# Algebra I

Thomas Keilen
Mathematics Institute
University of Warwick

Lecture Notes

October 2002

# Contents

CHAPTER I

# Finite Groups

## 0 Motivation - Check Digit Codes

Nowadays products in shops all carry bar codes and are identified by them. Moreover, at the cash desk the bar code is scanned or typed in and that way you get charged the price. Sometimes the bar codes are not recognised correctly or the wrong number has been typed in. However, the error is recognised by the machine and the bar code is not accepted.

A) Have you ever wondered how it comes, that you are always charged the right price?

Well, the machine looks the bar code up in some data base, and if the incorrect bar code was contained in that data base as well, then the machine could not possibly detect any error. So, when assigning bar codes, you have to make sure that no bar codes which - in a certain sense - are too similar are in the data base.
Is this difficult? Well, to decide on that question we should know, what bar codes in principle look like!
Bar codes are also called **EAN-13** codes, where EAN is short for European Article Number, and they consist of a thirteen digit number. The first 2 to 3 digits stand for the organisation which assigned the numbers to the producer, some of the next digits identify this producer and so on. So, the digits are not really arbitrary digits. In particular, for a fixed producer a large part of the bar code will always be the same. I. e. the numbers will have to be similar!
How can we get along with that problem?

**Idea:** Store some *redundant information* which is not needed to identify the article, but only to detect possible errors.

In the case of the EAN-13 only 12 digits characterise the article. Digit no. 13 is a so called **check digit**.

B) How is the check digit related to the (real) article number?

**Basic Idea:** It should be possible to calculate the check digit from the remaining digits in an easy way, but such that (common) errors are possibly detected.

**First Idea:** Repeat the whole number! This is a bit too much redundancy and increases the risk of falsely scanned numbers.

**Second Idea:** Take the cross sum of the digits of the real product number as check "digit".

E. g. if the product number is 013412547180, then the check digit would be

$$0+1+3+4+1+2+5+4+7+1+8+0=36.$$

This will usually be several digits long, and is still too much redundancy.

**Third Idea:** Let's just take the last digit of the cross sum!
E. g. in the above example the check digit would then be 6.
This can be formulated in a more mathematical way by saying that

we take the remainder of the cross sum by division with remainder modulo 10.

And that's where groups come into play as a nice way to formulate the procedure. We may identify the digits $0,\dots,9$ with the elements of the additive group $(\mathbb{Z}/10\mathbb{Z},+)$, just via the map

$$\{0,\dots,9\}\to\mathbb{Z}/10\mathbb{Z}:a\mapsto\overline{a}=a+10\mathbb{Z}=\{a+10z\mid z\in\mathbb{Z}\},$$

i. e. identifying the digit with the residue class represented by the number. Viewing the digits in the article number as elements of $\mathbb{Z}/10\mathbb{Z}$ that way, the check digit becomes just the sum of the "digits".
E. g. $\overline{0}+\overline{1}+\overline{3}+\overline{4}+\overline{1}+\overline{2}+\overline{5}+\overline{4}+\overline{7}+\overline{1}+\overline{8}+\overline{0}=\overline{36}=\overline{6}$.

C) Does this allow to detect errors? Otherwise it is of no use.

Certainly we will not be able to detect all errors, thus we have to distinguish certain types of errors! Some statistics tell us that the following two types are the most common ones.

- **Type I:** "Single Digit Errors" – i. e. just one digit is wrong. These are roughly 80% of the occuring errors.
- **Type II:** "Neighbour Transpositions" – i. e. two neighbouring digits have been interchanged. These are about 10% of the errors.

It is fairly obvious that the cross-sum-mod-10-approach cannot detect errors of Type II, since the addition in $\mathbb{Z}/10\mathbb{Z}$ is commutative. However, does it detect errors of Type I?
Suppose the correct number was $a_1a_2\cdots a_{13}$ and instead of some $a_i$ we read $a_i'\in\{0,\dots,9\}$ with $a_i\neq a_i'$. Then

$$\overline{a_{13}}-\left(\sum_{j\neq i,13}\overline{a_j}+\overline{a_i'}\right)=\sum_{j=1}^{12}\overline{a_j}-\left(\sum_{j\neq i,13}\overline{a_j}+\overline{a_i'}\right)=\overline{a_i-a_i'}\neq\overline{0},\quad(1)$$

since $a_i-a_i'$ is number between $-9$ and $9$ which is non-zero and thus 10 does not divide $a_i-a_i'$. That means "Single Digit Errors" are detected.

D) Back to EAN-13.

The encoding of EAN-13 is, however, slightly different. The check digit in $a_1a_2\cdots a_{13}$ satisfies

$$\overline{a_{13}}=(-1)\cdot\overline{a_1}+(-3)\cdot\overline{a_2}+(-1)\cdot\overline{a_3}+(-3)\cdot\overline{a_4}+\ldots+(-1)\cdot\overline{a_{13}}$$

or equivalently

$$\overline{a_1}+3\cdot\overline{a_2}+\overline{a_3}+\ldots+\overline{a_{13}}=\overline{0}.$$

We call these equations **check digit equations**.
Does this still detect errors of Type I?
Let's go back to Equation (1) for this. The question finally comes down to checking whether $a_i \neq a_i'$ implies that $\overline{a_i - a_i'}$ and $\overline{3 \cdot (a_i - a_i')}$ are not equal to $\overline{0}$, which is the case since $a_i - a_i'$ is not divisible by 10 and thus also three times this number is not. Thus we are lucky.
How about errors of Type II?
If $a_i$ and $a_{i+1}$ have been interchanged, then this comes down to the question whether

$$\begin{aligned} & 3 \cdot \overline{a_i} + \overline{a_{i+1}} = 3 \cdot \overline{a_{i+1}} + \overline{a_i} \\ \Leftrightarrow \quad & \overline{2 \cdot (a_i - a_{i+1})} = \overline{0} \\ \Leftrightarrow \quad & 5 \mid a_i - a_{i+1}. \end{aligned}$$

Thus even errors of Type II will quite frequently be detected, but not all of them. We achieved this by multiplying the digits in the cross sum by certain weights $w_i$ – here $w_i = 1$ and $w_i = 3$.

E) WHICH WEIGHTS $w_i$ WOULD HAVE BEEN SUITABLE IN THE CHECK DIGIT EQUATION IN ORDER NOT TO LOOSE THE PROPERTY THAT ERRORS OF TYPE I ARE DETECTED?

The important point was that

$$\overline{a_i} \neq \overline{a_i'} \quad \Rightarrow \quad \omega_i \cdot \overline{a_i} \neq \omega_i \cdot \overline{a_i'},$$

i. e. that the map

$$\mu_{\omega_i} : \mathbb{Z}/10\mathbb{Z} \to \mathbb{Z}/10\mathbb{Z} : \overline{a} \mapsto \omega_i \cdot \overline{a}$$

is injective, and hence bijective since $\mathbb{Z}/10\mathbb{Z}$ is a finite set. In other words, $\mu_{\omega_i}$ is a permutation of the set $\mathbb{Z}/10\mathbb{Z}$.
This leads to the following generalisation and definition.

**0.1 Definition**
Let $(G, \cdot)$ be a group, $g_0 \in G$ a fixed element, and let $\pi_1, \ldots, \pi_n \in \mathrm{Sym}(G)$ be permutations.

a. We call

$$C = C_G(\pi_1, \ldots, \pi_n, g_0) = \left\{ (g_1, \ldots, g_n)^t \in G^n \;\middle|\; \pi_1(g_1) \cdots \pi_n(g_n) = g_0 \right\}$$

a *check digit code* (CDC) of *length* $n$ on the *alphabet* $G$.

b. We say that $C$ detects errors of Type I if and only if $(g_1, \ldots, g_n)^t \in C$ and $g_i' \in G$ with $g_i' \neq g_i$ implies that $(g_1, \ldots, g_{i-1}, g_i', g_{i+1}, \ldots, g_n)^t \notin C$.

c. We say that $C$ detects errors of Type II if and only if $(g_1, \ldots, g_n)^t \in C$ with $g_i \neq g_{i+1}$ implies that $(g_1, \ldots, g_{i-1}, g_{i+1}, g_i, g_{i+2}, \ldots, g_n)^t \notin C$.

**0.2 Example (EAN-13)**
Let $(G, \cdot) = (\mathbb{Z}/10\mathbb{Z}, +)$, $g_0 = \overline{0}$, $n = 13$, $\pi_i = \mu_1$ if $i$ is odd and $\pi_i = \mu_3$ if $i$ is even. This then describes the EAN-13 code $C = C_{\mathbb{Z}/10\mathbb{Z}}(\mu_1, \mu_3, \ldots, \mu_1, \overline{0})$.

Actually, $C = \ker(\phi)$, where $\phi : (\mathbb{Z}/10\mathbb{Z})^{13} \to \mathbb{Z}/10\mathbb{Z}$ is the group homomorphism defined by multiplication with the matrix $(\overline{1}, \overline{3}, \overline{1}, \ldots, \overline{1})$.

Having introduced check digit codes over arbitrary groups it would be nice to know something about their error detecting properties.

**0.3 Proposition (Error Detecting Properties)**
*Let* $C = C_G(\pi_1, \ldots, \pi_n, g_0)$ *be a CDC over the alphabet* $(G, \cdot)$.

a. $C$ *detects errors of Type I.*

b. *If* $n \geq 3$, *then* $C$ *detects errors of Type II if and only if* $\forall\ i = 1, \ldots, n-1$, $\forall\ g, h \in G$ *s. t.* $g \neq h$*:*

$$g \cdot \left(\pi_{i+1} \circ \pi_i^{-1}\right)(h) \neq h \cdot \left(\pi_{i+1} \circ \pi_i^{-1}\right)(g).$$

**Proof:** a. Let $(g_1, \ldots, g_n)^t \in C$, $g_i' \in G$ such that $g_i' \neq g_i$, and suppose $(g_1, \ldots, g_i', \ldots, g_n)^t \in C$. Then

$$\pi_1(g_1) \cdots \pi_n(g_n) = g_0 = \pi_1(g_1) \cdots \pi_i\left(g_i'\right) \cdots \pi_n(g_n).$$

By the cancellation law we thus deduce that

$$\pi_i(g_i) = \pi_i\left(g_i'\right).$$

But then also $g_i = g_i'$, since $\pi_i$ is injective. This, however, is a contradiction to our assumption.

b. Let's first assume that the condition of the proposition is satisfied and let's show that then $C$ detects errors of Type II. For this let $(g_1, \ldots, g_n)^t \in C$ be given with $g_i \neq g_{i+1}$ and set $g = \pi_i(g_i)$ and $h = \pi_i(g_{i+1})$. Since $\pi_i$ is injective we have $g \neq h$. Thus by the condition of the proposition we also have

$$\pi_i(g_i) \cdot \pi_{i+1}(g_{i+1}) = g \cdot \left(\pi_{i+1} \circ \pi_i^{-1}\right)(h) \neq h \cdot \left(\pi_{i+1} \circ \pi_i^{-1}\right)(g) = \pi_i(g_{i+1}) \cdot \pi_{i+1}(g_i).$$

Multiplying both sides with the same element of $G$ the inequality is preserved and we get

$$\pi_1(g_1) \cdots \pi_i(g_i) \cdot \pi_{i+1}(g_{i+1}) \cdots \pi_n(g_n) \neq \pi_1(g_1) \cdots \pi_i(g_{i+1}) \cdot \pi_{i+1}(g_i) \cdots \pi_n(g_n).$$

This means that $C$ detects errors of Type II.
Let's now suppose that $C$ detects errors of Type II and then prove the above condition. For this let $g, h \in G$ with $g \neq h$, and set $g_i = \pi_i^{-1}(g)$ and $g_{i+1} = \pi_i^{-1}(h)$. Since $\pi_i$ is bijective $g_i \neq g_{i+1}$. Choose now $g_j \in G$, $j \neq i, i+1$ such that $(g_1, \ldots, g_n)^t \in C$ (here we need $n \geq 3$). Thus by assumption

$$(g_1, \ldots, g_{i+1}, g_i, \ldots, g_n)^t \notin C.$$

But then

$$\pi_1(g_1) \cdots \pi_n(g_n) = g_0 \neq \pi_1(g_1) \cdots \pi_i(g_{i+1}) \cdot \pi_{i+1}(g_i) \cdots \pi_n(g_n).$$

Using the cancellation law we derive

$$g \cdot \left(\pi_{i+1} \circ \pi_i^{-1}\right)(h) = \pi_i(g_i) \cdot \pi_{i+1}(g_{i+1}) \neq \pi_i(g_{i+1}) \cdot \pi_{i+1}(g_i) = h \cdot \left(\pi_{i+1} \circ \pi_i^{-1}\right)(g).$$

This finishes the proof.

□

**Note:** If $(G,\cdot)$ is *abelian* and $\mathrm{inv} : G \to G : g \mapsto g^{-1}$ denotes the inversion map, then the condition in Proposition 0.3 comes down to

$$g \cdot \left(\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1}\right)(g) \neq h \cdot \left(\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1}\right)(h). \tag{2}$$

Since $\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1} \in \mathrm{Sym}(G)$ is a permutation of $G$, it seems that maps of the form $g \mapsto g\pi(g)$ for some permutation $\pi \in \mathrm{Sym}(G)$ are connected to the error detecting properties of codes.

**0.4 Definition**
Let $(G,\cdot)$ be a group and $\pi \in \mathrm{Sym}(G)$. We call $\pi$ a *complete mapping* if and only if the map

$$\pi^* : G \to G : g \mapsto g \cdot \pi(g)$$

is again a permutation of $G$.

So far we know how to check whether a given CDC detects errors of Typer II or not, but we have no means to find such a code – or possibly to decide that their is non.

**0.5 Corollary**
*Let $(G,\cdot)$ be a finite abelian group, $n \geq 3$. Then there is a CDC of length $n$ which detects errors of Type II if and only if $G$ admits a complete mapping.*

**Proof:** Let's first suppose that $G$ admits a complete mapping $\pi \in \mathrm{Sym}(G)$. Set $g_0 = e_G$ and $\pi_i = (\mathrm{inv} \circ \pi)^i$ for $i = 1, \ldots, n$.
**Claim:** $C = C_G(\pi_1, \ldots, \pi_n, g_0)$ detects errors of Type II.
For this we only have to check that Equation (2) is satisfied. Let $g, h \in G$ such that $g \neq h$. Then

$$\begin{aligned} g \cdot \left(\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1}\right)(g) &= g \cdot \left(\mathrm{inv} \circ (\mathrm{inv} \circ \pi)^{i+1-i}\right)(g) = g \cdot \pi(g) = \pi^*(g) \\ &\neq \pi^*(h) = h \cdot \pi(h) = h \cdot \left(\mathrm{inv} \circ (\mathrm{inv} \circ \pi)^{i+1-i}\right)(h) = h \cdot \left(\mathrm{inv} \circ \pi_{i+1} \circ \pi_i^{-1}\right)(h). \end{aligned}$$

Thus Equation (2) is fulfilled.
Let's now suppose that there is a CDC $C_G(\pi_1, \ldots, \pi_n, g_0)$ which detects errors of Type II. We define $\pi = \mathrm{inv} \circ \pi_2 \circ \pi_1^{-1} \in \mathrm{Sym}(G)$ and we claim that this is then a complete mapping. In order to check this we let $g, h \in G$ such that $g \neq h$. Thus by Equation (2) we have

$$\pi^*(g) = g \cdot \pi(g) = g \cdot \left(\mathrm{inv} \circ \pi_2 \circ \pi_1^{-1}\right)(g) \neq h \cdot \left(\mathrm{inv} \circ \pi_2 \circ \pi_1^{-1}\right)(h) = h \cdot \pi(h) = \pi^*(h).$$

Hence $\pi^*$ is injective and thus bijective, since $G$ is finite. But then $\pi$ is a complete mapping. □

**0.6 Remark**

a. If $|G| = 2 \cdot m$ with $m$ odd, then there exists *no* complete mapping on $G$.[1]
In particular, there is no CDC on $\mathbb{Z}/10\mathbb{Z}$ which detects all errors of Type II.

b. If $|G|$ is odd, then the identity mapping $\mathrm{id}_G$ is a complete mapping.

[1]The proof is elementary, but lengthy. We refer the reader to H. Siemon, Anwendungen der elementaren Gruppentheorie in der Zahlentheorie, 1981.

**Proof:** Let $|G| = 2m+1$, then by the Theorem of Lagrange we have $e_G = g^{|G|} = g^{2m+1}$. Multiplying by $g$ we get $\left(g^{m+1}\right)^2 = g$, and thus the mapping $\mathrm{id}_G^* : g \mapsto g \cdot \mathrm{id}_G(g) = g^2$ is surjective. But since $G$ is finite, it is then bijective. □

c. **Problem:** There is no CDC on $(\mathbb{Z}/10\mathbb{Z}, +)$ which detects errors of Type II! How can we deal with that?

**Solution 1:** Use an odd number of digits, i. e. calculate over $\mathbb{Z}/m\mathbb{Z}$ with an odd $m$.
E. g. the ISBN code works over $(\mathbb{Z}/11\mathbb{Z}, +)$, where the element $\overline{10} = 10 + 11\mathbb{Z}$ is denoted by $X$ and is only used as check digit. The ISBN code is a $C_{\mathbb{Z}/11\mathbb{Z}}(\pi_1, \ldots, \pi_{10}, \overline{0})$ code, where $\pi_i : \mathbb{Z}/11\mathbb{Z} \to \mathbb{Z}/11\mathbb{Z} : \overline{a} \mapsto i \cdot \overline{a}$. We leave it as an exercise to check that the code actually detects errors of Type II. You only have to check that Equation (2) is satisfied.

**Solution 2:** Use a non-abelian group with ten elements! There the non-existence of a complete mapping is not related to the error detecting property.

**0.7 Example (German Currency)**
The check digits of the serial numbers of the German currency where actually encoded by a $C_{\mathbb{D}_{10}}\left(\pi_1, \ldots, \pi_{10}, \mathrm{id}_{\mathbb{D}_{10}}, (1)\right)$ code.
Consider the *dihedral group*

$$\mathbb{D}_{10} = \left\langle (1\ 2\ 3\ 4\ 5), (1\ 5)(2\ 4) \right\rangle \leq \mathbb{S}_5 = \mathrm{Sym}\left(\{1, \ldots, 5\}\right).$$

In the exercises you show that, setting $\sigma = (1\ 2\ 3\ 4\ 5)$ and $\tau = (1\ 5)(2\ 4)$, we may describe $\mathbb{D}_{10}$ as the set

$$\mathbb{D}_{10} = \left\{\sigma^0 = (1), \sigma^1, \ldots, \sigma^4, \tau \circ \sigma^0 = \tau, \tau \circ \sigma^1, \ldots, \tau \circ \sigma^4\right\}.$$

And since $\tau \circ \sigma = \sigma^{-1} \circ \tau \neq \sigma \circ \tau$, the group is indeed not abelian.
Verhoeff showed that the permutation $\pi : \mathbb{D}_{10} \to \mathbb{D}_{10}$ of $\mathbb{D}_{10}$ defined by

| $x$ | $\sigma^0$ | $\sigma^1$ | $\sigma^2$ | $\sigma^3$ | $\sigma^4$ | $\tau \circ \sigma^0$ | $\tau \circ \sigma^1$ | $\tau \circ \sigma^2$ | $\tau \circ \sigma^3$ | $\tau \circ \sigma^4$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $\pi(x)$ | $\sigma^1$ | $\tau \circ \sigma^0$ | $\tau \circ \sigma^2$ | $\tau \circ \sigma$ | $\sigma^2$ | $\tau \circ \sigma^3$ | $\sigma^3$ | $\sigma^0$ | $\tau \circ \sigma^4$ | $\sigma^4$ |

satisfies that $g, h \in \mathbb{D}_{10}$ with $g \neq h$ implies $g \circ \pi(h) \neq h \circ \pi(g)$. Hence, setting $\pi_i = \pi^i \in \mathrm{Sym}(\mathbb{D}_{10})$, the code $C_{\mathbb{D}_{10}}\left(\pi_1, \ldots, \pi_{10}, (1)\right)$ detects errors of Type II by Proposition 0.3.
Of course for the serial numbers on the German currency they did not use such fancy symbols like $\sigma$. They used the usual 10 digits and in addition 10 letters. However, they were identified with the elements in $\mathbb{D}_{10}$ in the following way

| $\sigma^0$ | $\sigma^1$ | $\sigma^2$ | $\sigma^3$ | $\sigma^4$ | $\tau \circ \sigma^0$ | $\tau \circ \sigma^1$ | $\tau \circ \sigma^2$ | $\tau \circ \sigma^3$ | $\tau \circ \sigma^4$ |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| A | D | G | K | L | N | S | U | Y | Z. |

Thus, if you wanted to check whether a serial number on a German bank note was valid, you replaced the digits and letters by the appropriate elements of $\mathbb{D}_{10}$ and looked whether this element belonged to $C_{\mathbb{D}_{10}}\left(\pi_1, \ldots, \pi_{10}, \mathrm{id}_{\mathbb{D}_{10}}, (1)\right)$.

