y_m = 0$. Then $a.\overline{z} + \ldots + a_m.\overline{y_m} = 0$. We have $a.\overline{z} = 0$ because $z \in < x_1 >$. But then $a_1.\overline{y_1} + \ldots + a_m.\overline{y_m} = 0$ which is $a_.\overline{y_1} = 0, \ldots, a_m.\overline{y_m} = 0$ by assumption. This yields $\text{per}(\overline{y_i}) = \text{per}(y_i)|a_i$, therefore $a_i.y_i = 0$ and finally $a.z = 0$. □

Corollary 3.18 *Assume that* $\mathbf{M}/ < x_1 >=< \overline{y_1} > \oplus \ldots \oplus < \overline{y_m} >$. *Then one has for suitable* $y_i \in \overline{y_i}$ *that* $\mathbf{M} =< x_1 > \oplus < y_1 > \oplus \ldots < y_m >$. □

Let $\mathbf{M}$ be a module over $\mathbf{D}$ and let p be a prime in $\mathbf{D}$. Then

$$\mathbf{M}(p) = \{x \mid \text{per}(x)|p\}$$

is a vector space over the field $\mathbf{D}/(p)$ with respect to

$$d + (p).x = d.x$$

We need to show that the multiplication is well defined: if $d \equiv d' \mod (p)$ then $d = d' + a \cdot p$ and $d.x = d'.x + apx = d'.x$

Lemma 3.19 *Assume that the elements $y_i, i = 1, \ldots, y_m$ of $\mathbf{M}(p)$ are different from zero. Then one has that $y_1, \ldots, y_m$ are independent in the $\mathbf{D}-$module $\mathbf{M}(p)$ iff $y_1, \ldots, y_m$ are linearly independent in the vector space $\mathbf{M}(p)$ over $\mathbf{D}/(p)$.*

PROOF. We first note that $y_i \neq 0$ is the same as $\text{per}(y_i) = p$. Assume first that $y_1, \ldots, y_m$ are independent. Then if $[d_1].y_1 + \ldots + [d_m].y_m = 0$ one has that $d_1.y_1 + \ldots + d_m.y_m = 0 d_i.y_i = 0$ and therefore $\text{per}(y_i)|d_i$. This is $p|d_i$ or $[d_i] = [0]$ in the vector space $\mathbf{D}$ over $\mathbf{D}/(p)$.

For the converse, assume that $y_1, \ldots, y_m$ are linerarly independent in $\mathbf{M}(p)$ over $\mathbf{D}/(p)$. Then $d_1.y_1 + \ldots + d_m.y_m = 0$ yields $d_1 \equiv 0(p)$. This is $d_i.y_i = 0$ because $y_i \in \mathbf{M}(p)$. □

Lemma 3.20 *Let $\mathbf{M}$ be a finitely generated $p-$module where per$(\mathbf{M}) = p^r, r \geq 1$. Let $x_1 \in \mathbf{M}$ where per$(x_1) = p^r$. Then:*

$$\dim_{\mathbf{D}/(p)} \mathbf{M}(\mathbf{p}) > \dim_{\mathbf{D}/(p)}(\mathbf{M}/ < x_1 >)(p)$$

PROOF. $\mathbf{M}(p)$ is as a submodule of a finitely generated module finitely generated. A generating set of $\mathbf{M}$ over $\mathbf{D}$ is a generating set of $\mathbf{M}(p)$ over $\mathbf{D}/(p)$. Thus $\mathbf{M}(p)$ has a finite basis over $\mathbf{D}/(p)$. $\mathbf{M}/ < x_1 >$ is as a homomorphic image of $\mathbf{M}$ also finitely generated.

Let now $\overline{y_1}, \ldots, \overline{y_m}$ be a base of $\mathbf{M}/ < x_1 > (p)$ over $\mathbf{D}/(p)$. Then $\overline{y_1}, \ldots, \overline{y_m}$ are independent over $\mathbf{D}$. We pick $y_i \in \overline{y_i}$ where $\text{per}(y_i) = \text{per}(\overline{y_i})$ and some $z \in < x_1 >$ where $\text{per}(z) = p$, e.g., $z = p^{r-1}.x_1$ such that $x, y_1, \ldots, y_m$ are independent in $\mathbf{M}$ and different from zero. All x, y_i belong to $\mathbf{M}(p)$ and we get that $x, y_1, \ldots, y_m$ are linearly independent in $\mathbf{M}(p)$ over $\mathbf{D}/(p)$. This is:

$$\dim_{\mathbf{D}/(p)} \mathbf{M}(p) \geq \dim_{\mathbf{D}/(p)}(\mathbf{M}/ < x_1 >)(p) + 1 > \dim_{\mathbf{D}/(p)}(\mathbf{M}/ < x_1 >)(p)$$

□

Theorem 3.21 *Let $\mathbf{M}$ be a finitely generated $p-$module over the principal ideal domain $\mathbf{D}$. Then $\mathbf{M}$ is a direct sum of cyclic $p-$modules:*

$$\mathbf{M} = < x_1 > \oplus \ldots \oplus < x_m >$$

where

$$per(x_i) = p^{\nu_i}, \nu_1 \geq \nu_2 \ldots \nu_m \geq 1 \text{ and } per(\mathbf{M}) = p^{\nu_1}.$$

PROOF. We proceed by induction on $\dim_{\mathbf{D}/(p)} \mathbf{M}(p)$.
$\dim_{\mathbf{D}/(p)} \mathbf{M}(p) = 0$ is $\mathbf{M} = \mathbb{O}$ because otherwise $\mathbb{M}$ has an element of power p.
Let $\dim_{\mathbf{D}/(p)} \mathbf{M}(p) = n$ and pick $x_1 \in \mathbf{M}, \text{per}(x_1) = \text{per}(\mathbf{M}) = p^{\nu_1}$. Then $\dim_{\mathbf{D}/(p)}(\mathbf{M}/ < x_1 >)(p) < n$. By induction hypothesis we get:

$$\mathbf{M}/ < x_1 > = < \overline{x_2} > \oplus \ldots \oplus < \overline{x_m} >$$

where $\mathrm{per}(\overline{x_2}) = \mathrm{per}(\mathbf{M}/ < x_1 >) = p^{\nu_2}, \mathrm{per}(\overline{x_j}) = p^{\nu_j}, \nu_2 \geq \ldots \geq \nu_m \geq 1$. Clearly, $\nu_2 \leq \nu_1$. We already know that we can pick $x_j \in \overline{x_j}$ such that $\mathrm{per}(x_j) = \mathrm{per}(\overline{x_j})$ and

$$\mathbf{M} = < x_1 > \oplus \ldots < x_m >$$

□

We have $< x > \cong \mathbf{D}/(\mathrm{per}(x))$ and therefore one has for any finitely generated $p-$module $\mathbf{M}$:

$$\mathbf{M} \cong \mathbf{D}/(p^{\nu_1}) \oplus \ldots \oplus \mathbf{M}/(p^{\nu^m})$$

According to Theorem (3.12) every finitely generated torsion module is the direct sum of its $p-$modules $\mathbf{T}_{p^i}$ and therefore one has

Theorem 3.22 *Every finitely generated torsion module over the principal ideal domain* $\mathbf{D}$ *is isomorphic to a direct sum of primary cyclic modules:*

$$\mathbf{M} \cong (\mathbf{D}/(p_1^{\nu_{11}}) \oplus \ldots \oplus \mathbf{D}/(p_1^{\nu_{1m_1}})) \oplus \ldots \oplus (\mathbf{D}/(p_k^{\nu_{k1}}) \oplus \ldots \oplus \mathbf{D}/(p_k^{\nu_{km_k}}))$$

The $p_i^{\nu_{ij}}$ *are called the elementary divisors of* $\mathbf{M}$*. Note that*

$$per(\mathbf{M}) = p_1^{\nu_{11}} \cdot \ldots \cdot p_k^{\nu_{k1}}$$

□

We can also group the cyclic modules $\mathbf{D}/p_i^{\nu_{ij}}$ according to the invariant factors. Let

$$m = \max_{i=1}^{k} m_i \text{ and put } \nu_{ij} = 0 \text{ if } m_i < j \leq m$$

and define

$$\begin{aligned}
\mathbf{C_1} &= \mathbf{D}/(p_1^{\nu_{11}}) \oplus \mathbf{D}/(p_2^{\nu_{21}}) \oplus \ldots \oplus \mathbf{D}/(p_k^{\nu_{k1}}), & q_1 &= p_1^{\nu_{11}} \cdot p_2^{\nu_{21}} \cdot \ldots p_k^{\nu_{k1}} \\
\mathbf{C_2} &= \mathbf{D}/(p_1^{\nu_{12}}) \oplus \mathbf{D}/(p_2^{\nu_{22}}) \oplus \ldots \oplus \mathbf{D}/(p_k^{\nu_{k2}}), & q_2 &= p_1^{\nu_{12}} \cdot p_2^{\nu_{22}} \cdot \ldots \cdot p_k^{\nu_{k2}} \\
&\ldots & & \\
\mathbf{C_m} &= \mathbf{D}/(p_1^{\nu_{1m}}) \oplus \mathbf{D}/(p_2^{\nu_{2m}}) \oplus \ldots \oplus \mathbf{D}/(p_k^{\nu_{km}}), & q_m &= p_1^{\nu_{1m}} \cdot p_2^{\nu_{21}} \cdot \ldots \cdot p_k^{\nu_{km}}
\end{aligned}$$

We clearly have that

$$q_m | q_{m-1} | \ldots | q_2 | q_1, q_m \neq 1$$

and the q_i are called the *invariant factors*. We are going to show that the submodules $\mathbf{C}_i$ are actually cyclic. This is a consequence of the

Theorem 3.23 *(*Chinese Remainder Theorem*) Let* $d_1, d_2, \ldots, d_m$ *be pairwise relatively prime elements of the principal ideal domain* $\mathbf{D}$*. Then the system of linear congruences:*

$$\begin{aligned}
x &\equiv x_1 \mod d_1 \\
x &\equiv x_2 \mod d_2 \\
&\ldots \\
x &\equiv x_m \mod d_m
\end{aligned}$$

has modulo $d = d_1 \cdot d_2 \cdot \ldots \cdot d_m$ *a unique solution* x*.*

PROOF. Assume that $x$and x' are congruent module $d_1, \ldots, d_m$. This is $d_1|x - x', \ldots, d_m|x - x'$. Because the d_i are pairwise relatively prime we can conclude that $d|x - x'$, i.e., $x \equiv x' \mod d$.

In order to show existence of such an x, we define

$$d_1^* = d_2 \cdot \ldots \cdot d_m$$

Then d_1 and d_1^* are relatively prime. Hence, we have $a_1 \cdot d_1^* + b_1 \cdot d_1 = 1$. This is $a_1 \cdot d_1^* \equiv 1 \mod d_1$. Of course, d_1^* is zero module $d_2, \ldots, d_m$. We similarly define d_2^* as the product of the d_i with d_2 missing. We then get an a_2 such that $a_2 \cdot d_2^* \equiv 1 \mod d_2$, etc. Then:

$$x = x_1 \cdot a_1 \cdot d_1^* + x_2 \cdot a_2 \cdot d_2^* + \ldots + x_m \cdot a_m \cdot d_m^*$$

is obviously a solution. □

For elements $d_1, \ldots, d_m$ of the principal ideal domain consider the homomorhism

$$\mathbf{D} \to \mathbf{D}/(d_1) \times \ldots \times \mathbf{D}/(d_m), x \mapsto ([x]_1, \ldots, [x]_m)$$

The kernel of this homomorphism is (l.c.m.$(d_1, \ldots, d_m)$ and, according to the Chinese Remainder Theorem, the map is surjective in case the d_i are pairwise relatively prime. In this formulation the Chinese Remainder Theorem invites generalization to more general rings and even to universal algebra. We also see now that $\mathbf{C_i} \cong \mathbf{D}/(q_i)$ and get

Theorem 3.24 *Every finitely generated torsion module over the principal ideal domain* $\mathbf{D}$ *is isomorphic to a direct sum of cyclic modules*

$$\mathbf{M} \cong \mathbf{D}/(q_1) \oplus \ldots \mathbf{D}/(q_k)$$

where the q_i *are the invariant factors.*

3.5 Uniqueness of the Invariant Factors and Elementary Divisors

We are going to show that the invariant factors determine a finitely generated $\mathbf{D}-$module $\mathbf{M}$ up to isomorphism uniquely. The spaces $\mathbf{M}(p)$ over $\mathbf{D}/(p)$ are again a basic tool.

Lemma 3.25 *Let* $q \neq 0$. *Then for any prime* $p \in \mathbf{D}$ *one has:*

$$\dim_{\mathbf{D}/(p)} \mathbf{D}/(q)(p) = \begin{cases} 1, & \text{if } p|q \\ 0, & \text{otherwise.} \end{cases}$$

PROOF. Assume first that p does not divide q. Then one has that $p.(d + (q)) = (q)$ iff $p \cdot d \in (q)$ iff $pd = cq$ iff $d = (c/p) \cdot q$. This is $d \in (q)$ which is $d + (q) = (q)$.

Now assume that $p|q$, that is $q = p \cdot b$ where $b \neq 0$. We have that $b|q$ but if we also had $q|b$ then $b \sim q$ and it would follow that p is a unit. Thus:

$$b + (q) \neq (q),\ p.(b + (q)) = (q)$$

This is $b + (q) \in (\mathbf{D}/(q))(p)$ and $b + (q) \neq (q)$. Now let $d + (q) \in (\mathbf{D}/(q))(p)$. Then $pd = cq = cpb$ and therefore $d = cb$ which is $d + (q) = c.(b + (q))$. Thus, $b + (q)$ is a generator $\neq \mathbb{O}$ of $(\mathbf{D}/(q))(p)$ over $\mathbf{D}$ and therefore $\dim_{\mathbf{D}/(p)}(\mathbf{D}/(q))(p) = 1$. □

Lemma 3.26 *Let $q = ab \neq 0$. Then $a.\mathbf{D}/(q) \cong \mathbf{D}/(b)$ as $\mathbf{D}-$modules.*

PROOF. $a.(d + (q)) \mapsto d + (b)$ is the desired isomorphism. The map is well defined: $ad_1 \equiv ad_2(q)$ is $a(d_1 - d_2) = cq$ and therefore $d_1 - d_2 = cb$, i.e., $d_1 \equiv d_2(b)$. Linearity and surjectivity are obvious. Assume $d \in (b)$, i.e., $d = cb$. Then $ad = cq \in (q)$ is injectivity. □

Theorem 3.27 *Let $\mathbf{M}$ be a finitely generated torsion module over the principal ideal domain $\mathbf{D}$ and let*

$$\mathbf{M} = \mathbf{C_1} \oplus \ldots \oplus \mathbf{C_k} = \mathbf{C'_1} \oplus \ldots \oplus \mathbf{C'_l}$$

be two decompositions of $\mathbf{M}$ as direct sums of non-zero cyclic modules $\mathbf{C_i}$ and $\mathbf{C'_j}$, respectively, with periods q_i and q'_j, and forming divisor chains:

(i) $\mathbf{C}_i \cong \mathbf{D}/(q_i)$, *and* $\mathbf{C'_j} \cong \mathbf{D}/(q'_j)$

(ii) $q_k| \ldots |q_1$ *and* $q'_l| \ldots |q_1$

Then $k = l$ and $q_i \sim q'_i$ for $i = 1 \ldots k$.

PROOF. Let $p = p_1$ be any prime that divides q_k. Then p divides $q_{k-1}, \ldots, q_1$. Assume that $x \in \mathbf{M}(p)$. We have:

$$x = x_1 + \ldots + x_k, p.x = 0 = p.x_1 + \ldots + p.x_k$$

and therefore

$$p.x_1 = \ldots = p.x_k = 0$$

This is

$$\mathbf{M}(p) = \mathbf{C_1}(p) \oplus \ldots \oplus C_k(p)$$

and therefore

$$\dim_{\mathbf{D}(p)}(p) = k$$

Hence there must be exactly $k-$many q'_j for which $p_1|q'_j$. This is $k \leq l$ and by symmetry we conclude $k = l$. We also see that $p_1|q'_k$. Hence: $q_k = p_1 b_k,\ q'_k = p_1 b'_k$ and $p_1|q_i,\ p_1|q'_i$, say $q_i = p_1 b_i,\ q'_i = p_1 b'_i,\ i = 1, \ldots, k$.
By the last lemma,

$$p_1\mathbf{M} \cong \mathbf{D}/(b_1) \oplus \ldots \oplus \mathbf{D}/(b_k) \cong \mathbf{D}/(b'_1) \oplus \ldots \oplus \mathbf{D}/(b'_k) \text{ where } b_k| \ldots |b_1 \text{ and } b'_k| \ldots |b'_1$$

Now, a prime p_2 which divides b_k divides b'_k, and vice versa, and we get:

$$p_1p_2\mathbf{M} \cong \mathbf{D}/(c_1) \oplus \ldots \mathbf{D}/(c_k) \cong \mathbf{D}/(c'_1) \oplus \ldots \mathbf{D}/(c'_k) \text{ where } c_k| \ldots |c_1 \text{ and } c'_k| \ldots |c'_1$$

and where we have:

$$p_1p_2c_i = q_i \text{ and } p_1p_2c'_i = q'_i$$

If $q_k = p_1 \cdot \ldots \cdot p_s$ then

$$q_k\mathbf{M} \cong \mathbf{D}/(e_1) \oplus \ldots \oplus \mathbf{D}/(e_{k-1}) \cong \mathbf{D}/(e'_1) \oplus \ldots \oplus \mathbf{D}/(e'_k) \text{ where } e_iq_k = q_i,\ e'_iq_k = q'_i$$

and

$$1 \sim e_k|e_{k-1}| \ldots |e_1,\ e'_k|e'_{k-1}| \ldots |e'_1$$

On the left-hand side for the sum decomposition of $q_k\mathbf{M}$ there are at most $k-1$ summands different from zero, so also on the right-hand side. Thus:

$$e'_k \sim 1 \text{ which is } q_k \sim q'_k.$$

If we assume the theorem for $\mathbf{D}-$modules with decompositions into $< k$ cyclic modules then we get: $e_i \sim e_i'$ and $q_i \sim q_i'$ for $i = 1, \ldots, k-1$. The case $k = 1$ is trivial: $\mathbf{D}/(q) \cong \mathbf{D}/(q')$ as $\mathbf{D}-$modules yields $\mathrm{per}(\mathbf{D}/(q)) = q = \mathrm{per}(\mathbf{D}/(q')) = q'$. □

The invariance of the elementary divisors is now an easy consequence: We have that $\mathbf{M}$ is a direct sum of its maximal primary submodules, and each of these primary modules is a direct sum of cyclic primary modules subject to the divisor chain condition.

3.6 Finitely Generated Abelian Groups

Let $\mathbf{A}$ be a finitely generated abelian group. Then:

$$\mathbf{A} \cong \mathbb{Z}^r \oplus \mathrm{Tor}(\mathbb{A}) \text{ where } r = \mathrm{rank}(\mathbb{A})$$

and

$$\mathrm{Tor}(\mathbf{A}) \cong \mathbb{Z}_{m_1} \oplus \ldots \oplus \mathbb{Z}_{m_l} \text{ where } m_l | m_{l-1} | \ldots | m_1,\ m_l > 1$$

The isomorphism type of $\mathbf{A}$ is uniquely determined by r and the invariant factors m_i. Notice:

$$\mathrm{ord}(\mathrm{Tor}(\mathbf{A}) = m_1 \cdot \ldots \cdot m_l,\ \mathrm{per}(\mathbf{A}) = m_1$$

We also have a unique decomposition according to the elementary divisors:

$$\mathrm{Tor}(\mathbf{A}) \cong (\mathbb{Z}_{p_1^{\nu_{11}}} \oplus \ldots \oplus \mathbb{Z}_{p_1^{\nu_{1l_1}}}) \oplus \ldots \oplus (\mathbb{Z}_{p_k^{\nu_{k1}}} \oplus \ldots \oplus \mathbb{Z}_{p_k^{\nu_{kl_k}}})$$

The product of all elementary divisors $p_i^{\nu_{ij}}$ is the order of the torsion and its period is the product of the $p_i^{\nu_{i1}}$ for $i = 1, \ldots, k$.

Theorem 3.28 ***(Kronecker)*** *A finite abelian group $\mathbf{G}$ of order m is cyclic if and only if there are for every divisor d of m at most d elements x for which $x^d = e$.*

Proof. While we have formulated the theorem for multiplicative notation we use for the proof additive notation, e.g., we consider $\mathbf{G}$ as a module over $\mathbb{Z}$.
Let $\mathbf{G}$ be cyclic, i.e., $\mathbf{G} \cong \mathbb{Z}_m$. Let d be a divisor of m. Then $\{x \mid d.x = 0\}$ is a subgroup $\mathbf{H}$ of $\mathbf{G}$.Now, $\mathbf{H}$ is generated by the class $[\frac{m}{d}]_m$ and this group is isomorphic to $\mathbb{Z}_d$.
On the other hand, if $\mathbb{G}$ is not cyclic then one of the primary components, say the one for the prime p_1 is not cyclic, i.e., $T_{p_1} \cong \mathbb{Z}_{p_1^{p_{11}}} \oplus \mathbb{Z}_{p_1^{p_{11}}} \oplus \ldots$ and we have in the first component at least p_1-many elements of order p_1 but also in the second component. That is, $\mathbf{G}$ contains for the divisor p_1 of the order m at least $2p_1-$many elements of order p_1. □

Corollary 3.29 ***(Gauss)*** *The multiplicative group of a finite field is cyclic.*

Proof. This is now obvious: $x^d = 1$ is the same as $x^d - 1 = 0$ and there are at most d solutions in $\mathbf{F}$. □

Recall that if $\mathbf{F}$ is a finite field of order m then $m = p^n$ where p is the characteristic of $\mathbf{F}$.