**0.8 Exercise**
Check if AA6186305Z2 is a valid serial number for a German bank note.

**Question:** Could we have used some other group with 10 elements as alphabet?

**Answer:** No! Not really. The group only matters up to isomorphism, and we will show at the end of the part on finite groups that up to isomorphism there are only two groups with 10 elements – $(\mathbb{Z}/10\mathbb{Z}, +)$ and $(\mathbb{D}_{10}, \circ)$.

# 1 Basics

Let's recall some of the basic definitions and results from first year courses.

A) GROUPS

**1.1 Definition**
A *group* is a tuple $(G, \cdot)$ consisting of a non-empty set $G$ and a binary operation

$$\cdot : G \times G \to G : (g, h) \mapsto g \cdot h$$

such that the following axioms are fulfilled:

(i) $g \cdot (h \cdot k) = (g \cdot h) \cdot k$ for all $g, h, k \in G$, (Associativity)

(ii) $\exists\, e \in G : \forall\, g \in G \;:\; e \cdot g = g$, (Existence of a Neutral)

(iii) $\forall\, g \in G\ \exists\, h \in G : h \cdot g = e$. (Existence of an Inverse)

If moreover

(iv) $g \cdot h = h \cdot g$ for all $g, h \in G$

is satisfied, then we call $(G, \cdot)$ *abelian*.
If $|G| < \infty$, we call the group *finite* and $|G| = o(G) = \#G$ is called its *order*.

**Notation:** Instead of $g \cdot h$ we will usually just write $gh$. If a group is abelian, then we will usually denote the operation by "+" instead of "$\cdot$".
If no ambiguity concerning the group operation can arise, we will just write $G$ instead of $(G, \cdot)$ in order denote a group.

**1.2 Proposition**
*Let $(G, \cdot)$ be a group.*

a. *The neutral element $e_G$ is uniquely determined and satisfies $g \cdot e_G = g$ for all $g \in G$ as well. Instead of $e_G$ we also write $1_G$.*

b. *For any element $g \in G$ the inverse element is uniquely determined and is denoted by $g^{-1}$ or $\mathrm{inv}_G(g)$. It satisfies $g \cdot g^{-1} = e_G$ as well.*

c. Cancellation Rule*: If $g, h, k \in G$ with $gh = gk$ or with $hg = kg$, then $h = k$.*

d. *For $g, h \in G$ we have $(g \cdot h)^{-1} = h^{-1} \cdot g^{-1}$ and $\left(g^{-1}\right)^{-1} = g$.*

e. *If we set $g^0 = e_G$ and, recursively, $g^{i+1} = g \cdot g^i$ and $g^{-i} = \left(g^i\right)^{-1}$ for $i \geq 0$, then the* exponential laws *are fulfilled, i. e. for $g \in G$ and $i, j \in \mathbb{Z}$ we have*

$$g^i \cdot g^j = g^{i+j} \quad \text{and} \quad \left(g^i\right)^j = g^{i \cdot j}.$$

**Proof:**

a./b. Let's prove Parts a. and b. together in several steps, where $e_G$ denotes a fixed (left-)neutral in $G$ for which every element has a (left-)inverse as indicated by the group axioms.

***Step 1*:** If $h \cdot g = e_G$ for $g, h \in G$, then also $g \cdot h = e_G$.

Since $G$ is a group, there is some $k \in G$ such that $k \cdot h = e_G$. Hence

$$g \cdot h = e_G \cdot (g \cdot h) = (k \cdot h) \cdot (g \cdot h) = k \cdot (h \cdot g) \cdot h = k \cdot e_G \cdot h = k \cdot h = e_G.$$

***Step 2*:** We also have $g \cdot e_G = g$ for all $g \in G$.

Let $g \in G$ and let $h \in G$ such that $h \cdot g = e_G$. Then, using Step 1,

$$g \cdot e_G = g \cdot (h \cdot g) = (g \cdot h) \cdot g = e_G \cdot g = g.$$

***Step 3*:** Let $e' \in G$ such that for all $g \in G$ we have, $e' \cdot g = g$, then $e' = e_G$.

Using Step 2, we have $e' = e_G \cdot e' = e' \cdot e_G = e_G$.

***Step 4*:** Let $k, h \in G$ such that $k \cdot g = e_G = h \cdot g$, then $k = h$.

By Step 1 we know that $g \cdot h = e_G$, thus we get with the aid of Step 2

$$k = k \cdot e_G = k \cdot (g \cdot h) = (k \cdot g) \cdot h = e_G \cdot h = h.$$

c. Let $g, h, k \in G$ such that $h \cdot g = k \cdot g$. Then

$$h = h \cdot e_G = h \cdot \left(g \cdot g^{-1}\right) = (h \cdot g) \cdot g^{-1} = (k \cdot g) \cdot g^{-1} = k \cdot \left(g \cdot g^{-1}\right) = k \cdot e_G = k.$$

The other way round works analogously.

d. Let $g, h \in G$. In order to see that $(g \cdot h)^{-1} = h^{-1} \cdot g^{-1}$, it suffices to show that the right hand side has the property of the inverse element of $g \cdot h$. Knowing that that one is uniquely determined we are then done.

$$\left(h^{-1} \cdot g^{-1}\right) \cdot (g \cdot h) = h^{-1} \cdot \left(g^{-1} \cdot g\right) \cdot h = h^{-1} \cdot e_G \cdot h = h^{-1} \cdot h = e_G.$$

Thus $h^{-1} \cdot g^{-1}$ is the unique inverse of $g \cdot h$, i. e. $h^{-1} \cdot g^{-1} = (g \cdot h)^{-1}$.
Analogously, for $g \in G$ we have by Part b.

$$g \cdot g^{-1} = e_G,$$

and hence $g$ satisfies the property of the inverse element of $g^{-1}$. Hence by unicity we get $\left(g^{-1}\right)^{-1} = g$.

e. Note that the definition implies right away

$$g^k = \left(g^{-1}\right)^{-k} \quad \forall\, g \in G, \forall\, k \in \mathbb{Z}.$$

Let's now prove the first exponential law, and for this let $i, j \in \mathbb{Z}$.
*1st Case:* Let $g \in G$ be arbitrary, $i \geq 0$. We do the proof by induction on $i$.
$i = 0$ : Then $g^i \cdot g^j = g^0 \cdot g^j = e_G \cdot g^j = g^j = g^{i+j}$.
$i \mapsto i + 1$ : By definition and induction hypothesis:

$$g^{i+1} \cdot g^j = \left(g \cdot g^i\right) \cdot g^j = g \cdot \left(g^i \cdot g^j\right) = g \cdot g^{i+j} = g^{i+1+j}.$$

*2nd Case:* Let $g \in G$ arbitrary, $i < 0$. Apply Case 1 to the element $g^{-1}$, then by definition we get (since $-i > 0$!)

$$g^i \cdot g^j = \left(g^{-1}\right)^{-i} \cdot \left(g^{-1}\right)^{-j} = \left(g^{-1}\right)^{-i-j} = g^{i+j}.$$

Let's now turn to the second exponential law, and let again $i, j \in \mathbb{Z}$, $g \in G$.
*1st Case:* $j \geq 0$. We do the proof by induction on $j$.
$j = 0$ : Then $\left(g^i\right)^j = \left(g^i\right)^0 = e_G = g^0 = g^{i \cdot j}$.
$j \mapsto j + 1$ : By definition, induction hypothesis and the first exponential law we get:

$$\left(g^i\right)^{j+1} = \left(g^i\right) \cdot \left(g^i\right)^j = g^i \cdot g^{i \cdot j} = g^{i+i \cdot j} = g^{i \cdot (j+1)}.$$

*2nd Case:* $j < 0$. By the first exponential law we have $g^{-i} \cdot g^i = g^{-i+i} = g^0 = e_G$, and thus $\left(g^i\right)^{-1} = g^{-i}$. By Case 1 and definition we get (since $-j > 0$!):

$$\left(g^i\right)^j = \left(\left(g^i\right)^{-1}\right)^{-j} = \left(g^{-i}\right)^{-j} = g^{(-i)\cdot(-j)} = g^{i \cdot j}.$$

□

**Notation:** If the group is abelian and the group operation is denoted by $+$, then we denote the neutral element rather by $0_G$ and the inverse of $g \in G$ by $-g$. Moreover, instead of $g^i$ we then write $i \cdot g$.

**1.3 Example** a. $(\mathbb{Z}, +)$ is an abelian group with neutral element $0$.

b. Let $(R, +, \cdot)$ be a ring (e. g. the integers) and $n \geq 1$ an integer. $\mathrm{Mat}(n \times n, R)$, the set of all $n \times n$-matrices with entries in $R$ forms a group with respect to matrix addition as binary operation. The neutral element is the zero matrix, and the inverse of $(a_{ij})$ is just $(-a_{ij})$.

c. Let $(K, +, \cdot)$ be any field (e. g. the real numbers) and let $n \geq 1$ be an integer. $\mathrm{Gl}_n(K) = \{(a_{ij}) \in \mathrm{Mat}(n \times n, K) \mid (a_{ij}) \text{ is invertible}\}$, the set of all invertible $n \times n$-matrices with entries in the field $K$, forms a group with respect to matrix multiplication as binary operation. The neutral element is the identity matrix and the inverse of an element is just its inverse matrix.

d. Let $M$ be any set. $\mathrm{Sym}(M) = \{\varphi : M \to M \mid \varphi \text{ is bijective}\}$, the set of all *permutations* of $M$, is a group with respect to the composition of maps. The neutral element is the identity map $\mathrm{id}_M$, and the inverse of an element $\varphi$ is its inverse mapping.

**1.4 Example (The Symmetric Group $\mathbb{S}_n$)**
When studying finite groups one group attracts a particular interest as an infinite source for interesting examples – this is the symmetric group of $n$ letters

$$\mathbb{S}_n = \mathrm{Sym}\left(\{1, \ldots, n\}\right).$$

An element $\pi \in \mathbb{S}_n$ can be represented in the form

$$\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \pi(1) & \pi(2) & \pi(3) & \cdots & \pi(n) \end{pmatrix}$$

or, if $\{1,\dots,n\}=\{a_1,\dots,a_n\}$ by

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ \pi(a_1) & \pi(a_2) & \pi(a_3) & \cdots & \pi(a_n) \end{pmatrix}.$$

The elements of $\mathbb{S}_n$ are called *permutations*, and there is a particular type of permutations called *cycles* – a permutation of the form

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_{k-1} & a_k & a_{k+1} & \cdots & a_n \\ a_2 & a_3 & \cdots & a_k & a_1 & a_{k+1} & \cdots & a_n \end{pmatrix}$$

is called a $k$-cycle and we write instead just $(a_1\ a_2\ \cdots\ a_k)$. Very simple cycles are 2-cycles $(a\ b)$, and they are called *transpositions*.

**Note:** The representation of a $k$-cycle is not unique –

$$(a_1\ a_2\ \cdots\ a_k)=(a_2\ a_3\ \cdots\ a_k\ a_1)=\dots.$$

And the neutral element, i. e. the identity map on $\{1,\dots,n\}$, is represented by any 1-cycle, i. e. $(1)=(2)=\dots=(n)$. We usually denote it by $(1)$.

**Facts:** a. *Cycle-Decomposition*: Every permutation $\pi\in\mathbb{S}_n$ has a unique representation as a product of disjoint cycles (unique up to ordering).

E. g. $\pi=\begin{pmatrix} 1&2&3&4&5&6&7\\ 3&2&1&5&6&7&4\end{pmatrix}\in\mathbb{S}_7$, then $\pi=(1\ 3)(2)(4\ 5\ 6\ 7)=(1\,3)(4\,5\,6\,7)$.

b. Every permutation $\pi\in\mathbb{S}_n$ can be written as a product of transpositions, and the parity of the number of necessary transpositions is uniquely determined. If the parity is even, then we say $\pi$ has *sign* $\operatorname{sgn}(\pi)=1$ and we call the permutation *even*, otherwise $\operatorname{sgn}(\pi)=-1$ and $\pi$ is said to be *odd*.

E. g. for the above permutation we have $\pi=\begin{pmatrix} 1&2&3&4&5&6&7\\ 3&2&1&5&6&7&4\end{pmatrix}=(1\,3)(4\,7)(4\,6)(4\,5)=(1\,3)(4\,7)(4\,6)(4\,5)(2\,3)(2\,3)$ – the parity is even.

B) Subgroups

**1.5 Definition and Proposition**

Let $(G,\cdot)$ be a group.

a. A non-empty subset $\emptyset\neq U\subseteq G$ is called a *subgroup* of $G$ if and only if (one of the) following equivalent conditions is fulfilled:

(i) $U$ is itself a group with respect to the restriction of the binary operation $\cdot$ to $U\times U$.

(ii) For all $u,v\in U$ we have $u\cdot v\in U$ and $u^{-1}\in U$.

(iii) For all $u,v\in U$ we have $u\cdot v^{-1}\in U$.

If $|U|<\infty$, then these are also equivalent to:

(iv) For all $u,v\in U$ we have $u\cdot v\in U$.

We denote this by $(U,\cdot)\leq(G,\cdot)$ or simply by $U\leq G$.

Note, if $U\leq G$, then $e_G=e_U\in U$!

b. If $U,V\subseteq G$ are two subsets, then we define $U\cdot V=\{u\cdot v\mid u\in U,v\in V\}$.

c. If $M \subseteq G$ is any subset, then we call

$$\langle M \rangle = \bigcap_{M \subseteq U \leq G} U$$

the *subgroup generated by* $M$, and this is by Proposition 1.7 indeed a group.

**Proof:** We have to prove the equivalences in the definition of a subgroup. For this we denote throughout the proof for $g \in G$ by $\mathrm{inv}_G(g)$ the inverse of $g$ in $G$ and for $u \in U$ by $\mathrm{inv}_U(u)$ the inverse of $u \in U$. We will show, that indeed $\mathrm{inv}_U(u) = \mathrm{inv}_G(u)$ for all $u \in U$!
(i) $\Longrightarrow$ (ii): Let's first show that, if $(U, \cdot)$ is itself a group, then $e_U = e_G$ and $\mathrm{inv}_U(u) = \mathrm{inv}_G(u)$ for all $u \in U$. For this note that

$$e_G = \mathrm{inv}_G(e_U) \cdot e_U = \mathrm{inv}_G(e_U) \cdot (e_U \cdot e_U) = (\mathrm{inv}_G(e_U) \cdot e_U) \cdot e_U = e_G \cdot e_U = e_U,$$

and thus $\mathrm{inv}_U(u) \cdot u = e_U = e_G = \mathrm{inv}_G(u) \cdot u$, which then by the cancellation law implies $\mathrm{inv}_U(u) = \mathrm{inv}_G(u)$.
Thus for any $u \in U$ we have $\mathrm{inv}_G(u) = \mathrm{inv}_U(u) \in U$ as desired, and for $u, v \in U$ it follows $u \cdot v \in U$, since by assumption the restriction of "$\cdot$" to $U \times U$ takes values in $U$.
(ii) $\Longrightarrow$ (i): By assumption $u \cdot v \in U$ for all $u, v \in U$, and hence

$$\cdot : U \times U \to U$$

is actually a binary operation taking values in $U$. It, therefore, suffices to check that the group axioms are fulfilled. Associativity comes for free, since it is already satisfied for elements from the larger set $G$. Moreover, if we could show that $e_G \in U$ and for any $u \in U$ also $\mathrm{inv}_G(u) \in U$, then we are done, since these elements satisfy the properties of the neutral respectively the corresponding inverse element. However, for $u \in U$ we have $\mathrm{inv}_G(u) \in U$ by assumption, and since $U \neq \emptyset$, we may choose some $v \in U$, so that again $\mathrm{inv}_G(v) \in U$ and thus by the closedness assumption

$$e_G = v \cdot \mathrm{inv}_G(v) \in U.$$

(ii) $\Longrightarrow$ (iii): Let $u, v \in U$, then by assumption $\mathrm{inv}_G(v) \in U$ and thus also $u \cdot \mathrm{inv}_G(v) \in U$.
(iii) $\Longrightarrow$ (ii): Let again $u, v \in U$. By assumption $e_G = u \cdot \mathrm{inv}_G(u) \in U$, and thus $\mathrm{inv}_G(u) = e_G \cdot \mathrm{inv}_G(u) \in U$ and $u \cdot v = u \cdot \mathrm{inv}_G\big(\mathrm{inv}_G(v)\big) \in U$.
(ii) $\Longrightarrow$ (iv): This is obvious no matter whether $U$ is finite or not.
(iv) $\Longrightarrow$ (ii): It remains to show that for every element $u \in U$ also $\mathrm{inv}_G(u) \in U$. We claim, that $\mathrm{inv}_G(u) = u^k$ for some $k \geq 0$, which then by the closedness assumption implies that it belongs to $U$.
Since $|U| < \infty$, also the set $\{u^k \mid k > 0\}$ is finite. This implies that there are natural numbers $i > j > 0$, such that $u^i = u^j$. But then by the exponential laws we have $u^{i-j-1} = u^{-1}$ and $i - j - 1 \geq 0$. □

**1.6 Example** a. $\mathbb{1} = \{e_G\}$ and $G$ are the *trivial subgroups* of $G$.

b. **Claim:** $(U,+) \leq (\mathbb{Z},+)$ if and only if there is an integer $n \geq 0$ such that $U = n \cdot \mathbb{Z} = \{n \cdot z \mid z \in \mathbb{Z}\}$.

**Proof:** Let $U = n \cdot \mathbb{Z}$ for some $n \geq 0$, then $0 \in U$ and thus $U$ is non-empty. Let $u = nz, v = nz' \in U$, then $u+v = n \cdot (z+z') \in U$ and $-u = n \cdot (-z) \in U$. Thus $U \leq \mathbb{Z}$.
Let now $U \leq \mathbb{Z}$ be an arbitrary subgroup of $\mathbb{Z}$ and suppose $U \neq \{0\}$. We have to find $n > 0$, such that $U = n \cdot \mathbb{Z}$. Let's set

$$n = \min\{u \in U \mid u > 0\}.$$

For this note that $U$ contains some non-zero element $u$ and hence its inverse $-u$, and one of these will thus be strictly greater than zero.
We claim that $U = n \cdot \mathbb{Z}$. Note first of all that with $n \in U$ and $U$ being a subgroup, we have

$$n \cdot z = n + \overset{z-\text{times}}{\cdots} + n,\ n \cdot (-z) = (-n) + \overset{z-\text{times}}{\cdots} + (-n),\ n \cdot 0 = 0 \in U$$

for all $z > 0$. Hence, $n \cdot \mathbb{Z} \subseteq U$.
On the other hand, if we choose an arbitrary $0 \neq u \in U$, then division with remainder modulo $n$ gives uniquely determined integers $z, r \in \mathbb{Z}$ such that

$$u = n \cdot z + r \quad \text{and} \quad 0 \leq r < n.$$

Rearranging the equation and using the fact the $U$ is a subgroup containing $n \cdot \mathbb{Z}$ we find

$$r = u - n \cdot z \in U.$$

But then the minimality assumption on $n$ implies that $r = 0$ and hence $u = n \cdot z \in n \cdot \mathbb{Z}$. □

c. $\mathbb{A}_n = \{\pi \in \mathbb{S}_n \mid \operatorname{sgn}(\pi) = 1\} \leq \mathbb{S}_n$.
E. g. $\mathbb{A}_3 = \big\{(1), (1\,2\,3), (1\,3\,2)\big\} \leq \mathbb{S}_3 = \big\{(1), (1\,2), (1\,3), (2\,3), (1\,2\,3), (1\,3\,2)\big\}$.

d. $\big(\{1,-1\}, \cdot\big) \leq \big(\mathbb{R} \setminus \{0\}, \cdot\big)$.

**1.7 Proposition**
*Let* $(G, \cdot)$ *be a group,* $U, V, U_i \leq G$, $i \in I$, $M \subseteq G$.

a. $\bigcap_{i \in I} U_i \leq G$.

b. $U \cup V \leq G$ *if and only if* $U \subseteq V$ *or* $V \subseteq U$.

c. $\langle M \rangle = \big\{g_1^{\alpha_1} \cdots g_n^{\alpha_n} \mid n \geq 0, g_1, \ldots, g_n \in M, \alpha_1, \ldots, \alpha_n \in \mathbb{Z}\big\}$.

d. $U \cdot V \leq G$ *if and only if* $U \cdot V = V \cdot U$ *if and only if* $U \cdot V = \langle U \cup V \rangle$.

**Proof:** a. This is Exercise 2 on the Assignment Set 3.

b. If $U \subseteq V$ or $V \subseteq U$, then the union is obviously a subgroup. Let's therefore suppose that $U \not\subseteq V$ and $V \not\subseteq U$. Then there are elements $u \in U \setminus V$ and $v \in V \setminus U$. It suffices to show that $u \cdot v \notin U \cup V$. Suppose the contrary. If $u \cdot v \in U$, then $v = u^{-1} \cdot (u \cdot v) \in U$ as well in contradiction to the choice of $v$. And if $u \cdot v \in V$, then $u = (u \cdot v) \cdot v^{-1} \in V$, which gives again a contradiction.

c. We set $N = \{g_1^{\alpha_1} \cdots g_n^{\alpha_n} \mid n \geq 0, g_1, \ldots, g_n \in M, \alpha_1, \ldots, \alpha_n \in \mathbb{Z}\}$.
Let's first show that $N \subseteq \langle M \rangle$. If $U \leq G$ such that $M \subseteq U$, then $g_1^{\alpha_1} \cdots g_n^{\alpha_n} \in U$ for all $g_i \in M$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{Z}$. Thus $N \subseteq U$, and thus $N \subseteq \langle M \rangle$.
It remains to show $\langle M \rangle \subseteq N$. For this it suffices to show that $N \leq G$ with $M \subseteq N$. Since the empty product by convention is $e_G$, $N$ is non-empty. If $h = g_1^{\alpha_1} \cdots g_n^{\alpha_n}, h' = g_{i+1}^{\alpha_{i+1}} \cdots g_m^{\alpha_m} \in N$ are two arbitrary elements, then $h \cdot h' = g_1^{\alpha_1} \cdots g_m^{\alpha_m} \in N$ and $h^{-1} = g_n^{-\alpha_n} \cdots g_1^{-\alpha_1} \in N$. Thus $N \leq G$, and $M \subseteq N$ is fulfilled anyway.

d. Let's first show that $U \cdot V \leq G$ if and only if $U \cdot V = V \cdot U$.
If $U \cdot V \leq G$ and $u \in U$ and $v \in V$ are given, then $v \cdot u = \left(u^{-1} \cdot v^{-1}\right)^{-1} \in U \cdot V$. Hence $V \cdot U \subseteq U \cdot V$, and by symmetry $V \cdot U = U \cdot V$.
Suppose now $V \cdot U = U \cdot V$. Since $e_G \in U, V$, we have $e_G = e_G \cdot e_G \in U \cdot V$, and the latter is non-empty. Let $u, u' \in U$ and $v, v' \in V$ be given. Then by assumption $v \cdot u' \in V \cdot U = U \cdot V$, and thus there are elements $\overline{u} \in U$ and $\overline{v} \in V$ such that $v \cdot u' = \overline{u} \cdot \overline{v}$. Hence,

$$(u \cdot v) \cdot \left(u' \cdot v'\right) = u \cdot \overline{u} \cdot \overline{v} \cdot v' \in U \cdot V,$$

and $(u \cdot v)^{-1} = v^{-1} \cdot u^{-1} \in V \cdot U = U \cdot V$. But thus $U \cdot V \leq G$.
Let's now show that $U \cdot V \leq G$ if and only if $U \cdot V = \langle U \cup V \rangle$.
If $U \cdot V = \langle U \cup V \rangle$, then in particular $U \cdot V \leq G$.
It remains to show that $U \cdot V \leq G$ implies $U \cdot V = \langle U \cup V \rangle$. For this note that $U \cdot V$ contains $U \cup V$, since both $U$ and $V$ contain $e_G$. But being a subgroup of $G$, then

$$\langle U \cup V \rangle = \bigcap_{U \cup V \subseteq H \leq G} H \quad \subseteq \quad U \cdot V.$$

On the other hand

$$U \cdot V \subseteq \{g_1^{\alpha_1} \cdots g_n^{\alpha_n} \mid n \geq 0, g_1, \ldots, g_n \in U \cup V, \alpha_1, \ldots, \alpha_n \in \mathbb{Z}\} = \langle U \cup V \rangle.$$

□

**1.8 Example** a. Let $n, m \geq 0$, then due to the unique factorisation of natural numbers we have $n \cdot \mathbb{Z} \cap m \cdot \mathbb{Z} = \operatorname{lcm}(n, m) \cdot \mathbb{Z}$.

b. Let $n, m \geq 0$, then due to the so called Bézout identity we have $n \cdot \mathbb{Z} + m \cdot \mathbb{Z} = \operatorname{hcf}(n, m) \cdot \mathbb{Z}$.

c. $\langle (1\,2), (1\,2\,3) \rangle = \mathbb{S}_3$, since $(1\,2\,3) = (1\,2) \circ (1\,2\,3) \circ (1\,2)$, $(1\,3) = (1\,2\,3) \circ (1\,2)$ and $(2\,3) = (1\,2) \circ (1\,2\,3)$.

**1.9 Definition and Proposition**
Let $(G, \cdot)$ be a group, $U \leq G$.

a. For $g, h \in G$ we define

$$g \sim_{U,l} h \quad \Longleftrightarrow \quad g^{-1} \cdot h \in U.$$

This defines an equivalence relation on $G$, and the equivalence class of $g$ is just $g \cdot U = \{g \cdot u \mid u \in U\}$. We call the equivalence classes *(left) cosets*.

**Note:** For $g, h \in G$ we have either $gU = hU$ or $gU \cap hU = \emptyset$. Moreover, since any element of $G$ belongs to some coset, $G$ can be written as the disjoint union $G = \coprod_{i \in I} g_i \cdot U$ of certain cosets and we call $\{g_i \mid i \in I\}$ the a system of representatives.

b. Similarly, for $g, h \in G$ we define

$$g \sim_{U,r} h \iff g \cdot h^{-1} \in U.$$

This defines again an equivalence relation on $G$, and the equivalence class of $g$ is just $U \cdot g = \{u \cdot g \mid u \in U\}$. We call the equivalence classes *right cosets*.

**Note:** For $g, h \in G$ we have either $Ug = Uh$ or $Ug \cap Uh = \emptyset$. Moreover, since any element of $G$ belongs to some coset, $G$ can be written as the disjoint union of certain right cosets.

**Note:** $U = e_G \cdot U = U \cdot e_G$ itself is always a left and right coset! Moreover, $g \cdot U = U$ if and only if $g \in U$ if and only if $U \cdot g = U$.

**Proof:** By symmetry it suffices to prove Part a.

***Show*:** $\sim_{U,l}$ is an equivalence relation.

Let $g \in G$, then $g^{-1} \cdot g = e_G \in U$, and thus $g \sim_{U,l} g$, i. e. the relation is reflexive.
If $g, h \in G$ such that $g \sim_{U,l} h$, then $g^{-1} \cdot h \in U$. Thus $h^{-1} \cdot g = \left(g^{-1} \cdot h\right)^{-1} \in U$, which means $h \sim_{U,l} g$ and gives the symmetry of the relation.
If $g, h, k \in G$ such that $g \sim_{U,l} h$ and $h \sim_{U,l} k$, then $g^{-1} \cdot h, h^{-1} \cdot k \in U$. But then also

$$\left(g^{-1} \cdot h\right) \cdot \left(h^{-1} \cdot k\right) = g^{-1} \cdot k \in U,$$

that is, the relation is transitive. So, finally, $\sim_{U,l}$ is an equivalence relation.

***Show*:** The equivalence class of $g \in G$ with respect to $\sim_{U,l}$ is just $g \cdot U$.

By definition, the equivalence class of $g \in G$ is just

$$\{h \in G \mid g \sim_{U,l} h\} = \left\{h \in G \;\middle|\; g^{-1} \cdot h \in U\right\} = \{h \in G \mid h \in g \cdot U\} = g \cdot U.$$

Taking general properties of equivalence relations into account we know that two equivalence classes, which are not disjoint, coincide, and we know that the disjoint union of the different equivalence classes is the whole set $G$. □

**1.10 Example** a. Consider the group $(\mathbb{Z}, +)$ and the subgroup $n\mathbb{Z} \leq \mathbb{Z}$ for $n \geq 0$ fixed. The cosets are then all of the form

$$x + n\mathbb{Z} \quad \text{with} \quad x \in \mathbb{Z}.$$

Since $\mathbb{Z}$ is abelian, left and right cosets coincide! A possible system of representatives is $\{0, 1, \ldots, n-1\}$. Note also that e. g. $4 + 11\mathbb{Z} = 15 + 11\mathbb{Z}$.

b. $G = \mathbb{S}_3$, $U = \langle (1\,2) \rangle$ and $\pi = (1\,2\,3)$, then

$$\pi \circ U = \left\{(1\,2\,3), (1\,3)\right\} \neq \left\{(1\,2\,3), (2\,3)\right\} = U \circ \pi.$$

Thus in general the left and right coset corresponding to an element will not coincide!

**1.11 Theorem (of Lagrange)**
*Let* $(G, \cdot)$ *be a* finite *group,* $V \leq U \leq G$.

a. $\#\{gU \mid g \in G\} = \#\{Ug \mid \ \in G\}$, *i. e. the number of different left cosets coincides with the number of different right cosets. This number is called the* index *of* $U$ *in* $G$ *and is denoted by* $|G : U|$.

b. $|G| = |U| \cdot |G : U|$.
*In particular, the order of a subgroup always divides the order of the group!*

c. $|G : V| = |G : U| \cdot |U : V|$.

d. $|U \cdot V| = \frac{|U| \cdot |V|}{|U \cap V|}$.

**Proof:** We note that for $g \in G$ fixed the map $\alpha : U \to g \cdot U : u \mapsto g \cdot u$ is bijective with inverse $\beta : g \cdot U \to U : v \mapsto g^{-1} \cdot v$. In particular we have for any $g \in G$

$$|U| = |g \cdot U|.$$

a./b. Since $G$ is finite, $\sim_{U,l}$ and $\sim_{U,r}$ lead to finite systems of representatives $\{g_1, \ldots, g_n\}$ and $\{h_1, \ldots, h_m\}$ for left respectively right cosets of $U$ in $G$. In particular, $n$ is the number of different left cosets and $m$ the number of different right cosets. It follows

$$\coprod_{i=1}^{n} g_i \cdot U = G = \coprod_{j=1}^{m} U \cdot h_j,$$

and hence

$$n \cdot |U| = \sum_{i=1}^{n} |g_i \cdot U| = |G| = \sum_{j=1}^{m} |U \cdot h_j| = m \cdot |U|.$$

This, however, implies $n = m = |G : U|$ and $|G| = |U| \cdot |G : U|$.

c. The proof is Exercise 1 on the Assignment Set 3.

d. By part b. it suffices to show $|U| = |V| \cdot |U : U \cap V|$. In order to see this, let $\{u_1, \ldots, u_n\}$ be a system of representatives of the cosets of $U \cap V$ in $U$, in particular $n = |U : U \cap V|$.

***Claim*:** $U \cdot V = \coprod_{i=1}^{n} u_i \cdot V$.

Let's show first that the union on the right hand side is disjoint. For that let's suppose that we have $v, w \in V$ such that $u_i \cdot v = u_j \cdot w \in u_i \cdot V \cap u_j \cdot V$ with $i \neq j$. Then $u_j^{-1} \cdot u_i = w \cdot v^{-1} \in U \cap V$, thus

$$u_i = u_j \cdot \left(w \cdot v^{-1}\right) \in u_j \cdot (U \cap V).$$

This, however, is a contradiction to the fact that the cosets $u_i \cdot (U \cap V)$ and $u_j \cdot (U \cap V)$ have no intersection.
We now show that indeed $U \cdot V = \bigcup_{i=1}^{n} u_i \cdot V$. Since $u_i \in U$, we have of course $\bigcup_{i=1}^{n} u_i \cdot V \subseteq U \cdot V$. Let now $u \in U$ be arbitrary. Then there is some $i \in \{1, \ldots, n\}$ such that $u \in u_i \cdot (U \cap V)$. Hence there is some $v \in U \cap V$ such that $u = u_i \cdot v$, and thus $u \cdot V = u_i \cdot v \cdot V = u_i \cdot V$. But this implies $U \cdot V \subseteq \bigcup_{i=1}^{n} u_i \cdot V$.

Having proved the claim, we deduce at once

$$|U| = \sum_{i=1}^{n} |u_i \cdot V| = n \cdot |V| = |U : U \cap V| \cdot |V| = \frac{|U| \cdot |V|}{|U \cap V|}.$$

□

**1.12 Remark** a. If $U \leq G$, then $|U| \,\Big|\, |G|$!

b. However, if $d \,\Big|\, |G|$, then there is not necessarily a subgroup of $G$ of order $d$. E. g. $d = 6$ and $G = \mathbb{A}_4$.

C) Normal Subgroups

**1.13 Definition and Proposition**

Let $(G, \cdot)$ be a group. A subgroup $N \leq G$ is called a *normal subgroup* of $G$ if and only if one of the following equivalent properties is fulfilled:

a. $n^g := g \cdot n \cdot g^{-1} \in N$ for all $n \in N$, $g \in G$.

b. $g \cdot N \cdot g^{-1} = N$ for all $g \in G$.

c. $g \cdot N = N \cdot g$ for all $g \in G$.

d. $(g \cdot N) \cdot (h \cdot N) = (g \cdot h) \cdot N$ for all $g, h \in G$.

We denote this by $(N, \cdot) \trianglelefteq (G, \cdot)$ or simply by $N \trianglelefteq G$.

**Proof:** a. $\Longrightarrow$ b.: By the assumption we have $g \cdot N \cdot g^{-1} \subseteq N$ for any $g \in G$. Let's now fix an arbitrary $g \in G$ and apply this inclusion to $g^{-1}$. We then get

$$g^{-1} \cdot N \cdot \left(g^{-1}\right)^{-1} \subseteq N,$$

and thus

$$N = e_G \cdot N \cdot e_G = g \cdot g^{-1} \cdot N \cdot \left(g^{-1}\right)^{-1} \cdot g^{-1} \subseteq g \cdot N \cdot g^{-1} \subseteq N.$$

This, however, implies $g \cdot N \cdot g^{-1} = N$.

b. $\Longrightarrow$ c.: Multiplying the equation $g \cdot N \cdot g^{-1} = N$ by $g$ on the desired equality.

c. $\Longrightarrow$ d.: Note that $N \cdot N = \{n_1 \cdot n_2 \mid n_1, n_2 \in N\} = N$, since $e_G \in N$! We thus get for $g, h \in G$

$$(gN) \cdot (hN) = (Ng) \cdot (hN) = N \cdot (gh) \cdot N = (gh) \cdot N \cdot N = ghN.$$

d. $\Longrightarrow$ a.: Let $g \in G$ and $n \in N$ be given, then

$$g \cdot n \cdot g^{-1} = g \cdot n \cdot g^{-1} \cdot e_G \in gN \cdot g^{-1}N = g \cdot g^{-1} \cdot N = e_G \cdot N = N.$$

□

**1.14 Example** a. The trivial subgroups $\mathbb{1}$ and $G$ of a group $(G, \cdot)$ are always normal subgroups.

b. If $(G, \cdot)$ is abelian, then every subgroup is a normal subgroup.

c. $\langle (1\ 2) \rangle$ is not a normal subgroup of $\mathbb{S}_3$ by Example 1.10 b., while $\langle (1\ 2\ 3) \rangle \trianglelefteq \mathbb{S}_3$ by Part d.
However, $(1\ 2) \circ (1\ 2\ 3) \circ (1\ 2)^{-1} = (1\ 3\ 2) \neq (1\ 2\ 3)$. Thus, $gNg^{-1} = N$ does not imply $gng^{-1} = n$ for all $n \in N$!

d. $\mathbb{A}_n \trianglelefteq \mathbb{S}_n$ by Proposition 1.15. For this note that for $n \geq 2$ we can write $\mathbb{S}_n = \mathbb{A}_n \cup (1\ 2) \circ \mathbb{A}_n$, where the first set contains all even permutations and the second one contains all odd ones.

**1.15 Proposition**
*Let* $(G, \cdot)$ *be a group,* $N \leq G$ *with* $|G : N| = 2$*, then* $N \trianglelefteq G$.

**Proof:** Let $g \in G$.

***1st Case*:** $gN = N$. Then $g = g \cdot e \in N$, and hence $Ng = N = gN$.

***2nd Case*:** $gN \neq N$. Then $g \notin N$, and hence $Ng \neq N$. However, since the index is two, the complement $G \setminus N$ of $N$ in $G$ must be a right and left coset. Hence $gN = G \setminus N = Ng$.

□

**1.16 Proposition**
*Let* $(G, \cdot)$ *be a group,* $N, N_1, N_2 \trianglelefteq G$, $U \leq G$.

a. $N \cdot U \leq G$.

b. $N_1 \cdot N_2 \trianglelefteq G$.

c. $N \cap U \trianglelefteq U$.

d. $N_1 \cap N_2 \trianglelefteq G$.

e. *If* $N_1 \cap N_2 = \mathbb{1}$*, then* $n_1 \cdot n_2 = n_2 \cdot n_1$ *for all* $n_i \in N_i$.

**Proof:** a. Since $N \trianglelefteq G$, we have $N \cdot u = u \cdot N$ for all $u \in U$, and hence

$$N \cdot U = \bigcup_{u \in U} N \cdot u = \bigcup_{u \in U} u \cdot N = U \cdot N.$$

Thus $N \cdot U \leq G$ by Proposition 1.7.

b. By Part a. $N_1 \cdot N_2 \leq G$, it thus remains to check one of the conditions for normality. Let $g \in G$. Taking into account, that $N_1$ and $N_2$ are normal, we get

$$g \cdot N_1 \cdot N_2 = N_1 \cdot g \cdot N_2 = N_1 \cdot N_2 \cdot g.$$

c. This is Exercise 3 on Assignment Set 3.

d. This is Exercise 3 on Assignment Set 3.

e. Let $n_i \in N_i$ for $i = 1, 2$ be given. Since $N_1$ and $N_2$ are normal subgroups, we have $n_1 \cdot n_2 \cdot n_1^{-1} \in N_2$ and $n_2 \cdot n_1^{-1} \cdot n_2^{-1} \in N_1$. But then

$$n_1 \cdot n_2 \cdot n_1^{-1} \cdot n_2^{-1} \in N_1 \cap N_2 = \{e_G\}.$$

Hence, $n_1 \cdot n_2 \cdot n_1^{-1} \cdot n_2^{-1} = e_G$, which implies $n_1 \cdot n_2 = n_2 \cdot n_1$.