3.7 The Structure of a Linear map over a Finite Dimensional Vector Space

Let $\mathbf{F}$ be a field and let $\mathbf{F}[x]$ be the ring of polynomials over $\mathbf{F}$. Let $\mathbf{A}$ be an $\mathbf{F}[x]-$module. Then

$$\mathbf{V_A} = (A, +, -, 0, (c.)_{c\in\mathbf{F}})$$

is a vector space over $\mathbf{F}$ and

$$\tau_x : \mathbf{V_A} \to \mathbf{V_A}, v \mapsto x.v$$

is a linear transformation on $\mathbf{V_A}$. Thus, we have a correspondence:

$$\mathbf{A} \mapsto (\mathbf{V_A}, \tau)$$

from the class of $\mathbf{F}[x]-$modules to pairs consisting of a vector space $\mathbf{V}$ over $\mathbf{F}$ and linear maps τ on $\mathbf{V}$.

On the other hand, let $(\mathbf{V}, \tau)$ be such a pair. For a polynomial $f(x) \in \mathbf{F}[x]$ define:

$$f(x).v = f(\tau)(v)$$

This makes $\mathbf{V}$ to an $\mathbf{F}[x]-$module $\mathbf{A}$ and it is easy to see that the correspondence $\mathbf{A} \mapsto (\mathbf{V_A}, \tau)$ is bijective.

Let $\phi : \mathbf{V_A} \to \mathbf{V_{A'}}$ be linear. Assume that ϕ is also $\mathbf{F}[x]-$linear. Then one has that $\phi(x.v) = x.\phi(v)$ which is $\phi(\tau(v)) = \tau'(\phi(v))$ or

$$\phi \circ \tau = \tau' \circ \phi$$

In particular, if ϕ is an isomorphism between $\mathbf{F}[x]-$modules then

$$\phi \circ \tau \circ \phi^{-1} = \tau'$$

shows that τ *and* τ' *are similar.*

Let $\mathbf{W}$ be a subspace of $\mathbf{V_A}$. Then $\mathbf{W}$ is an $\mathbf{F}[x]-$submodule of $\mathbf{A}$ if and only if $x.v \in \mathbf{W}$ in case that $v \in \mathbf{W}$, and this is that $\tau(v) \in \mathbf{W}$ for any $v \in \mathbf{W}$, or that $\mathbf{W}$ is $\tau-$*invariant.*

Let $\mathbf{C}$ be a cyclic submodule of the $\mathbf{F}[x]-$module $\mathbf{A}$:

$$\mathbf{C} =< v >_{\mathbf{F}[x]} = \{f(\tau).v \mid f \in \mathbf{F}[x]\} =< \{\tau^\nu.v \mid \nu \in \mathbb{N}\} >_{\mathbb{F}}$$

and one has:

$$\mathbf{F}[x]/(\mathrm{per}(v)) \cong \mathbf{C} \text{ as } \mathbf{F}[x]-\text{modules}$$

where $\mathrm{per}(v) = m_v(x)$ is the polynomial of smallest degree such that $m_v.v = 0$ or $m_v = 0$ in case that $\mathbf{F}[x] \cong \mathbf{C}$. We assume from now on that m_v has a degree, i.e., that it is different from zero, and let $\deg(m_v) = n$. Of course, $v = 0$ is equivalent to $n = \deg(m_v) = 0$

$m_v(x)$ is called the *minimal polynomial* of v. Of course, we may always assume that m_v is monic:

$$m_v = a_0 + a_1 x + \ldots + a_{n-1} + x^n, n = \deg(m_v).$$

Now, $\mathbf{F}[x]/(m_v)$ is also a vector space over $\mathbf{F}$ and as such has a base $[1], [x], \ldots, [x^{n-1}]$ where $[\,]$ denotes the congruence class modulo $(m_v(x))$ in the polynomial ring $\mathbf{F}[x]$. For $\mathbf{C}$ this means that

$$v, \tau(v), \ldots, \tau^{n-1}(v)$$

is a base of $< v >_{\mathbf{F}[x]}$ over $\mathbf{F}$. The map τ, restricted to the cyclic space $\mathbf{C}$ has with respect to the base $\{v, \tau(v), \ldots, \tau^{n-1}(v)\}$ the matrix:

$$\mathbf{T}_v = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{pmatrix}$$

and $\tau^n(v) = -a_0.v - a_1.\tau(v) - \ldots a_{n-1}.\tau^{n-1}(v)$

Theorem 3.30 *Let $\mathbf{V}$ be a finite dimensional vector space over the field $\mathbf{F}$ and let $\tau : \mathbf{V} \to \mathbf{V}$ be a linear transformation on $\mathbf{V}$. Then $\mathbf{V}$ is a direct sum of cyclic invariant subspaces.*
Let m_1 be the minimal polynomial of τ. Then there is a list

$$m_l(x)|m_{l-1}(x)|\ldots|m_1(x), \deg(m_l) > 0$$

of monic polynomials, each one dividing the next, such that

$$(\mathbf{V}, \tau) \cong \mathbf{F}[x]/(m_1(x)) \oplus \ldots \oplus \mathbf{F}[x]/(m_l(x))$$

The list $m_i(x)$ of invariant factors determines the similarity class of τ uniquely. One has

$$\dim(\mathbf{V}) = \deg(m_1) + \ldots + \deg(m_l) \text{ and } m(x) = m_1(x) = \text{ minimal polynomial of } \tau$$

One can find in $\mathbf{V}$ a base such that the matrix $\mathbf{T}$ of τ is in rational canonical form:

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}_1 & \ldots & \ldots & \ldots \\ \ldots & \mathbf{T}_2 \ldots & \ldots & \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \mathbf{T}_l \end{pmatrix}$$

The matrices $\mathbf{T}_\nu = \mathbf{T}_{v_\nu}$ are as before and all entries outside these blocks are zero.

For $\tau : \mathbf{V} \to \mathbf{V}$ and matrix $\mathbf{T}$ for τ one defines

$$\det(\mathbf{T} - x_1.\mathbf{E}) = c(x) \in \mathbf{F}[x]$$

as the *characteristic* polynomual for τ. Here

$$\mathbf{E} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ & & \ddots & \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

is the unit $n \times n-$matrix. The polynomial $c(x)$ does not depend on the chosen base for $\mathbf{V}$. If $\mathbf{T}'$ is the matrix for τ in any other base then one has some non-singular matrix $\mathbf{S}$, namely the matrix which encodes the coordinate transformation between the two bases, such that $\mathbf{T}' = \mathbf{S} \circ \mathbf{T} \circ \mathbf{S}^{-1}$ and $\det(\mathbf{T}' - x.\mathbf{E}) = \det(\mathbf{S}\mathbf{T}\mathbf{S}^{-1} - x.\mathbf{E}) = \det(\mathbf{S}(\mathbf{T} - x.\mathbf{E})\mathbf{S}^{-1}) = \det(\mathbf{T} - x.\mathbf{E})$

If $\mathbf{V} = \mathbf{V}_1 \oplus \ldots \oplus \mathbf{V}_k$ is a direct sum of $\tau-$invariant subspaces then $c(x) = c_1(x) \cdot \ldots \cdot c_k(x)$ where $c_i(x)$ is the characteristic polynomial of τ restricted to $\mathbf{V}_i$. For invariant subspaces $\mathbf{C}$ that are cyclic, i.e., $\mathbf{C} =< v >_{\mathbf{F}[x]} =< \{v, \tau(v), \ldots, \tau^{n-1}(v)\} >_{\mathbf{F}}$, τ has for $\mathbf{C}$ a matrix like $\mathbf{T}_v$, i.e., a matrix that has units on the lower diagonal and the the last column is given by coefficients $-a_i$ of the minimal polynomial $m_v(x)$ of v. It is a rather easy exercise to prove that the characteristic polynomial of such a matrix is $\pm m_v(x)$.

Theorem 3.31 ***(Hamilton-Cayley)*** *Let $c = c(x)$ be the characteristic polynomial for $\tau : \mathbf{V} \to \mathbf{V}$. Then*

$$c = \pm m_1 \cdot \ldots \cdot m_l$$

where $m_l | \cdots | m_1$ is the list of invariant factors for τ.

Corollary 3.32 *The minimal polynomial $m = m_1$ for τ divides the characteristic polynomial, i.e., $c(\tau) = 0$.*

Corollary 3.33 *The characteristic and the minimal polynomial for τ have the same prime factors $p_i(c)$.*

Theorem 3.34 *Let $\tau : \mathbf{V} \to \mathbf{V}$ be a linear transformation on the finite dimensional vector space. Let $m = p_1(x)^{\nu_1} \cdot \ldots \cdot p_k(x)^{\nu_k}$ be the prime factorization of the minimal polynomial $m_\tau(x)$ for τ. Then:*

$$(\mathbf{V}, \tau) = T_{p_1}(\tau) \oplus \ldots \oplus T_{p_k}(\tau) \text{ where}$$

$$T_{p_i}(\tau) = \{ v \mid m_v(x) | p_i(x)^r \text{ for some } r \geq 1 \} = \{ v \mid p_i(\tau)^r(v) = 0 \text{ for some } r \geq 1 \}$$

and each $T_{p_1}(\tau)$ is a direct sum of cyclic $p_i(x)-$modules:

$$T_{p_1}(\tau) \cong \mathbf{F}[x]/(p_i(x)^{\nu_{i1}} \oplus \ldots \oplus \mathbf{F}[x]/(p_i(x)^{\nu_{il_i}}, \nu_{i1} \geq \ldots \geq \nu_{il_i} \geq 1$$

where the list of elementary divisors $p_i(x)^{\nu_{ij}}$ determines the similarity type of τ uniquely.