□

**1.17 Definition and Proposition**
Let $(G, \cdot)$ be a group, $N \trianglelefteq G$. We denote by $G/N = \{gN \mid g \in G\}$ the set of (left) cosets of $N$ in $G$. We then define

$$\cdot : G/N \times G/N \to G/N : (gN, hN) \mapsto (gN) \cdot (hN) = ghN,$$

where the last equality is due to Proposition 1.13.
Then $(G/N, \cdot)$ is a group, the so called *quotient group* of $G$ by $N$.
The neutral element $e_{G/N}$ is just the coset $N = e_G N$, and the inverse of $gN$ is $g^{-1}N$.
If $(G, \cdot)$ is abelian, then $(G/N, \cdot)$ is abelian as well.

**Proof:** Note that the multiplication is well-defined by Part d. in Definition 1.13. Moreover, since $N = e_G N \in G/N$ is always a coset, $G/N$ is a non-empty set. It thus remains to verify the three group axioms.
Let $gN, hN, kN \in G/N$ be given. Then the associativity follows from the associativity of the multiplication in $G$:

$$\big(gN \cdot hN\big) \cdot kN = ghN \cdot kN = \big((gh) \cdot k\big) \cdot N = \big(g \cdot (hk)\big) \cdot N = gN \cdot hkN = gN \cdot (hN \cdot kN).$$

The coset $N = e_G N$ acts as neutral element:

$$e_G N \cdot gN = (e_G \cdot g) \cdot N = gN.$$

And for $gN \in G/N$ the inverse element is just $g^{-1}N$:

$$g^{-1}N \cdot gN = \big(g^{-1} \cdot g\big) \cdot N = e_G N.$$

If $G$ was abelian, then for $gN, hN \in G/N$ we have

$$gN \cdot hN = ghN = hgN = hN \cdot gN.$$

□

**1.18 Example** a. $\mathbb{Z}/n\mathbb{Z} = \{0 + n\mathbb{Z}, 1 + n\mathbb{Z}, \ldots, (n-1) + n\mathbb{Z}\}$, and we usually write $\overline{a}$ instead of $a + n\mathbb{Z}$ if no ambiguity can occur.
E. g. $(3 + 5\mathbb{Z}) + (4 + 5\mathbb{Z}) = 7 + 5\mathbb{Z} = 2 + 5\mathbb{Z}$, since $7 = 2 + 5 \cdot 1 \equiv 2 (\mathrm{mod}\ 5)$.

b. $\mathbb{S}_3/\mathbb{A}_3 = \big\{\mathbb{A}_3, (1\ 2) \circ \mathbb{A}_3\big\}$.

**1.19 Remark**
Let $(G, \cdot)$ be a group, $N \trianglelefteq G$. Then there is one-to-one correspondence between the subgroups $G/N$ and the subgroups of $G$ containing $N$ given by

$$\{U \leq G \mid N \subseteq U\} \longrightarrow \big\{\overline{U} \leq G/N\big\} : U \mapsto U/N.$$

Under this correspondence the normal subgroups of $G/N$ correspond precisely to the normal subgroups of $G$ containing $N$.

**Proof:** Proving this remark comes basically down to showing that, given $\overline{U} \leq G/N$, the set $\big\{u \in G \;\big|\; uN \in \overline{U}\big\}$ is a subgroup of $G$, containing $N$, and that it is normal, when $\overline{U}$ is normal. This establishes the inverse of the above map. We leave the details to the reader. □

D) HOMOMORPHISMS

**1.20 Definition**
Let $(G, \cdot)$ and $(H, \circ)$ be groups. A map $\varphi : G \to H$ is called a *(group-)homomorphism* if and only if for all $g, g' \in G$ we have $\varphi\big(g \cdot g'\big) = \varphi(g) \circ \varphi\big(g'\big)$.
If, moreover, $\varphi$ is injective / surjective / bijective, then we call $\varphi$ a *monomorphism / epimorphism / isomorphism.*

If $(H, \circ) = (G, \cdot)$, then the homomorphisms are also called *endomorphisms* and the isomorphisms are also called *automorphisms*.
We denote by $\mathrm{Hom}(G, H)$ the set of all homomorphisms from $G$ to $H$, and by $\mathrm{Aut}(G)$ the set of all automorphisms of $G$.
We say that $(G, \cdot)$ and $(H, \circ)$ are *isomorphic* if there is an isomorphism from $G$ to $H$, and we denote this by $(G, \cdot) \cong (H, \circ)$ or just $G \cong H$.

**1.21 Example** a. The map $\varphi : \mathbb{Z}/2\mathbb{Z} \to \{1, -1\}$ with $\varphi(0+2\mathbb{Z}) = 1$ and $\varphi(1+2\mathbb{Z})$ is an isomorphism from $(\mathbb{Z}/2\mathbb{Z}, +)$ to $\big(\{1, -1\}, \cdot\big)$.

b. $\mathrm{sgn} : (\mathbb{S}_n, \circ) \to \big(\{1, -1\}, \cdot\big) : \pi \mapsto \mathrm{sgn}(\pi)$ is an epimorphism if $n \geq 2$.
Note: $\mathrm{sgn}(a_1 \ \ldots \ a_k) = (-1)^{k-1}$.

c. $\det : \big(\mathrm{Gl}_n(K), \circ\big) \to \big(K \setminus \{0\}, \cdot\big) : A \mapsto \det(A)$ is an epimorphism by the determinant product rule.

d. $\exp : (\mathbb{R}, +) \to \big(\mathbb{R} \setminus \{0\}, \cdot\big) : x \mapsto e^x$ is a monomorphism by the exponential laws.

e. Let $(G, \cdot)$ be a group, and let $g \in G$ be some fixed element. We define a map

$$\alpha_g : G \to G : h \mapsto h^g = g \cdot h \cdot g^{-1}.$$

This map is an automorphism with inverse $\alpha_{g^{-1}}$, since $\alpha_g\big(hh'\big) = ghh'g^{-1} = ghgg^{-1}g'g^{-1} = \alpha_g(h) \cdot \alpha_g\big(h'\big)$.
Automorphisms of this type are called *inner automorphisms*. We denote by $\mathrm{Inn}(G)$ the set of all inner automorphisms of $G$.

f. Let $(G, \cdot)$ be a group, $N \trianglelefteq G$. The map $\nu : G \to G/N : g \mapsto gN$ is called the *quotient map* onto $G/N$ and is an epimorphism.

**1.22 Proposition**
*Let* $\alpha \in \mathrm{Hom}(G, H)$ *and* $\beta \in \mathrm{Hom}(H, K)$, *where* $(G, \cdot)$, $(H, *)$ *and* $(K, \diamond)$ *are groups.*

a. $\alpha(e_G) = e_H$ *and* $\alpha\big(g^{-1}\big) = \big(\alpha(g)\big)^{-1}$.

b. $\mathrm{Im}(\alpha) := \alpha(G) \leq G$ *and is called the* image *of* $\alpha$.

c. $\mathrm{Ker}(\alpha) := \alpha^{-1}(e_H) = \big\{g \in G \mid \alpha(g) = e_H\big\} \trianglelefteq G$ *and is called the* kernel *of* $\alpha$.

d. $\alpha$ *is a monomorphism if and only if* $\mathrm{Ker}(\alpha) = \{e_G\}$.

e. $\mathrm{Ker}(\nu : G \to G/N) = N$ *for* $N \trianglelefteq G$.

f. $\beta \circ \alpha \in \mathrm{Hom}(G, K)$.

g. *If* $\alpha$ *is bijective, then* $\alpha^{-1} \in \mathrm{Hom}(H, G)$.
*In particular,* $\big(\mathrm{Aut}(G), \circ\big)$ *is a subgroup of* $\big(\mathrm{Sym}(G), \circ\big)$.

**Proof:** a. Note

$$e_H * \alpha(e_G) = \alpha(e_G) = \alpha(e_G \cdot e_G) = \alpha(e_G) * \alpha(e_G),$$

and by the cancellation law we have $e_H = \alpha(e_G)$. Moreover, for $g \in G$ we then get

$$\alpha\big(g^{-1}\big) * \alpha(g) = \alpha\big(g^{-1} \cdot g\big) = \alpha(e_G) = e_H = \alpha(g)^{-1} * \alpha(g)$$

and applying the cancellation law once more we end up with $\alpha(g^{-1}) = \alpha(g)^{-1}$.

b. This is Exercise 1 on Assignment Set 4.

c. We note first that by $\alpha(e_G) = e_H$ we get $e_G \in \mathrm{Ker}(\alpha)$, so that the kernel is non-empty. Moreover, for $g, g' \in \mathrm{Ker}(\alpha)$ we have that $\alpha(g \cdot g') = \alpha(g) * \alpha(g') = e_H * e_H = e_H$, so that $g \cdot g' \in \mathrm{Ker}(\alpha)$. And $\alpha(g^{-1}) = \alpha(g)^{-1} = e_H^{-1} = e_H$, which implies that $g^{-1} \in \mathrm{Ker}(\alpha)$. Hence $\mathrm{Ker}(\alpha)$ is a subgroup of $G$. It remains to show that it satisfies the normality condition. Let $n \in \mathrm{Ker}(\alpha)$ and $g \in G$. Then

$$\alpha(g \cdot n \cdot g^{-1}) = \alpha(g) * \alpha(n) * \alpha(g^{-1}) = \alpha(g) * e_H * \alpha(g)^{-1} = e_H,$$

and therefore, $g \cdot n \cdot g^{-1} \in \mathrm{Ker}(\alpha)$.

d. Let's first suppose that $\alpha$ is injective. This implies that the kernel of $\alpha$, which is the preimage of $e_H$, contains at most one element. However, by Part a. $e_G \in \mathrm{Ker}(\alpha)$, hence $\mathrm{Ker}(\alpha) = \{e_G\}$.
Suppose now, that $\mathrm{Ker}(\alpha) = \{e_G\}$, and let $g, g' \in G$ such that $\alpha(g) = \alpha(g')$. We have to show that $g = g'$. By assumption we have

$$e_H = \alpha(g)^{-1} * \alpha(g') = \alpha(g^{-1}) * \alpha(g') = \alpha(g^{-1} \cdot g'),$$

which implies that $g^{-1} \cdot g' \in \mathrm{Ker}(\alpha) = \{e_G\}$. Hence $g^{-1} \cdot g' = e_G$, and thus $g = g'$.

e. $g \in \mathrm{Ker}(\nu)$ if and only if $gN = N$ if and only if $g \in N$.

f. Let $g, g' \in G$ be given.

$$\begin{aligned}(\beta \circ \alpha)(g \cdot g') = \beta\big(\alpha(g \cdot g')\big) &= \beta\big(\alpha(g) * \alpha(g')\big) \\ &= \beta\big(\alpha(g)\big) \diamond \beta\big(\alpha(g')\big) = (\beta \circ \alpha)(g) \diamond (\beta \circ \alpha)(g').\end{aligned}$$

g. This is Exercise 1 on Assignment Set 4.

□

**1.23 Theorem (Homomorphismtheorem)**
*Let $\alpha \in \mathrm{Hom}(G, H)$, then the map*

$$G/\mathrm{Ker}(\alpha) \to \mathrm{Im}(\alpha) : g\,\mathrm{Ker}(\alpha) \mapsto \alpha(g)$$

*is welldefined and an isomorphism.*
*In particular,* $G/\mathrm{Ker}(\alpha) \cong \mathrm{Im}(\alpha)$.

**Proof:** Let's do the proof in several steps.

***Step 1*:** $\overline{\alpha}$ is well-defined.

Let $g\,\mathrm{Ker}(\alpha) = g'\,\mathrm{Ker}(\alpha)$. We have to show that $\alpha(g) = \alpha(g')$, that is, $\alpha$ does not depend on the particular representative of the coset. By assumption we have $g^{-1} \cdot g' \in \mathrm{Ker}(\alpha)$. Hence

$$e_H = \alpha(g^{-1} \cdot g') = \alpha(g)^{-1} * \alpha(g'),$$

and thus $\alpha(g) = \alpha(g')$.

***Step 2*:** $\overline{\alpha}$ is a homomorphism.

Let $g\,\mathrm{Ker}(\alpha), g'\,\mathrm{Ker}(\alpha) \in G/\mathrm{Ker}(\alpha)$. Then

$$\begin{aligned}\overline{\alpha}\big(g\,\mathrm{Ker}(\alpha)\cdot g'\,\mathrm{Ker}(\alpha)\big) &= \overline{\alpha}\big(gg'\,\mathrm{Ker}(\alpha)\big) = \alpha\big(gg'\big)\\ &= \alpha(g) * \alpha\big(g'\big) = \overline{\alpha}\big(g\,\mathrm{Ker}(\alpha)\big)\cdot\overline{\alpha}\big(g'\,\mathrm{Ker}(\alpha)\big).\end{aligned}$$

Hence $\overline{\alpha}$ is a homomorphism.

***Step 3*:** $\overline{\alpha}$ is surjective.

Let $h \in \mathrm{Im}(\alpha)$ be given. Then there is some $g \in G$ such that $\alpha(g) = h$. But then $\overline{\alpha}\big(g\,\mathrm{Ker}(\alpha)\big) = \alpha(g) = h$, and thus $\overline{\alpha}$ is surjective.

***Step 4*:** $\overline{\alpha}$ is injective.

Let $g\,\mathrm{Ker}(\alpha) \in \mathrm{Ker}\,\big(\overline{\alpha}\big)$. Then

$$e_H = \overline{\alpha}\big(g\,\mathrm{Ker}(\alpha)\big) = \alpha(g).$$

Hence, $g \in \mathrm{Ker}(\alpha)$, and thus $g\,\mathrm{Ker}(\alpha) = \mathrm{Ker}(\alpha)$ is the neutral element of $G/\mathrm{Ker}(\alpha)$. This implies $\mathrm{Ker}\,\big(\overline{\alpha}\big)$ consists only of the neutral element, and therefore $\overline{\alpha}$ must be injective by the previous proposition. □

**1.24 Theorem (Isomorphismtheorems)**
*Let* $(G,\cdot)$ *be a group,* $N, N', M \trianglelefteq G$ *such that* $M \subseteq N$.

a. $\big(N\cdot N'\big)/N' \cong N/\big(N\cap N'\big)$.

b. $(G/M)/(N/M) \cong G/N$.

**Proof:** a. This is Exercise 2 on Assignment Set 4.

b. We note that $N/M$ is actually a normal subgroup of $G/M$, so that the double quotient group on the left hand side makes sense. And we do the proof in a similar way, defining a map by

$$\beta : G/M \to G/N : gM \mapsto gN,$$

showing that this is an epimorphism with kernel $N/M$ and then applying the Homomorphism Theorem.

***Step 0*:** $\beta$ is well-defined.

Since we define the map $\beta$ via the choice of a (non-unique) representative of the coset, we have to show that $\beta$ is well-defined, i. e. that the definition is independent of the chosen representative. Let therefore $gM = g'M$, then $g^{-1}\cdot g' \in M \subseteq N$, and thus $= gN = g'N$, i. e. $gN$ does not depend on the representative of $gM$.

***Step 1*:** $\beta$ is a homomorphism.

Let $gM, g'M \in G/M$ be given. Then

$$\beta\big(gM\cdot g'M\big) = \beta\big(gg'M\big) = gg'N = gN\cdot g'N = \beta(gM)\cdot\beta\big(g'M\big).$$

***Step 3*:** $\beta$ is surjective.

Let $gN \in G/N$ be given. Then $gN = \beta(gM) \in \mathrm{Im}(\beta)$, so that $\beta$ is surjective.

***Step 4*:** $\mathrm{Ker}(\beta) = N/M$.

$gM \in \mathrm{Ker}(\beta)$ if and only if $gN = N$ if and only if $g \in N$ if and only if $gM \in N/M$.

□

**1.25 Example** a. $\mathrm{Ker}(\mathrm{sgn}) = \mathbb{A}_n \trianglelefteq \mathbb{S}_n$ and $(\mathbb{S}_n/\mathbb{A}_n, \circ) \cong (\mathbb{Z}/2\mathbb{Z}, +)$.

b. $\mathrm{Sl}_n(K) := \mathrm{Ker}(\det) \trianglelefteq \mathrm{Gl}_n(K)$ and $\big(\mathrm{Gl}_n(K)/\mathrm{Sl}_n(K), \circ\big) \cong \big(K \setminus \{0\}, \cdot\big)$.

c. Consider the group$(\mathbb{Z}, +)$ and normal subgroups $N = n\mathbb{Z}$, $N' = n'\mathbb{Z}$. We then deduce from the Isomorphism Theorems:

$$\mathbb{Z}/\tfrac{n'}{\mathrm{hcf}(n,n')}\mathbb{Z} \cong \mathrm{hcf}(n,n')\mathbb{Z}/n'\mathbb{Z} = (n\mathbb{Z} + n'\mathbb{Z})/n'\mathbb{Z}$$
$$\cong n\mathbb{Z}/(n\mathbb{Z} \cap n'\mathbb{Z}) = n\mathbb{Z}/\mathrm{lcm}(n,n')\mathbb{Z} \cong \mathbb{Z}/\tfrac{\mathrm{lcm}(n,n')}{n}\mathbb{Z},$$

which corresponds to the fact that $n \cdot n' = \mathrm{hcf}\big(n,n'\big) \cdot \mathrm{lcm}\big(n,n'\big)$.

**1.26 Theorem (of Cayley)**
*Let* $(G,\cdot)$ *be a finite group of order* $n$*, then* $G$ *is isomorphic to a subgroup of* $(\mathbb{S}_n, \circ)$.

**Proof:** We first note that by Exercise 3 on Assignment Set 4 the groups $(\mathbb{S}_n, \circ)$ and $\big(\mathrm{Sym}(G), \circ\big)$ are isomorphic, so that it suffices to show that $G$ is actually isomorphic to a subgroup of the latter one. We define a map

$$\lambda : G \to \mathrm{Sym}(G) : g \mapsto \lambda_g,$$

where $\lambda_g : G \to G : h \mapsto g \cdot h$. Note that $\lambda_g$ is actually a permutation of $G$ with inverse $\lambda_{g^{-1}}$.

***Show*:** $\lambda$ is a monomorphism.

For $g, g' \in G$ we have

$$\lambda_{g\cdot g'}(h) = \big(g \cdot g'\big) \cdot h = g \cdot \big(g' \cdot h\big) = \lambda_g\big(\lambda_{g'}(h)\big) = (\lambda_g \circ \lambda'_g)(h)$$

for all $h \in G$, which implies $\lambda_{g\cdot g'} = \lambda_g \circ \lambda_{g'}$. But then

$$\lambda\big(g \cdot g'\big) = \lambda_{g\cdot g'} = \lambda_g \circ \lambda_{g'} = \lambda(g) \circ \lambda\big(g'\big)$$

and $\lambda$ is a homomorphism. It remains to show that $\lambda$ is injective, i. e. that the kernel of $\lambda$ consists only of the neutral element $e_G$.
$g \in \mathrm{Ker}(\lambda)$ if and only if $\lambda_g = \lambda(g) = \mathrm{id}_G$. However, the multiplication by $g$ on the left is the identity if and only if $g = e_G$. Thus $\mathrm{Ker}(\lambda) = \{e_G\}$.
Knowing that $\lambda$ is a monomorphism the Homomorphism Theorem gives

$$G \cong G/\{e_G\} = G/\mathrm{Ker}(\lambda) \cong \mathrm{Im}(\lambda) \leq \mathrm{Sym}(G).$$

□

This theorem says basically, that it would be sufficient to study subgroups of the symmetric groups $(\mathbb{S}_n, \circ)$, $n \in \mathbb{N}$, in order to get to know all possible finite groups up to isomorphism. This sounds in a certain sense very promising. However, the fact that the symmetric groups contain so much information is reflected by the fact that they are very complicated as well. The order of $\mathbb{S}_n$ is $n!$, so that already $\mathbb{S}_{10}$

has 3628800 elements. If we wanted to use this approach to study groups of order 10, we would have to cope with a very complicated object, while general methods of group theory allow us to show that there are, up to isomorphism, only two quite simple groups of order 10.