Of particular interest is the case where $\mathbf{F}$ is *algebraically closed*. Then $p(x)$ is prime if and only if $p(x) = x - \lambda$ for some $\lambda \in \mathbf{F}$. An example is $\mathbf{F} = \mathbb{C}$ where $\mathbb{C}$ is the field of complex numbers.
Let $\mathbf{C}$ be a cyclic invariant subspace, i.e.,

$$\mathbf{C} \cong \mathbf{F}[x]/(x - \lambda)^l$$

If $\mathbf{C}$ has v as a cyclic generator then $v, \tau(v), \ldots \tau^{l-1}(v)$ is a base of $\mathbf{C}$. We claim that also:

$$v_0 = v, v_1 = (\tau - \lambda)(v), v_2 = (\tau - \lambda)^2(v), \ldots, v_{l-1} = (\tau - \lambda)^{l-1}(v)$$

is a base. Clearly, $\mathbf{C}$ is invariant under $(\tau - \lambda)$ and therefore all v_i belong to $\mathbf{C}$. Hence, we only have to show that they are linearly independent. Assume that:

$$\rho_0.v + \rho_1.(\tau - \lambda)(v) + \ldots + \rho_{l-1}.(\tau - \lambda)^{l-1}(v) = 0$$

If we multiply this relation by $(\tau - \lambda)^{l-1}$ then we get $\rho_0(\tau - \lambda)^{l-1}(v) = 0$ and this yields $\rho_0 = 0$ because the degree of the minimal polynomial $m_v = (x - \lambda)^l$ is l. We get $\rho_1 = \ldots = \rho_{l-1} = 0$ similarly. Now,

$$\begin{array}{rclcrcl}
\tau(v_0) & = & v_1 + \lambda.v_0 & \text{and} & v_2 & = & (\tau - \lambda)(\tau - \lambda)v = (\tau - \lambda)v_1 \\
\tau(v_1) & = & v_2 + \lambda.v_1 & \text{and} & v_3 & = & (\tau - \lambda)(\tau - \lambda)^2 v = (\tau - \lambda)v_2 \\
\tau(v_2) & = & v_3 + \lambda.v_2 & \text{and} & v_4 & = & \ldots \\
 & & & \ldots & & & \\
\tau(v_{l-2}) & = & v_{l-1} + \lambda.v_{l-2} & \text{and} \quad (\tau - \lambda)v_{l-1} & = & & (\tau - \lambda)^l v = 0 \\
\tau(v_{l-1}) & = & \lambda.v_{l-1} & &
\end{array}$$

With respect to this basis the matrix of the map τ on $\mathbf{C}$ looks like:

$$\mathbf{E}_{\lambda l} = \begin{pmatrix} \lambda & 0 & \ldots & 0 \\ 1 & \lambda & 0 \ldots & 0 \\ & \ddots & & \vdots \\ 0 & 0 & \ldots & \lambda \end{pmatrix}$$

The *elementary Jordan matrix* $\mathbf{E}_{\lambda \mathbf{l}}$ has $l-$many $\lambda' s$ on the main diagonal and $l - 1$ units on the lower diagonal. It corresponds to the polynomial $(x - \lambda)^l$.

Corollary 3.35 ***(Jordan Normal Form)*** *If the minimal polynomial $m(x)$ of τ splits in $\mathbf{F}[x]$ into linear factors $p_i(x) = (x - \lambda_i), i = 1, \ldots, k$, i.e.,*

$$m(x) = (x - \lambda_1)^{\nu_1} \cdot \ldots \cdot (x - \lambda_k)^{\nu_k}$$

Then one can find in $\mathbf{V}$ a base such that the matrix $\mathbf{T}$ of τ is in Jordan Normal form:

$$\mathbf{T} = \begin{pmatrix} \mathbf{E}_{\lambda_1 \nu_{11}} & & & & \\ & \ddots & & & \\ & & \mathbf{E}_{\lambda_1 \nu_{1 l_1}} & & \\ & & & \ddots & \\ & & & & \mathbf{E}_{\lambda_k \nu_{k l_k}} \end{pmatrix}, \nu_i = \nu_{i1} \geq \ldots \geq \nu_{i l_1}$$

The primary component for $(x - \lambda_i)$ is the space

$$\mathbf{T}_{(x-\lambda_i)} = \{v \mid (\tau - \lambda_i)^r = 0 \text{ for some } r \geq 0\} = \{v \mid (\tau - \lambda_i)^{n_i} = 0\}, n_i = \nu_{i1} + \ldots + \nu_{i l_i} = \dim(T_{(x-\lambda_i)})$$

and its minimal and characteristic polynomial for τ are

$$m_i(x) = (x - \lambda)^{\nu_i}, c_i(x) = (x - \lambda)^{n_i}$$

□

For $\lambda \in \mathbf{F}$ we define

$$\mathbf{E}_\lambda = \{v \mid (\tau - \lambda)(v) = 0\} = \ker(\tau - \lambda)$$

as the *eigenspace* for λ. Clearly, $\mathbf{E}_\lambda \subseteq \mathbf{T}_{(x-\lambda)}$. If $\mathbf{E}_\lambda \neq \mathbb{O}$, then λ is called an *eigenvalue* for τ and $\mathbf{T}_{(x-\lambda)}$ is called a *generalized eigenspace*. We have that $\lambda = 0$ is an eigenvalue iff τ is singular.

We have $\mathbf{T}_{(x-\lambda_i)} = \mathbf{E}_{\lambda_i}$ iff $\{v \mid (\tau - \lambda_i)^r(v) = 0 \text{ for some } r \geq 1\} = \{v \mid (\tau - \lambda_i)(v) = 0\}$ iff $\nu_{i1} = 1$ iff $(x - \lambda_i)$ is the minimal polynomial of τ restricted to $\mathbf{T}_{(x-\lambda_i)}$ iff $\mathbf{T}_{(x-\lambda_i)}$ has an eigenbase iff the matrix $\mathbf{T}_i$ of τ restricted to $\mathbf{T}_{(x-\lambda_i)}$ is an $n_i \times n_i$ diagonal matrix with λ_i on the diagonal.

Corollary 3.36 *Assume that the minimal polynomial of τ splits into linear factors $(x - \lambda_i)$. Then the following statements are equivalent:*

(i) $\mathbf{V}$ *has a base of eigenvectors for τ.*

(ii) *τ admits a representation by a diagonal matrix.*

(iii) *The minimal polynomial of τ is without multiple roots, i.e.,*

$$m(x) = (x - \lambda_1) \cdot \ldots \cdot (x - \lambda_k).$$

(iv) *The characteristic polynomial of τ is*

$$c(x) = (x - \lambda_1)^{n_1} \cdot \cdots \cdot \ldots \cdot (x - \lambda_k)^{n_k}, \text{ where } n_i = \dim(\mathbf{E}_i).$$

□

Chapter 4

Sylow Theorems

4.1 Transformation Groups

Let S be any set. The group of bijective mappings on S is sometimes called the *symmetric* group $\mathrm{Sym}(S)$ on S. Any subgroup $\mathbf{G}$ of $\mathrm{Sym}(S)$ is called a *transformation group* on the set S and we say that $\mathbf{G}$ *acts on* S.
Examples

1. Let $S = n = \{0, \ldots, n-1\}$. $\mathrm{Sym}(n)$ is called the symmetry group of degree n.
2. Let $\mathbf{G}$ be any group. We already established the injective group homomorphism:

   $$\lambda : \mathbf{G} \to \mathrm{Sym}(\mathbf{G}), x \mapsto \lambda_x, \lambda_x(y) = x \cdot y$$

   Every group $\mathbf{G}$ acts on its carrier set by means of left-translations and $\mathbf{G} \cong \lambda(\mathbf{G}) \subseteq \mathrm{Sym}(\mathbf{G})$. This, of course, is Cayley's theorem. The group $\mathbf{G}$ acts on itself by left translations.
3. Let $\mathbf{G}$ be any group. We already established the group homomorphism

   $$\sigma : \mathbf{G} \ \to \mathrm{Sym}(\mathbf{G}), x \mapsto \sigma_x, \sigma_x(y) = x \cdot y \cdot x^{-1}$$

   The kernel of the homomorphism σ is $\mathrm{cent}(\mathbf{G})$. Each σ_x is an (inner) automorphism of $\mathbf{G}$ and $\mathbf{G}$ acts on itself by conjugation. We have that $\mathbf{G}/\mathrm{cent}(\mathbf{G}) \cong \mathrm{Aut}_i(\mathbf{G})$.

Definition 15 A *representation* of the group $\mathbf{G}$ by transformations on a set S is a homomorphism

$$\rho : \mathbf{G} \to \mathrm{Sym}(S)$$

Definition 16 For a representation $\rho : \mathbf{G} \to \mathrm{Sym}(S)$ the *equivalence under* ρ is an equivalence on S and defined by:

$$a \sim_\rho b \text{ iff } a = \rho(x)(b) \text{ for some } x \in \mathbf{G}$$

The equivalence classes for $\sim_\rho$ are called the *orbits* of ρ. We obviously have that the orbit of a under ρ is the set

$$\mathrm{orb}_\rho(a) = \{\rho(x).a \mid x \in \mathbf{G}\}$$

and it is also called the *trace* of a under ρ. For $a \in S$ the set

$$\mathrm{stab}_\rho = \{x \mid \rho(x)(a) = a\}$$

is called the *stabilizer of* a *under* ρ. It is a subgroup of $\mathbf{G}$.

Examples

1. We get for the representation λ of a subgroup $\mathbf{H}$ of $\mathbf{G}$ that

$$a \sim_\lambda b \text{ iff } a = h \cdot b \text{ for some } h \in \mathbf{H} \text{ iff } a \in \mathbf{H}b$$

   The orbits are the right-cosets of $\mathbf{H}$ in $\mathbf{G}$. The stabilizer is the trivial group $\{e\}$ for every $a \in G$.

2. For conjugation we get that

$$a \sim_\sigma b \text{ iff } a = x \cdot b \cdot x^{-1}$$

   for some $x \in \mathbf{G}$ and the orbits are the conjugacy classes $\{b \mid b = xax^{-1}, x \in \mathbf{G}\}, a \in \mathbf{G}$, of $\mathbf{G}$. The conjugacy class of a is the singleton $\{a\}$ iff $a \in \text{cent}(\mathbf{G})$. We have

$$\text{stab}_\sigma(a) = \{x \in \mathbf{G} \mid \sigma_x(a) = a\} = \{x \in \mathbf{G} \mid x \cdot a \cdot x^{-1} = a\} = \{x \in \mathbf{G} \mid xa = ax\} = \text{cent}_{\mathbf{G}}(a)$$

   Where $\text{cent}_{\mathbf{G}}(a)$ is called the *centralizer of a in $\mathbf{G}$*. It is the set of all $x \in \mathbf{G}$ which commute with a. The conjugacy class of a is the singleton $\{a\}$ iff $a \in \text{cent}(\mathbf{G})$ iff $\text{cent}_{\mathbf{G}}(a) = \mathbf{G}$. This is a special case of the next theorem.