## 2 Cyclic Groups

In this paragraph we would like to obtain a good understanding of the structure of the simplest type of groups, namely those who can be generated by a single element.

**2.1 Definition**
A group $(G, \cdot)$ is said to be *cyclic* if and only if there is a $g \in G$ such that $G = \langle g \rangle$. We call $g$ a *generator* of $G$.
If $(G, \cdot)$ is any group, we then call $o(g) := |\langle g \rangle| = \min\{n > 0 \mid g^n = e_G\}$ the *order* of $g$, and $o(g)$ divides every integer $n$ with $g^n = e_G$. (Cf. Exercise 1 on Assignment Set 2.)

**2.2 Theorem (Classification of Cyclic Groups)**
*Let $(G, \cdot)$ be a cyclic group.*

a. *If $|G| = \infty$, then $(G, \cdot) \cong (\mathbb{Z}, +)$.*

b. *If $|G| = n < \infty$, then $(G, \cdot) \cong (\mathbb{Z}/n\mathbb{Z}, +)$.*

**Proof:** Let $G = \langle g \rangle = \{g^z \mid z \in \mathbb{Z}\}$, where the last equality is due to Proposition 1.7. The map

$$\alpha : \mathbb{Z} \to G : z \mapsto g^z$$

is, due to the exponential laws, an epimorphism of groups. By the Homomorphism Theorem we thus have

$$\mathbb{Z}/\operatorname{Ker}(\alpha) \cong \operatorname{Im}(\alpha) = G.$$

Due to Example 1.6 there is an integer $n \geq 0$ such that $\operatorname{Ker}(\alpha) = n\mathbb{Z}$, so that $G \cong \mathbb{Z}/n\mathbb{Z}$. This implies the above statement, once we note that $0 \cdot \mathbb{Z} = \{0\}$ and $\mathbb{Z}/\{0\} \cong \mathbb{Z}$ in an obvious way. □

**2.3 Proposition (Subgroups of Cyclic Groups)**
*Let $(G, \cdot)$ be a cyclic group with generator $g$.*

a. *If $|G| = \infty$, then $U \leq G$ if and only if $\exists\, n \geq 0 : U = \langle g^n \rangle$.*

b. *If $|G| = n < \infty$, then $U \leq G$ if and only if $\exists\, m \mid n : U = \left\langle g^{\frac{n}{m}} \right\rangle$.*
*In particular, $G$ has for every divisor of $|G|$ precisely one subgroup of this order.*

**Proof:** Part a. is an immediate consequence of Theorem 2.2 and Example 1.6 b., where we classified the subgroups of $(\mathbb{Z}, +)$.
For Par b. we note that according to Remark 1.19 there is a one-to-one correspondence between the subgroups of $\mathbb{Z}/n\mathbb{Z}$ and the subgroups of $\mathbb{Z}$ which contain $n\mathbb{Z}$. However, $U \leq \mathbb{Z}$ with $n\mathbb{Z} \subseteq U$ if and only if $\exists\, m : U = m\mathbb{Z}$ and $n\mathbb{Z} \subseteq m\mathbb{Z}$ if and only if $\exists\, m : U = m\mathbb{Z}$ and $m \mid n$. This proves Part b. □

**2.4 Corollary**
*Every subgroup of a cyclic group is cyclic.*

**2.5 Proposition**
*If* $(G,\cdot)$ *is a group of prime order, then* $G$ *is cyclic.*

**Proof:** Let $e_G \neq g \in G$. Then $\mathbb{1} \neq \langle g\rangle \leq G$. By the Theorem of Lagrange we have

$$1 \;<\; |\langle g\rangle| \;\Big|\; |G|.$$

Since the latter is a prime number, this implies $|\langle g\rangle| = |G|$ and thus $\langle g\rangle = G$. □

**2.6 Proposition**
*Let* $(G,\cdot)$ *be a group,* $g \in G$ *with* $o(g) = n$ *and* $k \in \mathbb{Z}$. *Then* $o\big(g^k\big) = \frac{n}{\mathrm{hcf}(k,n)}$.

**Proof:** Recall that $\mathrm{lcm}(k,n) = k \cdot \frac{n}{\mathrm{hcf}(n,k)}$. We thus get by Exercise 1 on the Assignment Set 2

$$o\big(g^k\big) = \min\big\{a \geq 0 \;\big|\; g^{ka} = e_G\big\} = \min\big\{a \geq 0 \;\big|\; n \mid ka\big\}$$
$$= \min\big\{a \geq 0 \;\big|\; n \mid ka \;\text{ and }\; k \mid ka\big\} = \min\big\{a \geq 0 \;\big|\; \mathrm{lcm}(k,n) \mid ka\big\} = \frac{n}{\mathrm{hcf}(n,k)}.$$

□

**2.7 Proposition**
*Let* $(G,\cdot)$ *be a group, and let* $g,h \in G$ *such that* $gh = hg$ *and* $\mathrm{hcf}\big(o(g),o(h)\big) = 1$. *Then* $o(gh) = o(g)\cdot o(h)$.
*In particular,* $\langle g,h\rangle = \langle gh\rangle$ *is a cyclic group of order* $o(g)\cdot o(h)$.

**Proof:** Let $m = o(g)$, $n = o(h)$ and $k = o(gh)$.
Note that $(gh)^{mn} = \big(g^m\big)^n \cdot \big(h^n\big)^m = e_G$, since $gh = hg$. Hence, $k \mid mn$.
Moreover, $e_G = (gh)^k = g^k \cdot h^k$ implies that $g^k = \big(h^{-1}\big)^k$. Taking into account that $o(h) = o\big(h^{-1}\big)$, Proposition 2.6 gives

$$a := \frac{m}{\mathrm{hcf}(k,m)} = o\big(g^k\big) = o\Big(\big(h^{-1}\big)^k\Big) = \frac{n}{\mathrm{hcf}(k,n)}.$$

In particular, $a$ divides $m$ and $n$, hence $a \mid \mathrm{hcf}(m,n) = 1$, and thus $a = 1$, i. e. $g^k$ and $\big(h^{-1}\big)^k$ have order one. Thus $g^k = e_G = h^k$. This however, implies the order of $g$ and the order of $h$ divide $k$, hence their least common multiple divides $k$, i. e.

$$mn = \frac{mn}{\mathrm{hcf}(m,n)} = \mathrm{lcm}(m,n) \;\Big|\; k.$$

Thus we must have $mn = k$.
For the in particular part we note that by Proposition 1.7 and since $gh = hg$, we have

$$\langle g,h\rangle = \langle g\rangle \cdot \langle h\rangle.$$

Moreover, by the Product Formula in Theorem 1.11 we get

$$|\langle g,h\rangle| = \frac{|\langle g\rangle|\cdot|\langle h\rangle|}{|\langle g\rangle \cap \langle h\rangle|} \leq mn.$$

On the other hand $\langle gh\rangle \leq \langle g,h\rangle$ is a subgroup of order $mn$, thus $\langle gh\rangle = \langle g,h\rangle$ and $|\langle g,h\rangle| = mn$. □

**2.8 Proposition**
*Let* $G$ *be a finite subgroup of the multiplicative group* $(K^*, \cdot)$ *of a field* $K$*, then* $G$ *is cyclic.*
*In particular, for a prime* $p$ *the group* $\big(\mathbb{Z}/p\mathbb{Z} \setminus \{\overline{0}\}, \cdot\big)$ *is cyclic of order* $p-1$.

**Sketch of Proof:** Let $m = \operatorname{lcm}\big\{o(a) \;\big|\; a \in G\big\}$, then in particular $a^m = 1$ for all $a \in G$.
With the aid of Proposition 2.6 and of Proposition 2.7 we can find an element $g \in G$ such that $o(g) = m$. Hence $m = o(g) \leq |G|$.
Since $K$ is a field, the polynomial $f = x^m - 1$ has at most $m$ zeros in $K$. But since $a^m = 1$ for all $a \in G$, the elements of $G$ are zeros of $f$, and thus $|G| \leq m$.
We therefore have $|G| = m$, and $G = \langle g \rangle$ is cyclic. □

**2.9 Theorem (Automorphisms of Cyclic Groups)**
*Let* $(G, \cdot)$ *be a cyclic group of order* $n$ *with generator* $g$.
*Then* $\operatorname{Aut}(G) = \{\alpha_k \mid \operatorname{hcf}(k, n) = 1, 1 \leq k < n\}$*, where* $\alpha_k : G \to G : g^i \mapsto g^{ik}$.

**Proof:** Due to the exponential laws the maps $\alpha_k$ are actually group homomorphisms for all $k \in \mathbb{Z}$. Moreover, $\alpha_k$ is bijective if and only if it is surjective, since $G$ is finite. This is the case if and only if

$$n = |G| = |\operatorname{Im}(\alpha_k)| = \big|\langle \alpha_k(g) \rangle\big| = o\big(g^k\big) = \frac{n}{\operatorname{hcf}(k, n)}$$

by Proposition 2.6. Thus $\alpha_k$ is an automorphism if and only if $\operatorname{hcf}(k, n) = 1$.
It remains to show that any automorphism of $G$ is of this form. However, if $\beta : G \to G$ is any homomorphism, then $\beta$ is fixed once we know the image of the generator $g$, since any element of $G$ is a power of $g$. I. e. $\beta(g) = g^k$ for some $k$ implies $\beta = \alpha_k$. □

**2.10 Corollary**
*If* $|G| = p$ *has prime order, then* $\operatorname{Aut}(G)$ *is cyclic of order* $p-1$.
*In particular,* $\operatorname{inv}_G \in \operatorname{Aut}(G)$ *is the only automorphism of order* 2.

**Proof:** By Theorem 2.2 we may assume that $(G, \cdot) = (\mathbb{Z}/p\mathbb{Z}, +)$. Translating the result of Theorem 2.9 to the additive group gives $\operatorname{Aut}(\mathbb{Z}/p\mathbb{Z}) = \{\alpha_k \mid k = 1, \ldots, p-1\}$ with

$$\alpha_k : \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z} : \overline{a} \mapsto k \cdot \overline{a} = \overline{k} \cdot \overline{a}.$$

Thus $\alpha_k$ is just the multiplication with $\overline{k}$, and we have a natural identification

$$\big(\operatorname{Aut}(\mathbb{Z}/p\mathbb{Z}), \circ\big) \cong \big(\mathbb{Z}/p\mathbb{Z} \setminus \{\overline{0}\}, \cdot\big).$$

Hence, we are done by Proposition 2.8. □

# 3 Group Actions

**3.1 Definition and Proposition**

Let $(G, \cdot)$ be group, let $\Omega$ be a non-empty set and let $\alpha : G \to \mathrm{Sym}(\Omega)$ be a group homomorphism.

a. $\alpha$ is called an *action* of $G$ on $\Omega$.
Note that $\alpha(g) : \Omega \to \Omega$ is a permutation of $\Omega$ for all $g \in G$, in particular it is a map which may be evaluated at $\omega \in \Omega$.
We usually write $g \cdot \omega$ instead of $\alpha(g)(\omega)$ for $g \in G$ and $\omega \in \Omega$, if no ambiguity can arise.
Thus the fact that $\alpha$ is a group homomorphism translates to the rule

$$(g \cdot h) \cdot \omega = g \cdot (h \cdot \omega)$$

for $g, h \in G$ and $\omega \in \Omega$, and we have

$$e_G \cdot \omega = \omega.$$

b. Defining for $\omega, \omega' \in \Omega$

$$\omega \sim \omega' \iff \exists\, g \in G \,:\, g \cdot \omega = \omega'$$

gives an equivalence relation on $\Omega$. The equivalence class of $\omega$ is denoted by $\mathrm{orb}_G(\omega) = \{g \cdot \omega \mid g \in G\}$ and is called the *orbit* of $\omega$ under $G$.
Note, that in particular $\Omega$ is the disjoint union of the different orbits, i. e. there exist $\omega_i \in \Omega$, $i \in I$, such that

$$\Omega = \coprod_{i \in I} \mathrm{orb}_G(\omega_i).$$

c. $\mathrm{Stab}_G(\omega) = \{g \in G \mid g \cdot \omega = \omega\} \leq G$ is called the *stabiliser* of $\omega$.

d. If $\mathrm{Ker}(\alpha) = \mathbb{1}$, the action is said to be *faithful*.

e. If $\Omega = \mathrm{orb}_G(\omega)$ for some $\omega \in \Omega$, then the action is called *transitive*.

**Proof:** For Part b. we have to prove that $\sim$ is an equivalence relation. Since $\omega = e_G \cdot \omega$, we have $\omega \sim \omega$ and the relation is reflexive.
Suppose that $\omega \sim \omega'$, then there is a $g \in G$ such that $g \cdot \omega = \omega'$. Hence $\omega = e_G \cdot \omega = g^{-1} \cdot (g \cdot \omega) = g^{-1} \cdot \omega'$, and thus $\omega' \sim \omega$. The relation is therefore also symmetric.
Let $\omega \sim \omega'$ and $\omega' \sim \omega''$. Then there are $g, h \in G$ such that $g \cdot \omega = \omega'$ and $h \cdot \omega' = \omega''$. This implies $(gh) \cdot \omega = g \cdot (h \cdot \omega) = g \cdot \omega' = \omega''$ and $\omega \sim \omega''$. So finally the relation is transitive, and thus an equivalence relation.
It remains to show in Part c. that the stabiliser of $\omega$ is a subgroup of $G$. Since $e_G \cdot \omega = \omega$, the neutral element belongs ot $\mathrm{Stab}_G(\omega)$ and the latter is a non-empty set. Let now $g, h \in \mathrm{Stab}_G(\omega)$, then

$$(g \cdot h) \cdot \omega = g \cdot (h \cdot \omega) = g \cdot \omega = \omega \quad \text{and} \quad g^{-1} \cdot \omega = g^{-1} \cdot (g \cdot \omega) = e_G \cdot \omega = \omega$$

and hence $gh, g^{-1} \in \mathrm{Stab}_G(\omega)$ as required. □

**3.2 Example**

a. Consider the additive group of real numbers $(\mathbb{R}, +)$, the set $\mathbb{C}$ of complex numbers and the map

$$\alpha : \mathbb{R} \to \mathrm{Sym}(\mathbb{C}) : t \mapsto \big(\alpha(t) : \mathbb{C} \to \mathbb{C} : c \mapsto c \cdot e^{2t\pi i}\big).$$

Due to the exponential laws for complex numbers this map is a group homomorphism, which means that $\mathbb{R}$ acts on $\mathbb{C}$. More concretely, the real number $t$ acts on $\mathbb{C}$ by multiplication with $e^{2t\pi i}$, which is a rotation by the angle $2t\pi$.
The kernel of $\alpha$ is the set of numbers $t$ for which $\alpha(t) = \mathrm{id}_{\mathbb{C}}$, that is, for which the rotation by $2t\pi$ does not do anything. This is, of course, the case if and only if $t$ is an integer, so that $\mathrm{Ker}(\alpha) = \mathbb{Z}$. In particular, $\alpha$ is not faithful.
The orbit of $c \in \mathbb{C}$ is just

$$\mathrm{orb}_{\mathbb{R}}(c) = \big\{c \cdot e^{2t\pi i} \;\big|\; t \in \mathbb{R}\big\}$$

the circle of radius $|c|$ with the origin as centre, and the stabiliser of $c \neq 0$ is

$$\mathrm{Stab}_{\mathbb{R}}(c) = \big\{t \in \mathbb{R} \mid c \cdot e^{2t\pi i} = c\big\} = \big\{t \in \mathbb{R} \;\big|\; e^{2t\pi i} = 1\big\} = \mathbb{Z},$$

while for $c = 0$ we have

$$\mathrm{Stab}_{\mathbb{R}}(0) = \mathbb{R}.$$

b. Let $(G, \cdot)$ be any group, $k \geq 1$ and

$$\Omega = \big\{U \subseteq G \;\big|\; |U| = k\big\}.$$

We define a homomorphism

$$\lambda : G \to \mathrm{Sym}(\Omega) : g \mapsto \lambda_g$$

with $\lambda_g : \Omega \to \Omega : U \mapsto gU$. I. e. $\lambda(g) = \lambda_g$ is the left-multiplication by $g$, and we say for short $G$ *acts on* $\Omega$ *by left-multiplication.*
If $k < |G|$, then the action is faithful. For this just note that for a set $U$ of order $k$ containing $e_G$ but not $g$ we always have $U \neq gU$. This means that $\lambda_g \neq \mathrm{id}_\Omega$ and hence the kernel of $\lambda$ contains only the neutral element $e_G$.
If $k = |G| > 1$, then the action cannot be faithful, since $\Omega$ contains only one element, and thus $\mathrm{Sym}(\Omega)$ has order one, while $G$ has an order strictly greater than one.
Consider the special case of $U \in \Omega$ where $U$ is a *subgroup* of $G$. Then

$$\mathrm{orb}_G(U) = \{gU \mid g \in G\} = \text{set of left cosets of } U \text{ in } G.$$

Moreover,

$$\mathrm{Stab}_G(U) = \{g \in G \mid gU = U\} = U.$$

In particular

$$|\mathrm{orb}_G(U)| = |G : U| = |G : \mathrm{Stab}_G(U)|.$$

We will see in Theorem 3.3 that the last equality does not only hold by chance.

c. Using the notation of b. let $U \subseteq G$ with $|U| = k$ and let $H = \mathrm{Stab}_G(U)$. Then the group $H$ acts on the set $U$ by left-multiplication, i. e. the map

$$\mu : H \to \mathrm{Sym}(U) : h \mapsto \mu_h,$$

with $\mu_h : U \to U : u \mapsto hu$, is a group homomorphism. The important point for this is that $hu \in U$, since $h \in \mathrm{Stab}_G(U)$.
For $u \in U$ we find

$$\mathrm{orb}_H(u) = H \cdot u.$$

Since the action of $H$ divides the set $U$ into a disjoint union of orbits, there are $u_1, \ldots, u_r \in U$ such that

$$U = \coprod_{i=1}^{r} H \cdot u_i.$$

Knowing that $|H \cdot u_i| = |H|$ we get

$$|U| = r \cdot |H|,$$

which in particular means that $|H|$ divides $|U| = k!$ I. e. the stabiliser of $U \in \Omega$ in Part b. will be a subgroup of $G$ of an order which divides $k$.

d. Let $(G, \cdot)$ be a group and $A \subseteq G$ be a fixed subset. We then consider the set

$$\Omega = \left\{ A^g \;\middle|\; g \in G \right\},$$

where $A^g = gAg^{-1}$. The group homomorphism

$$\alpha : G \to \mathrm{Sym}(\Omega) : h \mapsto \alpha_h,$$

with $\alpha_h : \Omega \to \Omega : A^g \mapsto \left(A^g\right)^h = A^{hg}$, defines an action of $G$ on $\Omega$, which we call *conjugation*. $\Omega$ is called the *conjugacy class* of $A$.
The action is transitive, since $\Omega = \mathrm{orb}_G(A)$, and we call

$$N_G(A) := \mathrm{Stab}_G(A) = \left\{ g \in G \;\middle|\; A^g = A \right\}$$

the *normaliser* of $A$ in $G$.