Theorem 4.1 *Let $\mathbf{G}/\text{stab}_\rho(a)$ be the set of left-cosets of $\text{stab}_\rho(a)$ in $\mathbf{G}$. Then one has that:*

$$\mathbf{G}/\text{stab}_\rho(a) \to \text{orb}_\rho(a), x \cdot \text{stab}_\rho(a) \mapsto \rho(x)(a)$$

is a well defined bijective map.

PROOF. We have that

$$x_1 \,\text{stab}_\rho(a) = x_2 \,\text{stab}_\rho(a) \text{ iff } x_2^{-1}x_1 \in \text{stab}_\rho(a) \text{ iff } \text{ iff } \rho(x_2)^{-1}\rho(x_1)(a) = a \text{ iff } \rho(x_1)(a) = \rho(x_2)(a).$$

This proves that the map is well defined and injective. For surjectivity, let $b \in \text{orb}_\rho(a)$. This is $b = \rho(x)(a)$ for some $x \in \mathbf{G}$. But then $x\,\text{stab}_\rho(a) \mapsto \rho(x)(a) = b$. □

Corollary 4.2 (Counting Lemma) *One has for every $a \in S$ that* $\text{card}(\text{orb}_\rho(a)) = \text{card}(\mathbf{G}/\text{stab}_\rho(a))$.
□

Recall that for any subgroup $\mathbf{H}$ of $\mathbf{G}$, $\text{card}(\mathbf{G}/\mathbf{H})$ is called the index $[\mathbf{G} : \mathbf{H}]$. In case of a finite group $\mathbf{G}$ of order n and subgroup $\mathbf{H}$ of order m one has that $n/m = [\mathbf{G} : \mathbf{H}]$.
Let $\mathbf{G}$ act on S by ρ. The orbits of ρ partition the set S. Let Δ be a complete set of representatives for the partition into orbits. Then:

$$S = \biguplus_{a \in \Delta} \text{orb}_\rho(a) \text{ and } \text{card}(S) = \sum_{a \in \Delta} \text{card}(\text{orb}_\rho(a)) = \sum_{a \in \Delta} \text{card}(\mathbf{G} : \text{stab}_\rho(a))$$

We apply this equation to the action by conjugacy on a finite group and notice that the center is made up by the elements of the center.

Corollary 4.3 (Class-Equation) *Let $\mathbf{G}$ be a finite group. Then*

$$ord(\mathbf{G}) = ord(\text{cent}(\mathbf{G}) + \sum_{a \in \Delta'} [\mathbf{G} : \text{cent}_{\mathbf{G}}(a)]$$

where Δ' is a complete set of non-conjugate elements which are not in the center.

As an immediate consequence of the class equation we get

Theorem 4.4 *Any finite group of prime power has a non-trivial center.*

PROOF. Each term $[\mathbf{G} : \mathrm{cent}_{\mathbf{G}}(a)], a \in \Delta'$, is divisible by p and therefore $\mathrm{ord}(\mathrm{cent}(\mathbf{G}))$ must be divisible by p. □

Any action $\rho : \mathbf{G} \to \mathrm{Sym}(S)$ induces also an action on $\mathcal{P}(S)$ by means of:

$$\rho^* : \mathbf{G} \to \mathrm{Sym}(\mathcal{P}(S)), x \mapsto (A \mapsto \rho(x)(A) = \{\rho(x)(a) \mid a \in A\}, A \subseteq S$$

It is easy to see that $\rho(x \cdot y)(A) = \rho(x)(\rho(y)(A))$ that is, $\rho^*$is a group homomorphism. For every subset A of S,

$$\mathrm{orb}_{\rho^*}(A) = \{\rho(x)(A) \mid x \in \mathbf{G}\}$$

consists of all subsets B of S that are *similar to A with respect to ρ*. Of course, $\mathrm{card}(B) = \mathrm{card}(A)$ for every $B \in \mathrm{orb}_{\rho^*}(A)$.

For every subset $A \subseteq S$,

$$\mathrm{stab}_{\rho^*}(A) = \{x \in \mathbf{G} \mid \rho(x)(A) = A\}$$

is the subgroup of all $x \in \mathbf{G}$ for which $\rho(x)$ permutes A.

Let now $\sigma : \mathbf{G} \to \mathrm{Sym}(\mathbf{G})$ be the representation of $\mathbf{G}$ by inner automorphisms. Note that for every subgroup $\mathbf{H}$ of $\mathbf{G}$ the image of $\mathbf{H}$ under σ is also a subgroup of $\mathbf{G}$. Therefore,

$$\sigma^* : \mathbf{G} \to \mathrm{Sym}(\mathrm{Sub}(\mathbf{G})), x \mapsto \sigma^*_x, \mathbf{H} \mapsto x\mathbf{H}x^{-1}$$

is a representation of $\mathbf{G}$. One has that

$$\mathrm{orb}_{\sigma^*}(\mathbf{H}) = \{\mathbf{K} \mid \mathbf{K} = x\mathbf{H}x^{-1} \text{ for some } x \in \mathbf{G}\}$$

is the system of all subgroups $\mathbf{K}$ of $\mathbf{G}$ which are conjugate to $\mathbf{H}$. One has that

$$\mathbf{H} \text{ is normal iff } \mathrm{orb}_{\sigma^*}(\mathbf{H}) = \{\mathbf{H}\}$$

The stabilizer of $\mathbf{H}$ in $\mathbf{G}$ is

$$\mathrm{stab}_{\sigma^*}(\mathbf{H}) = \{x \in \mathbf{G} \mid x\mathbf{H}x^{-1} = \mathbf{H}\}$$

and is called the *normalizer* $N(\mathbf{H})$ of $\mathbf{H}$. Note:

(a) $N(\mathbf{H}) \supseteq \mathbf{H}$

(b) $\mathbf{H}$ is normal in $N(\mathbf{H})$

(c) Let $\mathbf{K} \supseteq \mathbf{H}$ be any subgroup of $\mathbf{G}$ in which $\mathbf{H}$ is normal. Then $x\mathbf{H}x^{-1} = \mathbf{H}$ for all $x \in \mathbf{K}$. But then $\mathbf{K} \subseteq N(\mathbf{H})$.

The normalizer $N(\mathbf{H})$ of $\mathbf{H}$ is the largest subgroup of $\mathbf{G}$ in which $\mathbf{H}$ is normal. Of course, $N(\mathbf{H}) = \mathbf{G}$ iff $\mathbf{H}$ is normal. The Counting Lemma 4.2 now yields:

$$\mathrm{card}\{\mathbf{K} \mid \mathbf{K} \text{ conjugate to } \mathbf{H}\} = \mathrm{card}(\mathbf{G}/N(\mathbf{H}))$$

If $\mathbf{G}$ is finite then the number of subgroups $\mathbf{K}$ that are conjugate to $\mathbf{H}$ equals the index of the normalizer $N(\mathbf{H})$ in $\mathbf{G}$, in particular this number is a divisor of $\mathbf{G}$.

Theorem 4.5 (Isomorphism Lemma) *Let* $\mathbf{H}$ *and* $\mathbf{K}$ *be subgroups of the group* $\mathbf{G}$. *Assume that* $\mathbf{H} \subseteq N(\mathbf{K})$. *Then one has:*

(a) $\mathbf{H} \cdot \mathbf{K} = \{h \cdot k \mid h \in H, k \in K\}$ *is a subgroup of* $\mathbf{G}$.

(b) $\mathbf{K}$ *is normal in* $\mathbf{H} \cdot \mathbf{K}$ *and* $\mathbf{H} \cap \mathbf{K}$ *is normal in* $\mathbf{H}$.

(c) $\mathbf{H} \cdot \mathbf{K}/\mathbf{K} \cong \mathbf{H}/\mathbf{H} \cap \mathbf{K}$.

PROOF. We have for any $h \in \mathbf{H}$ that $h\mathbf{K}h^{-1} = \mathbf{K}$, i.e., $hk = k'h$. Thus, $(hk)^{-1} = k^{-1}h^{-1} = h^{-1}k' \in \mathbf{HK}$. Similarly, $(hk)(h'k') = (hh')(k''k') \in \mathbf{HK}$. This proves (a). We have that $(hk)k'(hk)^{-1} = hkk'k^{-1}h^{-1} = hk''h^{-1} \in \mathbf{K}$. This proves that $\mathbf{K}$ is normal in $\mathbf{HK}$. Let $l \in \mathbf{H} \cap \mathbf{K}$. Then $m = hlh^{-1}$ is in $\mathbf{K}$ because $l \in \mathbf{K}$ and $\mathbf{H}$ is within the normalizer of $\mathbf{K}$; $m \in \mathbf{H}$ because $l \in \mathbf{H}$. Thus m is in $\mathbf{H} \cap \mathbf{K}$ and this proves (b). Finally,

$$\mathbf{H} \to \mathbf{HK}/\mathbf{K}, h \mapsto h\mathbf{K}$$

is surjective because $(hk)\mathbf{K} = h(\mathbf{K})$ and clearly homomorphic. We have $h\mathbf{K} = \mathbf{K}$ iff $h \in \mathbf{K}$, i.e., $h \in \mathbf{H} \cap \mathbf{K}$. This proves (c) by the homomorphism theorem. □

A special case is where $\mathbf{K}$ is a normal subgroup $\mathbf{N}$ of $\mathbf{G}$. Then

$$\mathbf{HN}/\mathbf{N} \cong \mathbf{H}/\mathbf{H} \cap \mathbf{N}$$

This is sometimes called the second isomorphism theorem.

4.2 Sylow Theorems

Let $\mathbf{G}$ be a finite group of order $n = p_1^{k_1} \cdots p_m^{k_m}$. As a partial converse of Lagrange's theorem we have

Theorem 4.6 *If the prime power* p^k *divides the order* n *of the finite group* $\mathbf{G}$ *then* $\mathbf{G}$ *contains a subgroup of order* p^k.

PROOF. We prove this by induction on the order n of the group $\mathbf{G}$. Of course, $n = 1$ and $n = 2$ are trivial. Assume that $\mathbf{G}$ has order n and that the theorem is true for all groups of order $< n$. We will use the class equation:

$$\text{ord}(\mathbf{G}) = \text{ord}(\text{cent}(\mathbf{G})) + \sum_{a \in \Delta'} [\mathbf{G} : \text{cent}_{\mathbf{G}}(a)]$$

Assume first that p **does not divide** ord(cent($\mathbf{G}$). Because p divides ord($\mathbf{G}$) one has some $a \in \Delta'$ such that p does not divide $[\mathbf{G} : \text{cent}_{\mathbf{G}}(a)]$. But

$$\text{ord}(\mathbf{G}) = \text{ord}(\text{cent}_{\mathbf{G}}(a)) \cdot [\mathbf{G} : \text{cent}_{\mathbf{G}}(a)]$$

and because p^k divides ord($\mathbf{G}$) one has that p^k divides ord(cent$_{\mathbf{G}}(a)$). Now, ord(cent$_{\mathbf{G}}(a)$) $< n$ because $a \notin$ cent($\mathbf{G}$). Hence, cent$_{\mathbf{G}}(a)$ contains a subgroup of order p^k by induction hypothesis.