### 3.3 Orbit Stabiliser Theorem

*Let $(G, \cdot)$ be a group acting on the* finite *set $\Omega$, and let $\omega \in \Omega$. Then*

$$|\mathrm{orb}_G(\omega)| = |G : \mathrm{Stab}_G(\omega)|.$$

*In particular, the order of an orbit always divides of the order of the group.*

**Proof:** Let $U = \mathrm{Stab}_G(\omega)$, and $\mathcal{M} = \{gU \mid g \in G\}$ be the set of left-cosets of $U$ in $G$. We then have to show that $|\mathrm{orb}_G(\omega)| = |\mathcal{M}|$, i. e. we have to find a bijection between the corresponding sets. Define

$$\gamma : \mathcal{M} \to \mathrm{orb}_G(\omega) : gU \mapsto g \cdot \omega.$$

Since the definition of $\gamma$ a priori depends on the chosen representative $g$ of the coset $gU$, we first have to show that $\gamma$ is well-defined, i. e. that it is independent of $g$. Suppose that $gU = hU$, then there is a $u \in U = \mathrm{Stab}_G(\omega)$ such that $g = hu$. Thus

$$g \cdot \omega = (h \cdot u) \cdot \omega = h \cdot (u \cdot \omega) = h \cdot \omega.$$

Hence, $\gamma$ is well-defined, and we may go on showing, that $\gamma$ is bijective.
Suppose that $gU, hU \in \mathcal{M}$ such that $\gamma(gU) = \gamma(hU)$, i. e. $g \cdot \omega = h \cdot \omega$. Then

$$\omega = e_G \cdot \omega = \left(g^{-1} \cdot g\right) \cdot \omega = g^{-1} \cdot (g \cdot \omega) = g^{-1} \cdot (h \cdot \omega) = \left(g^{-1} \cdot h\right) \cdot \omega.$$

Hence, $g^{-1} \cdot h \in \mathrm{Stab}_G(\omega) = U$, and thus $gU = hU$. This proves, $\gamma$ is injective.
Obviously, $\gamma$ is also surjective, since for $g \cdot \omega \in \mathrm{orb}_G(\omega)$ arbitrary, we have $g \cdot \omega = \gamma(gU)$.
This adds up to the fact

$$|\mathrm{orb}_G(\omega)| = |\mathcal{M}| = |G : \mathrm{Stab}_G(\omega)|.$$

□

**3.4 Corollary**
*Let $(G, \cdot)$ be a finite group, and $A \subseteq G$ any subset. Then*

$$|G : N_G(A)| = \left|\left\{A^g \;\middle|\; g \in G\right\}\right|,$$

*where the latter is the order of the conjugacy class of $A$ in $G$.*

**Proof:** By Exercise 3.2 d., the group $G$ acts transitively on the set $\Omega = \left\{A^g \;\middle|\; g \in G\right\}$ via conjugation, and $\Omega = \mathrm{orb}_G(A)$. Moreover, by definition the normaliser of $A$ in $G$ is just $\mathrm{Stab}_G(A) = N_G(A)$. Hence the statement follows from the Orbit Stabiliser Theorem 3.3. □

**3.5 Corollary**
*Let $(P, \cdot)$ be a group of order $p^n$ for some prime $p$, and suppose that $P$ acts on a finite set $\Omega$ with* $\mathrm{hcf}\left(|\Omega|, p\right) = 1$. *Then there is an $\omega \in \Omega$ such that*

$$\mathrm{orb}_P(\omega) = \{\omega\},$$

*i. e. $\omega$ is a fix point of the action.*

**Proof:** By the Orbit Stabiliser Theorem and the Theorem of Lagrange we have

$$|\mathrm{orb}_P(\omega)| = |P : \mathrm{Stab}_P(\omega)| \;\Big|\; |P| = p^n$$

for all $\omega \in \Omega$, and thus there is some integer $0 \leq m = m(\omega) \leq n$, depending on $\omega$, such that

$$|\mathrm{orb}_P(\omega)| = p^{m(\omega)}.$$

Suppose that $m(\omega) > 0$ for all $\omega \in \Omega$. We know that there are $\omega_1, \ldots, \omega_r$ such that

$$\Omega = \coprod_{i=1}^{r} \mathrm{orb}_P(\omega_i).$$

By assumption $p$ divides $|\mathrm{orb}_P(\omega_i)|$ for all $i = 1, \ldots, r$, thus it divides the sum of these numbers, i. e.

$$p \;\Big|\; \sum_{i=1}^{r} |\mathrm{orb}_P(\omega_i)| = |\Omega|.$$

This, however, is a contradiction to the fact that $\mathrm{hcf}\left(|\Omega|, p\right) = 1$.
Hence, there is at least one $\omega \in \Omega$, such that $m(\omega) = 0$, which is the same as saying the $|\mathrm{orb}_P(\omega)| = 1$, or that $\mathrm{orb}_P(\omega) = \{\omega\}$. □

# 4 The Theorem of Sylow

The Theorem of Lagrange was one of the highlights of this lecture so far, even though it was not hard to prove. When interested whether a certain subset $U$ of a group $G$ could be a subgroup, we may check first, if the number of elements in $U$ is a divisor of $G$. If $|U|$ does not divide $|G|$, then it cannot possibly be a subgroup, that's what we inherit from the Theorem of Lagrange.
Knowing that the order of a subgroup must divide the order of the group, it is quite natural to ask, whether for every divisor of $|G|$ there is a subgroup of that order in $G$? We know that this is true for cyclic groups and we will show later that it also holds for abelian groups. However, it does not hold in general, as the following example shows.

**4.1 Example**
The group $\mathbb{A}_4$ has no subgroup of order 6, even though its order is 12.

We postpone the proof for a moment, so that we can use the result of the next theorem, which is - so to say - a converse of the Theorem of Lagrange for powers of primes. We introduce the notation

$$N_G(k) = \left|\{U \leq G \mid |U| = k\}\right|$$

to denote the number of subgroups of $G$ of order $k$, when $G$ is a finite group.

**4.2 Theorem**
*Let $(G, \cdot)$ be a group of order $|G| = p^a \cdot m$ with $p$ a prime, $a \geq 0$ and $m > 0$. Then*

$$N_G(p^a) \equiv 1 \pmod{p}.$$

**Proof (due to Wieland):** Having introduced some notation we will do the proof in several steps. Define

$$\Omega = \{U \subseteq G \mid |U| = p^a\}.$$

Then $G$ acts on $\Omega$ by left-multiplication, as we have seen in Example 3.2 b., i.e.

$$\lambda : G \to \mathrm{Sym}(\Omega) : g \mapsto (\lambda_g : \Omega \to \Omega : U \mapsto gU)$$

is a group homomorphism.

***Step 1*:** $\forall\, U \in \Omega\ \exists\, 0 \leq b = b(U) \leq a\ :\ |\mathrm{Stab}_G(U)| = p^b$.

We note that by Example 3.2 c., the group $H = \mathrm{Stab}_G(U) \leq G$ acts on the set $U$ by left-multiplication. Moreover, we have shown there that there is a number $r$ such that

$$r \cdot |H| = |U| = p^a.$$

Hence, $|H| = p^b$ for some $0 \leq b \leq a$, of course depending on $U$.

***Step 2*:** For all $U \in \Omega$ we have either $|\mathrm{orb}_G(U)| = m$ or $|\mathrm{orb}_G(U)| \equiv 0 \pmod{pm}$.

By the Orbit Stabiliser Theorem and Step 1 be have

$$|\mathrm{orb}_G(U)| = |G : \mathrm{Stab}_G(U)| = \frac{p^a \cdot m}{p^{b(U)}} = p^{a-b(U)} \cdot m.$$

So, if $b(U) = a$, then $|\mathrm{orb}_G(U)| = m$, otherwise $|\mathrm{orb}_G(U)| \equiv 0 \pmod{pm}$.

***Step 3*:** $|\Omega| \equiv l \cdot m \pmod{pm}$, where $l = \left|\left\{ \mathrm{orb}_G(U) \;\middle|\; |\mathrm{orb}_G(U)| = m \right\}\right|$.

Since $G$ acts on $\Omega$, there are $U_1, \ldots, U_n \in \Omega$ such that

$$\Omega = \coprod_{i=1}^{n} \mathrm{orb}_G(U_i).$$

But then by Step 2

$$|\Omega| = \sum_{i=1}^{n} |\mathrm{orb}_G(U_i)| \equiv l \cdot m \pmod{pm}.$$

***Step 4*:** $l = N_G\left(p^a\right)$.

We set $\mathcal{M} = \{U \in \Omega \mid U \leq G\}$ and $\mathcal{N} = \left\{ \mathrm{orb}_G(U) \;\middle|\; |\mathrm{orb}_G(U)| = m \right\}$. Then $l = |\mathcal{N}|$ and $N_G\left(p^a\right) = |\mathcal{M}|$. It thus suffices to find a bijection between $\mathcal{M}$ and $\mathcal{N}$. We define

$$\beta : \mathcal{M} \to \mathcal{N} : U \mapsto \mathrm{orb}_G(U) = \{g \cdot U \mid g \in G\}$$

and we claim that $\beta$ is a bijection.

Note first of all, that for $U \in \mathcal{M}$ we have already shown in Example 3.2 b. that

$$|\mathrm{orb}_G(U)| = |G : U| = \frac{|G|}{|U|} = \frac{p^a \cdot m}{p^a} = m,$$

so that $\beta$ actually takes its values in $\mathcal{N}$!

Let's now show the injectivity of $\beta$. For this suppose that we have $U, U' \in \mathcal{M}$ such that $\mathrm{orb}_G(U) = \beta(U) = \beta\left(U'\right) = \mathrm{orb}_G\left(U'\right)$. Then there is a $g \in G$ such that $U = gU'$. However, since $e_G \in U = gU'$, we see that $g^{-1} \in U'$, and hence $g \in U'$ as well. But then $U' = gU' = U$.

It remains to show the surjectivity of $\beta$. For this let $\mathrm{orb}_G(U) \in \mathcal{N}$ be given. As we have seen in Example 3.2 c., the group $H = \mathrm{Stab}_G(U)$ acts on $U$ by left-multiplication and we have $u_1, \ldots, u_k \in U$ such that

$$U = \coprod_{i=1}^{k} \mathrm{orb}_G(u_i) = \coprod_{i=1}^{k} H \cdot u_i.$$

Note, that due to the Orbit Stabiliser Theorem and the Theorem of Lagrange we have

$$|H| = \frac{|G|}{|G : H|} = \frac{|G|}{|\mathrm{orb}_G(U)|} = \frac{p^a \cdot m}{m} = p^a.$$

This implies

$$p^a = |U| = k \cdot |H| = k \cdot p^a,$$

which is only possible if $k = 1$ and $U = Hu_1$. We set now $U' = u_1^{-1} H u_1$ which is a subgroup of $G$ of order $p^a$, hence is an element of $\mathcal{M}$. And that way we get

$$\begin{aligned} \mathrm{orb}_G(U) = \{gHu_1 \mid g \in G\} &= \{gu_1^{-1} H u_1 \mid g \in G\} \\ &= \left\{gU' \;\middle|\; g \in G\right\} = \mathrm{orb}_G\left(U'\right) = \beta\left(U'\right) \in \mathrm{Im}(\beta). \end{aligned}$$

Hence, $\beta$ is surjective.

***Step 5*:** $|\Omega| \equiv N_G\left(p^a\right) \cdot m \pmod{pm}$.

Just combine Step 3 and Step 4.

***Step 6*:** The trick of Shaw!

Note, that the number $|\Omega|$ does not depend on the group structure of $G$! It is the number of subsets with $p^a$ elements of a set with $p^a \cdot m$ elements. This number is for *any* group of order $p^a \cdot m$ the same. In particular, we may apply Step 5 to the given group $G$ as well as to the cyclic group $\big(\mathbb{Z}/p^a m\mathbb{Z}, +\big)$ and we find

$$N_G\big(p^a\big) \cdot m \equiv |\Omega| \equiv N_{\mathbb{Z}/p^a m\mathbb{Z}}\big(p^a\big) \cdot m = m \pmod{pm},$$

since this cyclic group has precisely one subgroup of order $p^a$ by Proposition 2.3. This implies

$$pm \,\Big|\, \big(N_G\big(p^a\big) - 1\big) \cdot m,$$

and hence $p \,\Big|\, N_G\big(p^a\big) - 1$, which is the same as saying

$$N_G\big(p^a\big) \equiv 1 \pmod{p}.$$

□

**Proof of Example 4.1:** Suppose $U \leq \mathbb{A}_4$ and $|U| = 6$.
We first note that $\mathbb{A}_4$ contains the Kleinian subgroup $\mathbb{K}_4$ and, apart from that, the eight 3-cycles in $\mathbb{S}_4$.
Moreover, we note that any subgroup $P \leq U$ of order 3 must be of the form

$$P = \big\{(1), (a\ b\ c), (a\ c\ b)\big\}$$

for some $\{a, b, c\} \subset \{1, 2, 3, 4\}$, since it is cyclic as a group of prime order.
By Theorem 4.2 $N_U(3) \equiv 1 \pmod 3$, and since $U$ cannot possibly contain 8 different 3-cycles, $N_U(3)$ cannot be 4 or even larger. Hence, $N_U(3) = 1$ and $U$ has a unique subgroup $P \leq U$ of order 3.
In particular, apart from the elements in $P$ the subgroup $U$ can only contain elements of order 2, since $\mathbb{A}_4$ only contains elements of order 1, 2 and 3. Thus we have

$$U = \big\{(1), (1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3), (a\ b\ c), (a\ c\ b)\big\}.$$

That implies that $\mathbb{K}_4 \leq U$, which is a contradiction to the Theorem of Lagrange, since $4 \nmid 6$. □

An immediate consequence of Theorem 4.2 is the Theorem of Cauchy.

**4.3 Corollary (Theorem of Cauchy)**
*Let* $(G, \cdot)$ *be a finite group such that* $p^a \,\Big|\, |G|$, *then* $G$ *has a subgroup of order* $p^a$.
*In particular, if* $p \,\Big|\, |G|$, *then* $G$ *contains an element of order* $p$.

**4.4 Corollary**
*Let* $(G, \cdot)$ *be a finite* abelian *group and* $d \,\Big|\, |G|$, *then* $G$ *has a subgroup of order* $d$.

**Proof:** We do the proof by induction on $d$, where the case $d = 1$ is obviously satisfied by the trivial subgroup $\mathbb{1}$.
Let's therefore assume that $d > 1$. If $d$ is a power of a prime number, we are done by the Theorem of Cauchy. Otherwise, there is a prime number $p$, $a > 0$ and $m > 0$ such that $d = p^a \cdot m$, where $p \nmid m$. Since $m < d$ and $p^a < d$, by induction there are subgroups $N_1, N_2 \leq G$ such that $|N_1| = m$ and $|N_2| = p^a$.

However, since $G$ is abelian, $N_1$ and $N_2$ are normal subgroups and thus $N_1 \cdot N_2 \leq G$ is a subgroup of $G$. We claim that its order is just $d$. For this note that $N_1 \cap N_2 \leq N_i$, $i = 1, 2$, and hence its order divides the orders of both, $N_1$ and $N_2$, i. e.

$$|N_1 \cap N_2| \;\Big|\; \operatorname{hcf}\left(|N_1|, |N_2|\right) = \operatorname{hcf}\left(m, p^a\right) = 1.$$

Thus $|N_1 \cap N_2| = 1$. Applying now the Product Formula 1.11 we may calculate the order of $N_1 \cdot N_2$ as

$$|N_1 \cdot N_2| = \frac{|N_1| \cdot |N_2|}{|N_1 \cap N_2|} = m \cdot p^a = d.$$

□

**4.5 Definition**

Let $(G, \cdot)$ be a group of order $p^a \cdot m$ with $\operatorname{hcf}(p, m) = 1$. We call the elements of

$$\operatorname{Syl}_p(G) := \left\{U \leq G \;\middle|\; |U| = p^a\right\}$$

$p$-*Sylow subgroups* of $G$.

**4.6 Theorem of Sylow**

*Let $(G, \cdot)$ be a finite group and let $p$ be a prime.*

a. $\left|\operatorname{Syl}_p(G)\right| \equiv 1 \pmod{p}$, *in particular,* $G$ *has* $p$*-Sylow subgroups.*

b. $G$ *acts transitively on the set* $\operatorname{Syl}_p(G)$ *by conjugation, i. e. if* $P, Q \in \operatorname{Syl}_p(G)$, *then there is some* $g \in G$ *such that* $Q = P^g = gPg^{-1}$.
*In particular,* $|\operatorname{Syl}_p(G)| = |G : N_G(P)|$ *is a divisor of* $|G|$.

**Proof:** Part a. is just a special case of Theorem 4.2. It thus only remains to prove Part b. Let $P, Q \in \operatorname{Syl}_p(G)$ be given, and consider the set

$$\Omega = \left\{P^g \;\middle|\; g \in G\right\}.$$

***Step 1*:** $\operatorname{hcf}\left(|\Omega|, p\right) = 1$.

By Corollary 3.4 and Theorem 1.11 we have

$$|\Omega| = |G : N_G(P)| = \frac{|G : P|}{|N_G(P) : P|} = \frac{m}{|N_G(P) : P|}.$$

In particular, $|\Omega|$ is a divisor of $m$, and thus the prime $p$ does not divide $|\Omega|$.

***Step 2*:** $\exists\, g \in G \;:\; Q = P^g$.

The $p$-group $Q$ acts on $\Omega$ by conjugation as well as $G$ does, and by Step 1 $\operatorname{hcf}\left(|\Omega|, p\right) = 1$. Thus Corollary 3.5 applies and we find a fix point of the action of $Q$ on $\Omega$, i. e. there is some $P^g \in \Omega$ such that

$$\operatorname{orb}_Q\left(P^g\right) = \left\{P^g\right\}.$$

This means $hP^gh^{-1} = P^g$ for all $h \in Q$, or equivalently $hP^g = P^gh$ for all $h \in Q$. Thus we have in particular

$$QP^g = P^gQ.$$

Applying Proposition 1.7 this implies that $QP^g$ is a subgroup of $G$. Since $Q \cap P^g$ is a subgroup of $Q$, its order is $p^b$ for some $0 \leq b \leq a$. By the Product Formula 1.11 we then find

$$|QP^g| = \frac{|Q| \cdot |P^g|}{|Q \cap P^g|} = \frac{p^a \cdot p^a}{p^b} = p^{2a-b}.$$

Since $p^a$ is the maximal power of $p$ which divides the order of $|G|$, $2a - b \leq a$, which is only possible if $a = b$. This however implies

$$Q = Q \cap P^g = P^g.$$

□

Since a subgroup $P$ is normal if and only if coincides with all its conjugates $P^g$, the following is an immediate corollary.

**4.7 Corollary**
*Let* $(G, \cdot)$ *be a finite group,* $p$ *a prime and* $P \in \mathrm{Syl}_p(G)$.
*Then* $P \trianglelefteq G$ *if and only if* $\mathrm{Syl}_p(G) = \{P\}$.

**4.8 Theorem**
*Let* $(G, \cdot)$ *be a group of order* $2p$*, where* $p > 2$ *is some prime.*
*Then either* $(G, \cdot) \cong (\mathbb{Z}/2p\mathbb{Z}, +)$ *is cyclic or* $(G, \cdot) \cong (\mathbb{D}_{2n}, \circ)$ *is not abelian.*

**Proof:** By Theorem 4.6 there are subgroups $P \in \mathrm{Syl}_p(G)$ of order $p$ and $U \in \mathrm{Syl}_2(G)$ of order 2. Since their orders are prime numbers, they must be cyclic by Proposition 2.5, i. e. $P = \langle g \rangle$ and $U = \langle u \rangle$ for some $g, u \in G$. Note, that $U \cap P = \mathbb{1}$, since the order of the intersection must be a divisor of 2 and of $p$.
By the Theorem of Lagrange $|G : P| = 2$, and thus by Proposition 1.15 $P \trianglelefteq G$. Moreover, by the Product Formula 1.11 we see that $|UP| = 2p$, and hence $G = UP = \langle u, g \rangle$.