Now assume that p **does divide** ord(cent($\mathbf{G}$). As an abelian group, cent($\mathbf{G}$)) is the direct sum of its primary components, and, clearly, cent($\mathbf{G}$) contains an element c of order p. We can form the factor group $\mathbf{G}/ < c >$ which has order n/p and therefore, p^{k-1} divides ord($\mathbf{G}/ < c >$). By induction hypothesis, $\mathbf{G}/ < c >$ contains a subgroup $\overline{H}$ of order p^{k-1}. Under the canonical projection

$$\phi : \mathbf{G} \twoheadrightarrow \mathbf{G}/ < c >$$

the counter image $\phi^{-1}(\overline{H})$ is a subgroup $\mathbf{H} \supseteq < c >$ and $\mathbf{H}/ < c >= \overline{H}$. We have that

$$\text{ord}(\mathbf{H}) = \text{ord}(< c >) \cdot \text{ord}(\mathbf{H}/ < c >) = p \cdot p^{k-1} = p^k$$

□

Let p^m be the **largest** power of p that divides $n = \mathrm{ord}(\mathbf{G})$. Then $\mathbf{G}$ contains a subgroup of order p^m. Any such subgroup is called a *p-Sylow* subgroup of $\mathbf{G}$. They are obviously maximal amongst $p-$subgropups,i.e., subgroups whose order is a power of p. In the abelian case there is only one $p-$Sylow subgroup, namely the primary component T_p .

Theorem 4.7 (M. Ludwig Sylow (1832-1910)) *Let* $\mathbf{G}$ *be a finite group. Then:*

(i) *If* $\mathbf{H}$ *is a* $p-$*subgroup of* $\mathbf{G}$ *then* $\mathbf{H}$ *is contained in some* $p-$*Sylow subgroup.*

(ii) *All* $p-$*Sylow subgroups are conjugate to each other.*

(iii) *The number* l *of* $p-$*Sylow subgroups is congruent to* 1 *modulo* p *and a divisor of ord*$(\mathbf{G})$*, hence*

$$l = 1 \text{ or } p+1 \text{ or } 2p+1, \ldots \text{ and } l|n.$$

Proof. Let M_p be the set of all $p-$Sylow subgroups of $\mathbf{G}$ and let $\mathbf{G}$ act on M_p by conjugation:

$$\sigma : \mathbf{G} \to \mathrm{Sym}(M_p), \sigma(x) : P \mapsto xPx^{-1}$$

Let $P_0 \in M_p$. Then $|\operatorname{orb}_\sigma(P_0)| = [\mathbf{G} : N(P_0)]$ where $N(P_0) \supseteq P_0$. Hence,

$$|\operatorname{orb}_\sigma(P_0)| = |\{xP_0x^{-1} \mid x \in \mathbf{G}\}|$$

is not divisible by p.

Now, let $\mathbf{H}$ be any $p-$subgroup of $\mathbf{G}$. Also, $\mathbf{H}$ acts on $\operatorname{orb}_\sigma(P_0)$ by conjugation:

$$\sigma_{\mathbf{H}} : \mathbf{H} \to \mathrm{Sym}(\operatorname{orb}_\sigma(P_0)), \sigma_{\mathbf{H}}(h) : P \mapsto hPh^{-1}$$

and

$$\operatorname{orb}_\sigma(P_0) = \operatorname{orb}_{\sigma_{\mathbf{H}}}(P_0) \uplus \cdots \uplus \operatorname{orb}_{\sigma_{\mathbf{H}}}(P_{k-1})$$

each of the $|\operatorname{orb}_{\sigma_{\mathbf{H}}}(P_\nu)|$ divides $\mathrm{ord}(\mathbf{H})$ and therefore they are powers of p. But p does not divide the sum $|\operatorname{orb}_\sigma(P_0)|$, and therefore

$$\operatorname{orb}_{\sigma_{\mathbf{H}}}(P_\nu) = \{P_\nu\}$$

for some ν, or, equivalently, $hP_\nu h^{-1} = P_\nu$ for all $h \in \mathbf{H}$. This is $\mathbf{H} \subseteq N(P_\nu)$. By the Isomorphism Lemma 4.5 ($\mathbf{K} = P_\nu$):

$$\mathbf{H}P_\nu/P_\nu \cong \mathbf{H}/\mathbf{H} \cap P_\nu$$

The group on the right-hand side is a group whose order is a power of p. The same must hold for the left-hand side and we conclude that

$$\mathrm{ord}(\mathbf{H}P_\nu) = \mathrm{ord}(\mathbf{H}/\mathbf{H} \cap P_\nu) \cdot \mathrm{ord}(P_\nu)$$

Hence, $\mathbf{H}P_\nu$ has an order which is a power of p. But P_ν is a maximal $p-$group and therefore $\mathbf{H}P_\nu = P_\nu$. This is

$$\mathbf{H} \subseteq P_\nu$$

and proves (i).

We saw that for any $p-$Sylow group P_0 and any $p-$subgroup $\mathbf{H}$ one has a $p-$Sylow group P_ν which is conjugate to P_0 such that $\mathbf{H} \subseteq P_\nu$. If we take now for $\mathbf{H}$ any $p-$Sylow group P then we see that P is contained in a Sylow group P_ν which is conjugate to P_0. But then, by maximality, $P_\nu = P$. Thus, all $p-$Sylow groups are conjugate. This proves (ii).

Thus, we have for any $p-$Sylow group P_0 and any $p-$subgroup $\mathbf{H}$

$$\mathrm{orb}_\sigma(P_0) = M_p = \mathrm{orb}_{\sigma_\mathbf{H}}(P_0) \uplus \cdots \uplus \mathrm{orb}_{\sigma_\mathbf{H}}(P_{k-1})$$

and $\mathbf{H} \subseteq P_\nu$ for some ν and $\mathrm{orb}_{\sigma_\mathbf{H}}(P_\nu)$ for all such ν. If we take $\mathbf{H} = P_0$ then $P_0 = P_\nu$ only for $\nu = 0$ and therefore,

$$l = 1 + \underbrace{\cdots\cdots}_{\text{powers of } p}$$

Of course, $l|n$ because l is an index. This proves (iii). □

Examples

1. No group of order 20 is simple. Any such group has exactly one Sylow subgroup of order 5, i.e., a non-trivial normal subgroup.

2. No group of order 30 is simple. Here the argument is somewhat more complicated. We have that $n = 2 \cdot 3 \cdot 5$. The possible numbers of Sylow subgroups for $p = 2$ are: 1, 5 and 15; for $p = 3$: 1 and 10; for $p = 5$: 1 and 6. All these Sylow subgroups are of prime order, thus cyclic and every element different from e is a generator. Thus, the intersection of of any two different Sylow groups is $\{e\}$. Hence, any simple group of order 30 must have at least 4 elements of order 2, $20 = 10 \cdot 2$ elements of order 3 and $24 = 6 \cdot 4$ elements of order 5. Of course, this is impossible. Hence, any group of order 30 has for some $p = 2$ or 3 or 5 a normal Sylow subgroup.

Chapter 5

Some Universal Constructions

5.1 Peano Algebras

Let $\mathcal{V}$ be a variety of similar algebras of type Δ. An algebra $\mathbf{F}(M)$ of $\mathcal{V}$ is called the *free $\mathcal{V}$-algebra, freely generated by M* if any map φ from M into any algebra $\mathbf{A}$ in $\mathcal{V}$ admits an extension to a homomorphism φ^* from $\mathbf{F}(M)$ into $\mathbf{A}$. Because we assume that M generates $\mathbf{F}(M)$, the extension is unique. A free algebra is uniquely determined up to an isomorphism over the generating set M.

Examples

1. Let $\mathcal{V}$ be the variety of vector spaces over the field $\mathbf{F}$. Because every vector space has a base, they are all free.
2. Let $\mathcal{M}$ be the variety of modules over the principal ideal domain $\mathbf{D}$. For the set M, $\mathbf{D}^{(M)}$ is free, freely generated by the M-many unit vectors.

We are going to show that varieties admit for any set M free algebras. Actually, this holds even for quasi-varieties of algebras. Recall, a quasi-variety is a class of algebras that is abstract, i.e., closed under isomorphic copies, and closed under subalgebras and direct products. A quasi-variety is non-trivial, if it contains algebras with more than one element. The direct product of the empty family of algbras of a given type Δ is the one element algebra with $\{0\}$ as carrier set. Before we prove the existence of free algebras for arbitrary quasi-varieties, we prove the existence of free algebras for the class $\mathcal{V}_\Delta$ of all algebras of type Δ.

Theorem 5.1 *For any set M and algebraic type $\Delta = (n_i)_{i\in I}$ there is an algebra $\mathbf{P} = \mathbf{P}(M, \Delta)$ such that the generalized Peano axioms hold:*

P1: $f_i(a_0, \ldots, a_{n_i-1}) \notin M$.

P2: $f_i(a_0, \ldots, a_{n_i-1}) = f_j(b_0, \ldots, b_{n_j-1})$ *only if* $i = j$ *and* $a_\nu = b_\nu$.

P3: *The set M generates $\mathbf{P}$.*

Proof. We first remark that we get the familiar Peano axioms for the natural numbers $\mathbb{N}$ for the case where the type contains only one unary operation, called the successor $'$ and where M is a one element set $\{0\}$. Then P1 is $n' \neq 0$, P2 is $n' = m'$ only if $n = m$ and P3 is induction.

For the proof, we take a set $F = \{f_i\}$ of *operation symbols* where $I \to F$ is a bijection. We also assume that F and M are disjoint. We define the algebra P as the algebra of terms. Terms are defined as the smallest set which is closed under the following conditions:

(a) Each element $m \in M$ is a term.

(b) If $t_0, \ldots, t_{n-1}$ are terms and if f_i is an $n-ary$ operation symbol, i.e., $n_i = n$, then $t = f_i(t_0, \ldots, t_{n-1})$ is a term. In particular, if $n_i = 0$, then f_i is a term.

According to property (b), P is an algebra of type Δ and the Peano Axioms are easily verified. □

Theorem 5.2 *The Peano algebra $\mathbf{P}(M, \Delta)$ is the free algebra, freely generated by M for the variety of all algebras of type Δ.*

PROOF. The proof is straight forward on the complexity of terms. □

We can define the Peano Algebra also as the set of well-formed words. Let $\mathbf{W}$ be the set of all strings of elements over the set $M \cup F$. If $w_0, \ldots, w_{n-1}$ are words and if $f = f_i$ is an $n-$ary operation symbol then $w = f_i w_0 \ldots w_{n-1}$ is a word. This way, $\mathbf{W}$ becomes an algebra of type Δ. The subalgebra that is generated by the set M is the algebra of terms, i.e., the Peano algebra over M. We can single out the terms amongst all words by a simple algorithm:
Define the *valency* function on the algebra of words to the integers by:

(a) $v(f_i) = -n_i + 1$, $v(m) = 1$.

(b) For a word w define v as the sum of valencies of its letters $c \in M \cup F$.

One has the following theorem of P. Hall:

Theorem 5.3 *A word $w = c_1 c_2 \ldots c_n$ is a term if and only if*

(a) $v(w) = 1$,

(b) $v(s_i) > 0$ *for every right segment $s_i = c_i \ldots c_n$.*

Moreover, w is a product of $r-$many terms if and only if the first condition is replaced by $v(w) = r$.

The proof is an interesting exercise.

Theorem 5.4 *Let $\mathcal{Q}$ be any non-trivial quasi-primitve class of algebras of type Δ. Then for any set M there exists a free $\mathcal{Q}-$algebra, freely generated by M.*

PROOF. Let $\mathcal{E}$ be the set of all congruences E of $\mathbf{P} = \mathbf{P}(M, \Delta)$ where the factor algebra $\mathbf{A}_E$ is a member of $\mathcal{Q}$. The family of projections:

$$q_E : \mathbf{P} \to \mathbf{A}_E, E \in \mathcal{E}$$

leads to a homomorphism

$$q : \mathbf{P} \to \mathbf{A} = \prod_{E \in \mathcal{E}} \mathbf{A}_E$$

such that

$$p_E \circ q = q_E$$

and where p_E are the canonical projections from the direct product $\mathbf{A}$ to the factors $\mathbf{A}_E$. We have that

$$\ker(q) = E_0 = \bigcap_{E \in \mathcal{E}} E$$

seperates the points of M. That is, if a and b are different points of M then $(a,b) \notin E_0$. Indeed, let $\mathbf{A}$ be any algebra in $\mathcal{Q}$ with at least two elements a' and b'. Then choose any map φ from M into $\mathbf{A}$ such that $\varphi(a) = a'$ and $\varphi(b) = b'$. For the kernel E of the homomorphic extension of φ we have that $(a,b) \notin E$ and therefore $(a,b) \notin E_0$. We have that the image $\mathbf{B} = \mathrm{im}(q)$ of q is a subalgebra of $\mathbf{A}$ which is a direct product of algebras in $\mathcal{Q}$. Thus $\mathbf{B}$ is an algebra that belongs to $\mathcal{Q}$. Because

$$\mathbf{B} \cong \mathbf{P}/E_0 = \mathbf{A}_{E_0}$$

we have that $\mathbf{A}/E_0$ is an algebra that belongs to $\mathcal{Q}$. We claim that $\mathbf{A}/E_0$ is the free $\mathcal{Q}-$algebra, freely generated by the set $\overline{M} = \{[m]_{E_0} | m \in M\}$ of equivalence classes of elements in M. Indeed, let $\overline{\varphi}$ be any map from $\overline{M}$ into an algebra $\mathbf{C} \in \mathcal{Q}$. Because an equivalence class for E_0 cannot have more than one element of M in it, we get a map $\varphi : M \to \mathbf{C}$ for which $\varphi(m) = \overline{\varphi}([m])$. The homomorphic extension φ^* has a kernel $E \supseteq E_0$ and therefore we have a unique homomorphism

$$\overline{\varphi}^* : \mathbf{A}_{E_0} \to \mathbf{C}$$

such that $\overline{\varphi}^* \circ q_{E_0} = \varphi^*$ Replacing the $[m]$ by m proves our claim. □

We mention some important facts and leave the proofs as exercises.

1. If $N \subseteq M$ then the subalgebra of $\mathbf{F}(M, \mathcal{Q})$ that is generated by N is $\mathbf{F}(N, \mathcal{Q})$. Of course, this does not imply that a subalgebra of a free algebra is free.

2. A subalgebra of a Peano Algebra is a Peano Algebra.

3. Let $M = \biguplus_{t \in T}$ be a partition of the set M into disjoint subsets N_t of M. Then

   $$i_t : \mathbf{F}(N_t, \mathcal{Q}) \to \mathbf{F}(M, \mathcal{Q})$$

   is a co-product system in $\mathcal{Q}$. The i_t are the homomorphic extensions of the inclusions $N_t \hookrightarrow M$ and are injective. In particular, the free algebra $\mathbf{F}(M, \mathcal{Q})$ is the co-product of free $\mathcal{Q}-$algebras that are freely generated by one element sets.

4. The free $\mathcal{Q}-$algebra, freely generated by the empty set is an *initial* object of $\mathcal{Q}$. That is, for every algebra $\mathbf{A} \in \mathcal{Q}$ there is exactly one map from $\mathbf{F}(\emptyset, \mathcal{Q})$ into $\mathbf{A}$. This map is the homomorphic extension of the empty map from $\emptyset$ into $\mathbf{A}$. For the class of groups, $\{e\}$ is initial. For the class of rings it is $\mathbb{Z}$. If the type does not contain constants then the empty set is an algebra in $\mathcal{Q}$ and it must be the initial algebra of $\mathcal{Q}$.

5. A trivial quasi-primitive class contains at any rate the one-element algebra as the empty product. It contains the empty set if the type is without constants. If the type is without constants then the empty set is the free $\mathcal{Q}-$algebra otherwise it is the one-element algebra.

Theorem 5.5 *A quasi-primitive class of algebras admits sums.*

PROOF. Let $\mathcal{Q}$ be quasi-primitive and let

$$\mathbf{A}_t = (A, (f_i)_{i \in I})$$

be given. We form:

$$M = \bigcup_{t \in T} \{t\} \times A_t = \{(t, a) \mid t \in T, a \in A_t\}$$

and extend the type Δ of $\mathcal{Q}$ by an $M-$indexed set of constants:

$$\Delta^* = ((n_i)_{i\in I}, (n_{(t,a)})_{(t,a)\in M}), n_{(t,a)} = 0$$

An algebra of type Δ^* is of the form

$$\mathbf{A}^* = (\mathbf{A}, (c_{(t,a)})_{(t,a)\in M})$$

where $\mathbf{A}$ is of type Δ and $c_{(t,a)} \in A$.

$$\varphi : (\mathbf{A}, (c_{(t,a)}) \to (\mathbf{B}, (d_{(t,a)})$$

is a homomorphism of Δ^*-algebras if

$$\varphi : \mathbf{A} \to \mathbf{B} \text{ is a } \Delta- \text{ homomorphism and } \varphi(c_{(t,a)}) = d_{(t,a)}, (t,a) \in M$$

A subalgebra of $\mathbf{A}^*$ is a subalgebra $\mathbf{B}$ of $\mathbf{A}$ that contains all $c_{(t,a)} \in M$.

The direct product of algebras $\mathbf{B}^*_s, s \in S$, is given by

$$\mathbf{B}^* = (\mathbf{B}, ((c^s_{(t,a)})_{s\in S})_{(t,a)\in M})$$

where $\mathbf{B}$ is the direct product of the algebras $\mathbf{B}_s$.

We call an algebra $\mathbf{B}^* = (\mathbf{B}, (c_{(t,a)})_{(t,a)\in M})$ *special* for $\mathbf{A}_t, t \in T$ if

$$\mathbf{B} \in \mathcal{Q}, \text{ and } \varphi_t : \mathbf{A}_t \to \mathbf{B}, a \mapsto c_{(t,a)} \text{ are } \Delta-\text{homomorpic.}$$

If c_i is a constant that belongs to the type Δ with values $c^t_i \in \mathbf{A}_t$ then one must have in $\mathbf{B}^*$:

$$\varphi_t(c^t_i) = c^{\mathbf{B}}_i = c_{(t,c^t_i)}, t \in T$$

and

$$\varphi_t(f^t_i(a_0, \ldots, a_{n_i-1}) = f^{\mathbf{B}}_i(c_{(t,a_0)}, \ldots, c_{(t,a_{n_i-1})}) = c_{(t,f^t_i(a_0,\ldots,a_{n_i-1})}, t \in T$$

which is

$$c^{\mathbf{B}}_i = c_{(t,c^t_i)} \text{ and } f^{\mathbf{B}}_i(c_{(t,a_0)}, \ldots, c_{(t,a_{n_i-1})}) = c_{(t,f^t_i(a_0,\ldots,a_{n_i-1}))}, t \in T$$

That is, an algebra $\mathbf{B}^*$ is special for the family $\mathbf{A}_t$ iff $\mathbf{B} \in \mathcal{Q}$ and these equations hold between the constants $c_{(t,a)}, (t,a) \in M$. It is therefore obvious that:

$$\mathcal{B}^* = \{\mathbf{B}^* \mid \mathbf{B}^* \text{ special for } \mathbf{A}_t, t \in T\}$$

is a quasi-primitive class of algebras. (It is primitive, in case that $\mathcal{Q}$ is.)

Now let $\mathbf{A}^* = \mathbf{F}(\emptyset, \mathcal{B}^*)$ be the free $\mathcal{B}^*-$algebra, freely generated by the empty set. The algebra $\mathbf{A}^*$ exists, even if $\mathcal{B}^*$ is trivial. For $\mathbf{A}^* = (\mathbf{A}, (c_{(t,a)})_{(t,a)\in M})$ we have

$$\mathbf{A} \in \mathcal{Q} \text{ and } i_t : \mathbf{A}_t \to \mathbf{A}, a \mapsto c_{(t,a)}, t \in T, \text{ are homomorphic}$$

We claim that

$$(\mathbf{A}, i_t : \mathbf{A}_t \to \mathbf{A})$$

is a sum system in $\mathcal{Q}$.

Let $f_t : \mathbf{A}_t \to \mathbf{B}$ be a family of homomorphisms in $\mathcal{Q}$. Then one has that

$$\mathbf{B}^* = (\mathbf{B}, (c_{(t,a)} = f_t(a))_{(t,a)\in M}) \in \mathcal{B}^*$$

Because $\mathbf{A}^*$ is free, there is a unique homomorphism

$$f : \mathbf{A}^* \to \mathbf{B}^*, c_{(t,a)} \mapsto f_t(a), (t,a) \in M$$

which is $f(i_t(a)) = f_t(a)$ for all $t \in T$ and $a \in \mathbf{A}_t$. Thus,

$$f \circ i_t = f_t, t \in T.$$

□

5.2 Ultraproducts

Let T be any set. A collection $\mathfrak{f}$ of subsets of T is called a filter if the following holds:

(a) If $A \in \mathfrak{f}$ and $B \in \mathfrak{f}$ then $A \cap B \in \mathfrak{f}$.