***Case 1*:** $U \trianglelefteq G$.

Then, since $U \cap P = \mathbb{1}$, Proposition 1.16 e. applies and we have

$$u \cdot g = g \cdot u.$$

However, then by Proposition 2.7 the element $ug$ has order $o(ug) = 2p$. Thus $G = \langle ug \rangle$ is cyclic of order $2p$, and by Theorem 2.2 its isomorphic to $(\mathbb{Z}/2p\mathbb{Z}, +)$.

***Case 2*:** $U$ is not normal in $G$.

Since $P$ is a normal subgroup of $G$, the map

$$\alpha_u : P \to P : h \mapsto h^u = uhu^{-1}$$

takes values in $P$ and is thus an automorphism of $P$. However, since $\alpha_u^2 = \alpha_{u^2} = \alpha_{e_G} = \mathrm{id}_G$, the automorphism $\alpha_u$ has order 2. By Corollary 2.10 the inversion

$$\mathrm{inv}_P : P \to P : h \mapsto h^{-1}$$

is the only automorphism of $P$ of order 2, i. e. $\alpha_u = \mathrm{inv}_P$. But then

$$u \cdot g \cdot u^{-1} = g^{-1}.$$

In particular, the generators $u$ and $g$ of $G$ satisfy the relations of the generators $\tau$ and $\sigma$ of the group $\mathbb{D}_{2p}$ in Exercise 1 on Assignment Set 1. This can be used to define an isomorphism

$$G \to \mathbb{D}_{2p} : u \mapsto \tau, g \mapsto \sigma.$$

We leave the details of this to the reader. □

**4.9 Remark**

We have just proved in particular that up to isomorphism there are only two groups with 10 elements, which may serve as alphabets for a CDC, namely the cyclic group $\mathbb{Z}/10\mathbb{Z}$ of order 10 and the dihedral group $\mathbb{D}_{10}$ of order 10. This answers the question at the end of Paragraph 0.

CHAPTER II

# Normal Forms of Linear and Bilinear Maps

## 1 Jordan Normal Form

### 1.0 General Assumptions and Reminder

Throughout this section $K$ will be a *field* and $V$ a finite-dimensional $K$-*vector space*. A *basis* $B = (v_1, \ldots, v_n)$ of $V$ is a family of vectors in $V$ such that

(i) $B$ is *linearly independent*, i. e. $\sum_{i=1}^{n} \lambda_i v_i = 0$, $\lambda_i \in K$, implies $\lambda_1 = \ldots = \lambda_n = 0$.

(ii) $B$ generates $V$, i. e. $\forall\, v \in V\ \exists\, \lambda_1, \ldots, \lambda_n \in K\ :\ v = \sum_{i=1}^{n} \lambda_i v_i$.

Moreover, if $B$ is a basis, then by (i) the coefficients $\lambda_i$ in (ii) for the representation of $v$ are uniquely determined, and we call

$$M_B(v) = (\lambda_1, \ldots, \lambda_n)^t \in K^n$$

the *basis representation* of $v$ w. r. t. $B$. Thus a basis $B$ determines an isomorphism

$$M_B : V \to K^n : v \mapsto M_B(v).$$

Recall also, that all bases of $V$ have the same number of elements, and this number $n = \dim_K(V)$ is called the *dimension* of $V$.
A typical example is the vector space $K^n$ with the standard basis

$$E = (e_1, \ldots, e_n),$$

where $e_i$ is the column vector which has entry $1$ in the $i$-th row and entry $0$ else.
When considering vectors of the form $(x_1, \ldots, x_n)^t \in K^n$, we will frequently denote them just by $\underline{x}$.

### 1.1 Definition and Proposition (Matrix Representation)

Let[1] $f \in \mathrm{End}_K(V) = \{f : V \to V \mid f \text{ is } K\text{-linear}\}$ and $B = (v_1, \ldots, v_n)$ be a basis of $V$.

a. As we have noticed above, there exist uniquely determined coefficients $a_{ij} \in K$, $i, j = 1, \ldots, n$, such that

$$f(v_j) = \sum_{i=1}^{n} a_{ij} v_i.$$

We call the $n \times n$-matrix

$$M_B^B(f) = (a_{ij})_{i,j=1,\ldots,n} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

[1] $K$-linear maps from $V$ to $V$ are called *endomorphisms*.

the *matrix representation* of $f$ w. r. t. the basis $B$.[2] In the lecture "Linear Algebra" it was shown that this establishes a one-to-one correspondence between linear maps from $V$ to $V$ and $n \times n$-matrices, i. e. the following map is a bijection[3]

$$M_B^B : \mathrm{End}_K(V) \to \mathrm{Mat}(n \times n, K) : f \mapsto M_B^B(f).$$

Note that in the case $V = K^n$ and $B = E$ the inverse map of $M_E^E$ is just

$$\mathrm{Mat}(n \times n, K) \to \mathrm{End}_K(B) : A \mapsto f_A,$$

where the map $f_A$ is defined by

$$f_A : K^n \to K^n : \underline{x} \mapsto A \cdot \underline{x}.$$

Moreover, it has been shown in "Linear Algebra" how the matrix representations of vectors and of linear maps fit together. Let $v \in V$ and $f \in \mathrm{End}_K(V)$, then

$$M_B\big(f(v)\big) = M_B^B(f) \cdot M_B(v), \tag{3}$$

i. e. if $v = \sum_{i=1}^n \lambda_i v_i$, $f(v) = \sum_{i=1}^n \mu_i v_i$ and $M_B^B(f) = (a_{ij})_{i,j=1,\dots,n}$ then

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

b. If $B' = (v_1', \dots, v_n')$ is another basis of $V$, then the matrix

$$T_{B'}^B = (t_{ij})_{i,j=1,\dots,n} \in \mathrm{Gl}_n(K)$$

with the property that

$$v_j' = \sum_{i=1}^n t_{ij} v_i$$

for $j = 1, \dots, n$ is called the *base change* from $B$ to $B'$.
In "Linear Algebra" it was shown that the matrix representation of $f$ w. r. t. the bases $B$ respectively $B'$ satisfy the following relation

$$M_{B'}^{B'}(f) = \big(T_{B'}^B\big)^{-1} \cdot M_B^B(f) \cdot T_{B'}^B. \tag{4}$$

Note also that

$$\big(T_{B'}^B\big)^{-1} = T_B^{B'}.$$

As an easy example let us consider $V = K^2$, $f : K^2 \to K^2 : (x, y)^t \mapsto (2x - y, -x)^t$, $B = E = (e_1, e_2)$ and $B' = (v_1', v_2')$ with $v_1' = (1, 1)^t$ and $v_2' = (1, -1)^t$. Since

$$f(v_1') = (1, -1)^t = 0 \cdot v_1' + 1 \cdot v_2' \quad \text{and} \quad f(v_2') = (3, -1)^t = 1 \cdot v_1' + 2 \cdot v_2'$$

[2] The j-th column of $M_B^B(f)$ is thus just the basis representation of the vector $f(v_j)$ w. r. t. the basis $B$.

[3] Actually, $\mathrm{End}_K(V)$ and $\mathrm{Mat}(n \times n, K)$ both carry the structure of a $K$-algebra, making $M_B^B$ a $K$-algebra isomorphism.

we have

$$M_{B'}^{B'} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

Dealing with the standard basis is even easier and leads to

$$M_B^B = \begin{pmatrix} 2 & -1 \\ -1 & 0 \end{pmatrix}.$$

We may calculate $T_{B'}^B$ by just taking the vectors of $B'$ as columns, since $B$ is the standard basis; calculating the inverse as well we get

$$T_{B'}^B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad (T_{B'}^B) = \tfrac{1}{2} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

It is now an easy exercise to verify Equation (4).

Let us also check Equation (3) for one vector, say $v = (2,0)^t$. Since $v = 1 \cdot v_1' + 1 \cdot v_2'$ and $f(v) = (4,-2)^t = 1 \cdot v_1' + 3 \cdot v_2'$ we get

$$M_{B'}^{B'}(f) \cdot M_{B'}(v) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} = M_{B'}\big(f(v)\big).$$

**1.2 Definition and Proposition (Base Change)**

We call two matrices $A, B \in \mathrm{Mat}(n \times n, K)$ *similar* and write $A \sim B$ if and only if there is a $T \in \mathrm{Gl}_n(K)$ such that

$$B = T^{-1} \cdot A \cdot T.$$

Similarity of matrices is an equivalence relation, as one easily sees, and the equivalence class of $A$ is called the *similarity class* of $A$.

**1.3 Aim**

> Given $f \in \mathrm{End}_K(V)$ we want to find a basis $B$ of $V$ such that $M_B^B(f)$ is as simple as possible.

Taking the interplay between matrices and linear maps into consideration, this is equivalent to the following problem:

> Given $A \in \mathrm{Gl}_n(K)$ find a $T \in \mathrm{Gl}_n(K)$ such that $T^{-1} \cdot A \cdot T$ is as simple as possible.

That is, we are looking for a simple representative of the similarity class of $A$. Such a representative would be called a *normal form* of $A$.

Of course, we have to specify, what we mean by *simple*! The precise meaning has to be looked up in Remark 1.21, where we describe what it means that a matrix is in *Jordan normal form*. For the moment it is sufficient to say that simple means it should be close to being diagonal.

**1.4 Definition and Proposition**

Let $f \in \mathrm{End}_K(V)$ and $A \in \mathrm{Mat}(n \times n, K)$ be given.

We call $\lambda \in K$ an *eigenvalue* of $f$ (resp. of $A$) if and only if one of the following equivalent conditions is fulfilled:

1) $\exists\, 0 \neq v \in V$ (resp. $0 \neq \underline{x} \in K^n$) such that $f(v) = \lambda \cdot v$ (resp. $A \cdot \underline{x} = \lambda \cdot \underline{x}$).
2) $\mathrm{Eig}(f, \lambda) := \mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V) \neq \{0\}$ (resp. $\mathrm{Eig}(A, \lambda) := \mathrm{Ker}(A - \lambda \cdot \mathbb{1}) \neq \{0\}$).

3) $\chi_f(\lambda) = 0$ (resp. $\chi_A(\lambda) = 0$), where $\chi_f = \det(f - t\cdot \mathrm{id}_V)$ (resp. $\chi_A = \det(A - t\cdot \mathbb{1})$) denotes the *characteristic polynomial* of $f$ (resp. $A$).

Vectors which satisfy the equation in 1) are called *eigenvectors* of $f$ (resp. $A$) w. r. t. the eigenvalue $\lambda$, and the kernel in 2) is called the corresponding *eigenspace*.

**1.5 Example**

Let $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Mat}(2\times 2, K)$. Then

$$\chi_A = \det \begin{pmatrix} 1-t & 1 \\ 0 & 1-t \end{pmatrix} = (1-t)^2.$$

That is, $\lambda = 1$ is the only eigenvalue of $A$, and we say, it has *multiplicity* two, since the factor $(1-t)$ occurs twice in the characteristic polynomial.

Let's now calculate the eigenspace of $A$ w. r. t. $\lambda$. By definition this is the set of solutions of the following homogeneous system of linear equations:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = (A - 1\cdot \mathbb{1}) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Of course $(x, y)^t$ satisfies this equation if and only if $x = 0$. Thus

$$\mathrm{Eig}(A, 1) = \mathrm{Ker}(A - \mathbb{1}) = \left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle.$$

**1.6 Proposition**

*Let $f \in \mathrm{End}_K(V)$ and $A \in \mathrm{Mat}(n\times n, K)$. Then $f$ (resp. $A$ )is diagonalisable*[4] *if and only if $V$ (resp. $K^n$) has a basis of eigenvectors of $f$ (resp. $A$).*

**Proof:** Let's first suppose that there is a basis $B = (v_1, \ldots, v_n)$ of eigenvectors, and let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues. Then

$$f(v_i) = \lambda_i \cdot v_i \qquad \text{resp.} \qquad A \cdot v_i = \lambda_i \cdot v_i.$$

This however implies

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix} \text{ resp. } T^{-1}\cdot A \cdot T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}, \quad (5)$$

where $T$ is the matrix whose columns are the vectors $v_1, \ldots, v_n$.

Let's now suppose that there is a basis $B = (v_1, \ldots, v_n)$ (resp. a matrix $T \in \mathrm{Gl}_n(K)$) such that Equation (5) is fulfilled, and call in the latter case the column vectors of $T$ just $v_1, \ldots, v_n$. Then

$$f(v_i) = \lambda_i v_i \qquad \text{resp.} \qquad A \cdot v_i = \lambda_i v_i.$$

Thus $(v_1, \ldots, v_n)$ is a basis of eigenvectors. □

[4]Recall, $f$ is said to be *diagonalisable* if and only if there is a basis $B$ of $V$ such that $M_B^B(f)$ is a diagonal matrix. Analogously, $A$ is *diagonalisable* if and only if it is similar to a diagonal matrix, i. e. if there is a $T \in \mathrm{Gl}_n(K)$ such that $T^{-1}\cdot A \cdot T$ is a diagonal matrix.

**1.7 Example (Example 1.5 continued)**
Since $\dim_K\left(\mathrm{Eig}(A,1)\right) = 1 < 2 = \dim_K\left(K^2\right)$ and since $A$ has no other eigenvalues, $K^2$ does not possess a basis of eigenvectors of $A$, and hence $A$ is *not diagonalisable*. In particular, not every endomorphism and not every square matrix can have a diagonal matrix as normal form!

**1.8 Corollary**
*Let* $f \in \mathrm{End}_K(V)$ *with* $\dim_K(V) = n$ *and let* $A \in \mathrm{Mat}(n\times n, K)$. *Suppose that* $f$ *(resp.* $A$*) has pairwise distinct eigenvalues* $\lambda_1, \ldots, \lambda_n$, *then* $f$ *(resp.* $A$*) is diagonalisable.*

**Proof:** [5] Let $v_1, \ldots, v_n \in V$ be the corresponding eigenvectors of $f$. By Proposition 1.6 it suffices to prove that $B = (v_1, \ldots, v_n)$ is a basis of $V$. For that, however, it suffices that $B$ is linearly independent, since $V$ has dimension $n$.
We proof by induction on $k$, that the vectors $v_1, \ldots, v_k$ are linearly independent.
For $k = 1$, there is nothing to show since by hypothesis an eigenvector is non-zero.
Let now $k > 1$ and let's assume that we have already shown that $v_1, \ldots, v_{k-1}$ are linearly independent.
Let $\mu_1, \ldots, \mu_k \in K$ such that $\sum_{i=1}^{k} \mu_i v_i = 0$. We have to show that then $\mu_1 = \ldots = \mu_k = 0$.

$$\sum_{i=1}^{k} \lambda_k \mu_i v_i = \lambda_k \sum_{i=1}^{k} \mu_i v_i = 0 = f(0) = f\left(\sum_{i=1}^{k} \mu_i v_i\right) = \sum_{i=1}^{k} \mu_i f(v_i) = \sum_{i=1}^{k} \mu_i \lambda_i v_i.$$

Subtracting both sides from each other, we get

$$0 = \sum_{i=1}^{k} (\lambda_k \mu_i - \lambda_i \mu_i) \cdot v_i = \sum_{i=1}^{k-1} (\lambda_k - \lambda_i) \cdot \mu_i \cdot v_i.$$

Since by induction $v_1, \ldots, v_{k-1}$ are linearly independent, this implies that

$$(\lambda_k - \lambda_i) \cdot \mu_i = 0$$

for $i = 1, \ldots, k-1$. And since by assumption $\lambda_k - \lambda_i \neq 0$, this implies

$$\mu_i = 0$$

for $i = 1, \ldots, k-1$. We are thus left with $0 = \sum_{i=1}^{k} \mu_i v_i = \mu_k v_k$, and since the eigenvector $v_k \neq 0$, we finally get $\mu_k = 0$. Thus $v_1, \ldots, v_k$ are linearly independent, and in particular $B$ is a basis of $V$. □

**1.9 Proposition**
*An endomorphism* $f \in \mathrm{End}_K(V)$ *(resp. a square matrix* $A \in \mathrm{Mat}(n \times n, K)$*) is triangulable*[6] *if and only if the characteristic polynomial* $\chi_f$ *(resp.* $\chi_A$*) factorises into linear factors.*

[5] We do the proof for endomorphisms, the proof for matrices is identical, if you replace $f$ by $A$ and $V$ by $K^n$.

[6] Recall, $f$ is said to be *triangulable* if and only if there is a basis $B$ of $V$ such that $M_B^B(f)$ is an upper triangular matrix. Analogously, $A$ is *triangulable* if and only if it is similar to an upper triangular matrix, i. e. if there is a $T \in \mathrm{Gl}_n(K)$ such that $T^{-1} \cdot A \cdot T$ is an upper triangular matrix.

**Proof:** [7] Let's first suppose we have a basis $B$ such that

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & * & \dots & \dots & * \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix},$$

where $*$ represents an entry which need not be zero. Then the characteristic polynomial of $f$ is

$$\chi_f = \det(f - t\,\mathrm{id}_V) = \det\left(M_B^B(f) - t\mathbb{1}\right) = \det\begin{pmatrix} \lambda_1 - t & * & \dots & \dots & * \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ 0 & \dots & \dots & 0 & \lambda_n - t \end{pmatrix} = \prod_{i=1}^{n} (\lambda_i - t),$$

in particular it factorises into linear factors.

The other direction is somewhat harder to proof. We do the proof by induction on $n = \dim_K(V)$. If $n = 1$, then there is nothing to prove, since every $n \times n$-matrix is automatically a "diagonal" matrix. Let therefore $n > 1$ and suppose that endomorphisms of vector spaces of dimension $n-1$ whose characteristic polynomial factorises are diagonalisable.

By assumption $\chi_f = (\lambda_1 - t)\cdots(\lambda_n - t)$, and $\lambda_1$ is therefore an eigenvalue of $f$. Let $0 \neq v_1 \in V$ be an eigenvector of $f$ w. r. t. $\lambda_1$. We set $U = \langle v_1 \rangle$, the subspace of $V$ generated by $v_1$. Since $v_1$ is an eigenvector of $f$, this space is $f$-invariant, i. e. $f(u) \in U$ for all $u \in U$. We may therefore consider the restriction of $f$ to $U$, denoted by

$$f_U : U \to U : u \mapsto f(u).$$

Moreover, $f$ induces an endomorphism on $V/U$ by

$$f_{V/U} : V/U \to V/U : v + U \mapsto f(v) + U,$$

which is well-defined since $U$ is $f$-invariant.

If we extend $B' = (v_1)$ to a basis of $V$ by vectors $v_2, \dots, v_n$, then the residue classes $B'' = (v_2 + U, \dots, v_n + U)$ form a basis of the quotient space $V/U$. It is straightforward to see that the matrix representations of $f$, $f_U$ and $f_{V/U}$ w. r. t. the bases $B$, $B'$ and $B''$ satisfy the following relation:

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \ \cdots \ * \\ \hline 0 & \\ \vdots & M_{B''}^{B''}(f_{V/U}) \\ 0 & \end{array}\right), \tag{6}$$

where the $*$'s are suitable entries. In particular, for the characteristic polynomials we have

$$(\lambda_1 - t)\cdots(\lambda_n - t) = \chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}} = (\lambda_1 - t) \cdot \chi_{f_{V/U}}.$$

[7] Once again we do the proof only for the case of endomorphisms and leave it to the reader to do the necessary replacements for matrices.

But then $f_{V/U}$ is an endomorphism of an $n-1$-dimensional vector space whose characteristic polynomial factorises. So, by induction, there is a basis of $V/U$ which triangulates $f_{V/U}$. W. l. o. g. we may assume that $B'' = (v_2+U, \ldots, v_n+U)$ is such a matrix. But then by Equation 6 we see that $B$ triangulates $f$. □

**1.10 Remark**

Proposition 1.9 says that, if we want a normal form for $f$ (resp. $A$) which is at least an upper triangular matrix, then we have to request that the characteristic polynomial factorises! We will therefore restrict in the theorems on the Jordan normal form to this case!