(b) If $A \in \mathfrak{f}$ and if $B \supseteq A$ then $B \in \mathfrak{f}$.

(c) $T \in \mathfrak{f}$.

A filter $\mathfrak{f}$ is *proper* if

(d) $\emptyset \notin \mathfrak{f}$.

Let T be an infinite set. Then the system of all co-finite subsets S of T, i.e., the sets for which the complement $T \setminus S$ is finite, is a proper filter. It is called the *Fréchet* filter on T.
For a topological space X and a point $x \in X$ the sets N which are neighborhoods of x form a filter $\mathfrak{n}(x)$.

Definition 17 A subset $\mathfrak{b}$ of the filter $\mathfrak{f}$ is called a *base* for $\mathfrak{f}$ if for every $A \in \mathfrak{b}$ one has some $B \in \mathfrak{b}$ such that $A \supseteq B$.

The open sets that contain x form a base of the neighborhood filter $\mathfrak{n}(x)$. For the set $\mathbb{N}$ the sets $N_k = \{m | m \geq k\}$ form a base of the Fréchet filter.

Let $\mathfrak{s}$ be a system of subsets of T. Then:

$$< \mathfrak{s} >= \{V \mid V \supseteq S_1 \cap \cdots \cap S_k, k \in \mathbb{N}, S_i \in \mathfrak{s}\}$$

is a filter. It is a proper filter if and only if $\mathfrak{s}$ has the finite intersection property.

A filter $\mathfrak{f}$ is *principal* if it is generated by one of its sets, i.e., $\mathfrak{f} =< \{S\} >$. For a finite set T, every filter is principal. More generally, a filter that contains a finite set has to be principal.

Definition 18 A proper filter $\mathfrak{u}$ is called an *ultrafilter* if for every subset S of T one has that either S or its complement $T \setminus S$ belongs to $\mathfrak{u}$.

Filters that have a singleton as a base are certainly ultrafilters. They are the trivial ultrafilters. Filters are partially ordered by inclusion. Because the union of a chain of proper filters is a proper filter, we have according to Zorn's Lemma that every filter $\mathfrak{f}$ is contained in a maximal filter $\mathfrak{m}$.

Theorem 5.6 *A filter is maximal if and only if it is an ultrafilter.*

PROOF. Assume that $\mathfrak{u}$ is an ultrafilter and that $\mathfrak{f} \supseteq \mathfrak{u}$ where $\mathfrak{f}$ is proper. Then for every $S \in \mathfrak{f}$ one has that $S \in \mathfrak{u}$ because otherwise one would have $T \setminus S \in \mathfrak{u}$ and therefore $S \cap T \setminus S = \emptyset \in \mathfrak{f}$. Thus $\mathfrak{u}$ is maximal.
Now assume that $\mathfrak{m}$ is maximal and that $S \notin \mathfrak{m}$. Then, because $\mathfrak{m}$ is maximal, $\mathfrak{m} \cup \{S\}$ cannot have the finite intersection property. This is, $S \cap F = \emptyset$ for some $F \in \mathfrak{m}$. But this is $F \subseteq T \setminus S$ and therefore $T \setminus S \in \mathfrak{m}$. Hence, $\mathfrak{m}$ is an ultrafilter. □

Ultrafilters are sometimes called *prime filters*. Recall, that an ideal P in a ring is prime if $ab \in P$ only if $a \in P$ or $b \in P$. Filters are the dual to ideals. Thus, a filter $\mathfrak{p}$ should be called prime if $A \cup B \in \mathfrak{p}$ only if $A \in \mathfrak{p}$ or $B \in \mathfrak{p}$. We then have that a filter is a prime filter if and only if it is an ultrafilter. Indeed, a prime filter $\mathfrak{p}$ is an ultrafilter because $A \cup (T \setminus A) = T \in \mathfrak{p}$, so $A \in \mathfrak{p}$ or $T \setminus A \in \mathfrak{p}$. Now assume for the ultrafilter $\mathfrak{u}$ that $S = S_1 \cup S_2 \in \mathfrak{u}$. If we had $S_1 \notin \mathfrak{u}$ and $S_2 \notin \mathfrak{u}$ then $(T \setminus S_1) \cap (T \setminus S_2) = T \setminus (S_1 \cup S_2) \in \mathfrak{u}$.

Theorem 5.7 *Every proper filter $\mathfrak{f}$ is the intersection of all maximal filters $\mathfrak{m}$ which contain $\mathfrak{f}$. In particular, for an infinite set, the Fréchet filter $\mathfrak{c}$ is the intersection of all proper ultrafilters.*

PROOF. If the set S does not belong to the proper filter $\mathfrak{f}$ then $\mathfrak{f} \cup \{T \setminus S\}$ has the finite intersection property. Thus, there is a maximal filter which contains $T \setminus S$ and $\mathfrak{f}$ and therefore not S.
Let $\mathfrak{u}$ be a proper ultrafilter. Because it cannot contain a finite subset of T it must contain $\mathfrak{c}$. □

Definition 19 Let $A_t, t \in T$, be a family of sets and let

$$A = \prod_{t \in T} A_t$$

be their direct product. Then for any filter $\mathfrak{f}$ on T one has that

$$\alpha = (a_t) \equiv_{\mathfrak{f}} \beta = (b_t) \text{ iff } \{t \mid a_t = b_t\} \in \mathfrak{f}$$

is an equivalence relation on A. The set

$$A_{\mathfrak{f}} = \{C \mid C = [\alpha], \alpha \in A\}$$

of equivalence classes is called the $\mathfrak{f}-$*reduced* product of the sets A_t.

It is quite easy to see that $\equiv_{\mathfrak{f}}$ is an equivalence on A. Reflexivity holds because $T \in \mathfrak{f}$ and symmetry is totally obvious. If $S_1 = \{t \mid a_t = b_t\} \in \mathfrak{f}$ and if $S_2 = \{t \mid b_t = c_t\} \in \mathfrak{f}$ then

$$S = \{t \mid a_t = c_t\} \supseteq \{t \mid a_t = b_t \text{ and } b_t = c_t\} = S_1 \cap S_2 \in \mathfrak{f}$$

and this shows transitivity.

For $\mathfrak{f} = \{T\}$ the reduced product becomes the ordinary product. If $\mathfrak{f} =< \{t\} >$, then the reduced product is equivalent to the factor A_t.

Assume that the $\mathbf{A}_t$ are algebras of a fixed type Δ and let $f = f_i$ be one of the operations defined on every $\mathbf{A}_t$. Assume, for simplicity, that thea rity of f is two. We then define on the reduced product an operation f by

$$f([\alpha], [\beta]) = [f(\alpha, \beta)]$$

This is a well-defined operation. Assume that $\alpha \equiv \alpha', \beta \equiv \beta'$. We then have that $S_1 = \{t \mid \alpha(t) = \alpha'(t)\} \in \mathfrak{f}$ and that $S_2 = \{t \mid \beta(t) = \beta'(t)\}\mathfrak{f}$. But then on $S = S_1 \cap S_2$ one has that $\alpha = \alpha'$ and $\beta = \beta'$ and therefore $f(\alpha(t), \beta(t)) = f(\alpha'(t), \beta'(t))$. This is $f(\alpha, \beta) \equiv f(\alpha', \beta')$. We have on the reduced product the operations defined in a way that $\alpha \mapsto [\alpha]$ is a homomorphism. Thus we have

Theorem 5.8 *Primitive classes are closed under reduced products.*

If the algebras $\mathbf{A_t}$ carry also a relational structure of type Δ', i.e., on every algebra relations R_j of arity m_j are defined then one defines on the reduced product a relation similarly as we have done for equality. For example, if $R = R_j$ is binary then $R_j([\alpha],[\beta])$ should hold if and only if $\{t \mid R(\alpha(t),\beta(t)) \text{ holds on } \mathbf{A_t}\} \in \mathfrak{f}$. Again, it is easy to see that this is a proper definition.

We have already defined the algebra of terms over a set M. In order to define elementary propositions, we take for M a countable set X of variables $x_1, x_2, \ldots$.

Atomic formulas are defined inductively:

(a) $t = s$ is atomic for terms s and t.

(b) Let $R = R_j$ be an $m-$ary relational symbol and let $t_1, \ldots, t_m$ be terms. Then $R(t_1, \ldots, t_m)$ is atomic.

Formulas then are defined according to the following rules. Of course, atomic formulas are formulas. And then

(c) If α and β are formulas then

$$(\alpha \vee \beta), (\alpha \wedge \beta), \neg\alpha$$

are formulas.

(d) If α is a formula then

$$\forall x\alpha, \exists x\alpha$$

are formulas.

A formula without free variables is called a (first order) sentence p. A sentence is either true or false. One now has the famous

Theorem 5.9 (Ultraproduct Theorem) *A sentence holds in an ultraproduct* $\mathbf{A}_\mathfrak{u}$ *if an only if it holds in* $\mathfrak{u}-$*almost all* $\mathbf{A}_t$.

That a property holds for $\mathfrak{f}-$almost all $t \in T$ means that the subset S of those t for which the property holds is an element of $\mathfrak{f}$. The proof of the ultraproduct theorem goes by induction on the complexity of formulas and is omitted. As a typical example one might want to prove directly that an ultraproduct of fields is a field.

If we take for T the set $\mathbb{N}$ of natural numbers then an ultrapower ${}^*\mathbb{R}$ of $\mathbb{R}$ is a totally ordered field that contains a copy of $\mathbb{R}$. The map

$$d : a \mapsto [(a = a_n)_{n\in\mathbb{N}}]$$

is an order embedding. The field ${}^*\mathbb{R}$ is non-archimedian. We have that

$$\omega = [(n)_{n\in\mathbb{N}}] > [(k)]$$

for every $k \in \mathbb{N}$ and is therefore an infinite hyper-natural number. It's multiplicative inverse is an infinitesimal, i.e., a number that is smaller than any ordinary positive real number. Any ultrapower of the reals can serve as a simple model for a calculus with infinitesimals.

Bibliography

[1] Hungerford, Thomas W., *Algebra*, Graduate Texts in Mathematics 74, Springer Verlag, New York, Heidelberg, Berlin.

[2] Jacobson, Nathan, *Basic Algebra*, W.H. Freeman and Company, San Francisco.