**1.11 Definition**

Let $f \in \mathrm{End}_K(V)$ and $A \in \mathrm{Mat}(n \times n, K)$ be given with characteristic polynomial $\chi_f = (t-\lambda)^n \cdot p$ resp. $\chi_A = (t-\lambda)^k \cdot p$, where $p \in K[t]$ is a polynomial such that $p(\lambda) \neq 0$.

We then call $\mathrm{mult}_{alg}(f,\lambda) := k$ resp. $\mathrm{mult}_{alg}(A,\lambda) := k$ the *algebraic multiplicity* of $\lambda$ as an eigenvalue of $f$ resp. $A$. That is, the algebraic multiplicity is the multiplicity of $\lambda$ as a zero of the characteristic polynomial.

And we call $\mathrm{mult}_{geo}(f,\lambda) := \dim_K\big(\mathrm{Eig}(f,\lambda)\big)$ resp. $\mathrm{mult}_{geo}(A,\lambda) := \dim_K\big(\mathrm{Eig}(A,\lambda)\big)$ the *geometric multiplicity* of $\lambda$ as an eigenvalue of $f$ resp. of $A$.

The following proposition gives a direct relation between the geometric and the algebraic multiplicity of an eigenvalue, so that a look at the factorised characteristic polynomial suffices to find upper bounds for the dimension of the eigenspaces.

**1.12 Proposition (Geometric and Algebraic Multiplicity)**

*Let $f \in \mathrm{End}_K(V)$ and $A \in \mathrm{Mat}(n \times n, K)$, and let $\lambda \in K$. The geometric multiplicity of $\lambda$ as an eigenvalue of $f$ resp. $A$ is less than or equal to the algebraic multiplicity of $\lambda$ as an eigenvalue of $f$ resp. $A$, i. e.*

$$\mathrm{mult}_{geo}(f,\lambda) \leq \mathrm{mult}_{alg}(f,\lambda) \qquad \textit{and} \qquad \mathrm{mult}_{geo}(A,\lambda) \leq \mathrm{mult}_{alg}(A,\lambda).$$

**Proof:** We do the proof for the case of an endomorphism only.

Let $m := \mathrm{mult}_{geo}(f,\lambda)$ and let $(v_1,\ldots,v_m)$ be a basis of $\mathrm{Eig}(f,\lambda)$. Extend this to a basis $B = (v_1,\ldots,v_n)$ of $V$. We know that

$$f(v_j) = \lambda \cdot v_j$$

for $j = 1,\ldots,m$ and that there $a_{ij} \in K$, $j = m+1,\ldots,n$ and $i = 1,\ldots,n$, such that

$$f(v_j) = \sum_{i=1}^{n} a_{ij} \cdot v_i$$

for $j = m+1,\ldots,n$. If we now set

$$M = \begin{pmatrix} a_{m+1,m+1} & \ldots & a_{m+1,n} \\ \vdots & & \vdots \\ a_{n,m+1} & \ldots & a_{nn} \end{pmatrix} \quad \text{and} \quad M' = \begin{pmatrix} a_{1,m+1} & \ldots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,m+1} & \ldots & a_{mn} \end{pmatrix}$$

then

$$M_B^B(f) = \left(\begin{array}{c|c} \lambda \cdot \mathbb{1}_m & M' \\ \hline 0_{(n-m)\times m} & M \end{array}\right).$$

Thus we have

$$\chi_f = \det\left(\begin{array}{c|c} (\lambda - t)\cdot \mathbb{1}_m & M' \\ \hline 0_{(n-m)\times m} & M - t\cdot \mathbb{1}_{n-m} \end{array}\right) = (\lambda - t)^m \cdot \chi_M.$$

In particular, the multiplicity $\mathrm{mult}_{alg}(f,\lambda)$ of $\lambda$ as a zero of $\chi_f$ is at least $m = \mathrm{mult}_{geo}(f,\lambda)$. □

The Theorem of Cayley-Hamilton is one of the key ingredients in the proof of the existence of the Jordan normal form and at the same time it helps actually calculating them.

**1.13 Theorem (Cayley-Hamilton)**
*Let* $f \in \mathrm{End}_K(V)$ *and* $A \in \mathrm{Mat}(n\times n, K)$, *then* $\chi_f(f) = 0$ *and* $\chi_A(A) = 0$.

**Proof:** [8] Let's first prove the statement for matrices.
Consider the matrix $B_t = A - t\cdot \mathbb{1} \in M := \mathrm{Mat}(n\times n, K)[t] = \mathrm{Mat}\left(n\times n, K[t]\right)$. Linear Algebra tells us that the adjoint matrix $\mathrm{adj}(B_t) \in M$ of $B_t$ satisfies the following equation

$$B_t \cdot \mathrm{adj}(B_t) = \det(B_t)\cdot \mathbb{1} = \chi_A \cdot \mathbb{1}. \tag{7}$$

However, remembering how the adjoint of a matrix is actually defined, we find that if $\mathrm{adj}(B_t) = (b_{ij})_{i,j=1,\ldots,n}$ then

$$b_{ij} = (-1)^{i+j}\cdot \det(C_{ji}),$$

where $C_{ji}$ is derived from the matrix $B_t$ by erasing the $j$-th row and the $i$-th column. In particular, $b_{ij}$ is a polynomial in $t$ of degree at most $n-1$, since $C_{ji}$ is an $(n-1)\times(n-1)$-matrix where each row contains at most one entry which is a non-constant polynomial in $t$, and this entry is then linear in $t$.
Thus the entries of $\mathrm{adj}(B_t)$ are all polynomials of degree at most $n-1$ and we may therefore consider $\mathrm{adj}(B_t)$ as a polynomial of degree at most $n-1$ with matrix coefficients, as indicated in Footnote 8, i. e. there are matrices $B_0,\ldots,B_{n-1} \in \mathrm{Mat}(n\times n, K)$ such that

$$B_t = B_{n-1}\cdot t^{n-1} + \ldots + B_1\cdot t + B_0.$$

Let now $\chi_A = (-1)^n\cdot t^n + \alpha_{n-1}\cdot t^{n-1} + \ldots + \alpha_1\cdot t + \alpha_0$, then Equation (7) implies

$$(A - t\cdot\mathbb{1})\cdot\left(B_{n-1}\cdot t^{n-1} + \ldots + B_0\right) = (-1)^n\cdot\mathbb{1}\cdot t^n + \alpha_{n-1}\cdot\mathbb{1}\cdot t^{n-1} + \ldots + \alpha_1\mathbb{1}\cdot t + \alpha_0\cdot\mathbb{1}.$$

[8] Note that the we have the following equality of sets $\mathrm{Mat}\left(n\times n, K[t]\right) = \mathrm{Mat}(n\times n, K)[t]$, where the first one is the set of $n\times n$-matrices whose entries are polynomials, while the second one is the set of polynomials whose coefficients are $n\times n$-matrices. Let's illustrate by an example how elements in the two sets are identified:

$$\begin{pmatrix} t^2-2 & t^2-t \\ t+3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\cdot t^2 + \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\cdot t + \begin{pmatrix} -2 & 0 \\ 3 & 0 \end{pmatrix}.$$

As usually with polynomial identities, we may compare the coefficients and get:

$$\begin{array}{lll} A \cdot B_0 & = & \alpha_0 \cdot \mathbb{1} \\ A \cdot B_1 - B_0 & = & \alpha_1 \cdot \mathbb{1} \\ \vdots & & \vdots \\ A \cdot B_{n-1} - B_{n-2} & = & \alpha_{n-1} \cdot \mathbb{1} \\ - B_{n-1} & = & (-1)^n \cdot \mathbb{1} \end{array} \tag{8}$$

Multiplying the $i$-th row in Equation (8) by $A^{i-1}$ we get

$$\begin{array}{lll} A \cdot B_0 & = & \alpha_0 \cdot \mathbb{1} \\ A^2 \cdot B_1 - A \cdot B_0 & = & \alpha_1 \cdot A \\ \vdots & & \vdots \\ A^n \cdot B_{n-1} - A^{n-1} \cdot B_{n-2} & = & \alpha_{n-1} \cdot A^{n-1} \\ - A^n \cdot B_{n-1} & = & (-1)^n \cdot A^n \end{array} \tag{9}$$

Adding the terms on the left hand side in Equation (9) we get the zero-matrix, adding the terms on the right hand side, we get $\chi_A(A)$. This proves the statement. For an endomorphism $f$ we note that

$$M_B^B\big(\chi_f(f)\big) = \chi_f\big(M_B^B(f)\big) = \chi_{M_B^B(f)}\big(M_B^B(f)\big) = 0,$$

by the previously shown result for matrices. However, if the endomorphism $\chi_f(f)$ has matrix representation $0$ w. r. t. some basis $B$, then it must be the zero endomorphism. □

**1.14 Remark**

Would not the following be a much shorter proof of the above theorem?

$$\chi_A(A) = \det(A - A \cdot \mathbb{1}) = \det(0) = 0.$$

What is wrong with this proof?

Note, that $\chi_A(A)$ is by definition a $n \times n$-matrix, while $\det(0)$, the determinant of the zero matrix, is just a number! They can hardy coincide!

The problem is, that we substituted $A$ for the variable $t$ in the above equation in the wrong way!

**1.15 Example (Example 1.7 continued)**

The characteristic polynomial of $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ was calculated as $\chi_A = (1-t)^2$. Let's now plug in $A$:

$$\chi_A(A) = \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right)^2 = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

**1.16 Lemma**

*Let* $g \in \mathrm{End}_K(V)$, *then there is an* $m \geq 1$ *such that*

$$\mathrm{Ker}(g) \subsetneq \mathrm{Ker}\left(g^2\right) \subsetneq \ldots \subsetneq \mathrm{Ker}\left(g^m\right) = \mathrm{Ker}\left(g^k\right) \quad \textit{for all} \quad k \geq m.$$

**Proof:** This is Exercise 4 on Assignment Set 6. □

**1.17 Theorem (Jordan Normal Form – $2 \times 2$-Case)**

a. *Let* $f \in \mathrm{End}_K(V)$ *with* $\dim_K(V) = 2$ *such that* $\chi_f = (\lambda_1 - t) \cdot (\lambda_2 - t)$. *Then there exists either a basis* $B$ *of* $V$ *such that*

$$J(f) := M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

*or a basis* $B$ *of* $V$ *such that*

$$J(f) := M_B^B(f) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \qquad \textit{and} \qquad \lambda = \lambda_1 = \lambda_2.$$

*We call* $J(f)$ *a* Jordan normal form *of* $f$.

b. *Let* $A \in \mathrm{Mat}(2 \times 2, K)$ *such that* $\chi_A = (\lambda_1 - t) \cdot (\lambda_2 - t)$. *Then there exists either a* $T \in \mathrm{Gl}_2(K)$ *such that*

$$J(A) := T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

*or a* $T \in \mathrm{Gl}_2(K)$ *such that*

$$J(A) := T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \qquad \textit{and} \qquad \lambda = \lambda_1 = \lambda_2.$$

*We call* $J(A)$ *a* Jordan normal form *of* $A$.

**Proof:** We do the proof for endomorphisms by considering different cases, and we leave it to the reader to translate this proof to the case of matrices.

***1st Case*:** $\lambda_1 \neq \lambda_2$. Then by Corollary 1.8 $f$ is diagonalisable, and we are done.

***2nd Case*:** $\lambda_1 = \lambda_2$ and $\dim_K\big(\mathrm{Eig}(f, \lambda)\big) = 2$. Then by Proposition 1.6 $f$ is diagonalisable, and we are done.

***3rd Case*:** $\lambda_1 = \lambda_2 =: \lambda$ and $\dim_K\big(\mathrm{Eig}(f, \lambda)\big) = 1$. By the Theorem of Cayley-Hamilton we have $(f - \lambda \cdot \mathrm{id}_V)^2 = \chi_f(f) = 0$ and thus the kernel of this map is the whole vector space $V$. Taking the dimension of the eigenspace into account, we get

$$\{0\} \subsetneq \mathrm{Ker}(f - \lambda \cdot \mathrm{id}_V) = \mathrm{Eig}(f, \lambda) \subsetneq \mathrm{Ker}\big((f - \lambda \cdot \mathrm{id}_V)^2\big) = V$$

Choose any $w \in V \setminus \mathrm{Eig}(f, \lambda) = \mathrm{Ker}\big((f - \lambda \cdot \mathrm{id}_V)^2\big) \setminus \mathrm{Eig}(f, \lambda)$ and set $v = (f - \lambda \cdot \mathrm{id}_V)(w)$. Then, by the choice of $w$,

$$0 \neq v \in \mathrm{Ker}\big(f - \lambda \cdot \mathrm{id}_V\big) = \mathrm{Eig}(f, \lambda).$$

In particular, $B = (v, w)$ is linearly independent and thus a basis of $V$. Moreover, we have $f(v) = \lambda v$, since $v$ is an eigenvector, and $f(w) = v + \lambda w$. This leads to the following matrix representation

$$M_B^B(f) = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

□

The proof of the Theorem was constructive and allows to calculate the Jordan normal form and the basis resp. transformation matrix leading to the Jordan normal form. In the first two cases we just have to calculate a basis of the eigenspaces and they give either the desired basis $B$ or the transformation matrix $T$, if we take them as columns of $T$.

**1.18 Example**

a. Let $A = \left(\begin{smallmatrix} -1 & 3 \\ 3 & -1 \end{smallmatrix}\right) \in \mathrm{Mat}(2 \times 2, K)$. Then the characteristic polynomial is

$$\chi_A = \det(A - t \cdot \mathbb{1}) = (-1-t)^2 - 9 = t^2 + 2t - 8 = (t+4) \cdot (t-2).$$

Thus we are in Case 1 and the Jordan normal form will be

$$J(A) = \begin{pmatrix} -4 & 0 \\ 0 & 2 \end{pmatrix}.$$

Let's now calculate the transformation matrix $T \in \mathrm{Gl}_2(K)$.
For this we first have to calculate the eigenspace of $A$ w. r. t. $-4$. Solving the system of linear equations

$$\begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = (A + 4\mathbb{1}) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

leads to

$$\mathrm{Eig}(A, -4) = \left\langle (1, -1)^t \right\rangle.$$

Similarly we solve the system of linear equations

$$\begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = (A - 2\mathbb{1}) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

in order to find that the eigenspace of $A$ w. r. t. $2$ is

$$\mathrm{Eig}(A, 2) = \left\langle (1, 1)^t \right\rangle.$$

The transformation matrix is thus the matrix having these two vectors as columns:

$$T = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

b. Let $A = \left(\begin{smallmatrix} 3 & 1 \\ -1 & 1 \end{smallmatrix}\right) \in \mathrm{Mat}(2 \times 2, K)$. Then the characteristic polynomial is

$$\chi_A = \det(A - t \cdot \mathbb{1}) = (3-t) \cdot (1-t) + 1 = t^2 - 4t + 4 = (t-2)^2.$$

Thus we may be in Case 2 or in Case 3. Let's therefore calculate the eigenspace of $A$ w. r. t. $2$.
Solving the system of linear equations

$$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = (A - 2\mathbb{1}) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

leads to

$$\mathrm{Eig}(A, 2) = \left\langle (1, -1)^t \right\rangle.$$

It thus has dimension 1 and we are actually in Case 3. The Jordan normal form will therefore be

$$J(A) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Let's now calculate the transformation matrix $T \in Gl_2(K)$. For this we may choose any vector $w \in K^2 \setminus \mathrm{Eig}(A, 2)$, e. g.

$$w = (1, 0)^t \qquad \text{and} \qquad v = (A - 2\mathbb{1}) \cdot w = (1, -1)^t.$$

Then the transformation matrix will have the vectors $v$ and $w$ as columns:

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

**1.19 Theorem (Jordan Normal Form – $3 \times 3$-Case)**

a. *Let* $f \in \mathrm{End}_K(V)$ *with* $\dim_K(V) = 3$ *such that* $\chi_f = (\lambda_1 - t) \cdot (\lambda_2 - t) \cdot (\lambda_3 - t)$. *Then there exists either a basis* $B$ *of* $V$ *such that*

$$J(f) := M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

*or a basis* $B$ *of* $V$ *such that*

$$J(f) := M_B^B(f) = \left(\begin{array}{cc|c} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ \hline 0 & 0 & \lambda' \end{array}\right)$$

*or a basis* $B$ *of* $V$ *such that*

$$J(f) := M_B^B(f) = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

*We call* $J(f)$ *a* Jordan normal form *of* $f$.

b. *Let* $A \in \mathrm{Mat}(3 \times 3, K)$ *such that* $\chi_A = (\lambda_1 - t) \cdot (\lambda_2 - t) \cdot (\lambda_3 - t)$. *Then there exists either a* $T \in Gl_3(K)$ *such that*

$$J(A) := T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

*or a* $T \in Gl_3(K)$ *such that*

$$J(A) := T^{-1} \cdot A \cdot T = \left(\begin{array}{cc|c} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ \hline 0 & 0 & \lambda' \end{array}\right)$$

*or a* $T \in Gl_3(K)$ *such that*

$$J(A) := T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

*We call* $J(A)$ *a* Jordan normal form *of* $A$.

**Proof:** Again we do the proof for endomorphisms by considering different cases, and we leave it to the reader to translate this proof to the case of matrices.

***1st Case*:** $\lambda_1, \lambda_2, \lambda_3$ are pairwise distinct, or $\lambda_1 = \lambda_2 = \lambda_3$ and $\dim_K \big(\mathrm{Eig}(f, \lambda_1)\big) = 3$. Then by Proposition 1.6 and Corollary 1.8 $f$ is diagonalisable, and we are done.

***2nd Case*:** Just two of the eigenvalues coincide and the corresponding eigenspace has dimension 2.

W. l. o. g. we may assume $\lambda_1 = \lambda_2 \neq \lambda_3$. Let $(v_1, v_2)$ be a basis of $\mathrm{Eig}(f, \lambda_1)$ and let $v_3$ be an eigenvector of $f$ w. r. t. $\lambda_3$. Since $v_3 \notin \mathrm{Eig}(f, \lambda_3) = \langle v_1, v_2 \rangle$, the vector space $\langle v_1, v_2, v_3 \rangle$ has dimension 3. This implies that $B = (v_1, v_2, v_3)$ is a basis of $V$ of eigenvectors, thus

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

***3rd Case*:** Just two of the eigenvalues coincide and the corresponding eigenspace has dimension 1. Again w. l. o. g. we may assume $\lambda_1 = \lambda_2 \neq \lambda_3$.

***Claim*:** $\mathrm{Eig}(f, \lambda_1) = \mathrm{Ker}(f - \lambda_1 \mathrm{id}_V) \subsetneq \mathrm{Ker}\big((f - \lambda_1 \mathrm{id}_V)^2\big)$.

With the aid of the dimension formula for linear maps and taking Proposition 1.12 into account we may calculate the dimension of $\mathrm{Im}(f - \lambda_3 \mathrm{id}_V)$ as

$$\dim_K \big(\mathrm{Im}(f - \lambda_3 \mathrm{id}_V)\big) = \dim_K(V) - \dim_K \big(\mathrm{Ker}(f - \lambda_3 \mathrm{id}_V)\big) = 3 - \dim_K \big(\mathrm{Eig}(f, \lambda_3)\big) = 2.$$

Since by the Theorem of Cayley-Hamilton we have

$$(f - \lambda_1 \mathrm{id}_V)^2 \circ (f - \lambda_3) = \chi_f(f) = 0,$$

we have

$$\mathrm{Im}(f - \lambda_3 \mathrm{id}_V) \subseteq \mathrm{Ker}\big((f - \lambda_1 \mathrm{id}_V)^2\big),$$

so that the latter vector space has dimension at least 2. Thus, since the eigenspace of $f$ w. r. t. $\lambda_1$ has by assumption only dimension 1, the claim follows in view of Lemma 1.16.

Choose now any $v_2 \in \mathrm{Ker}\big((f - \lambda_1 \cdot \mathrm{id}_V)^2\big) \setminus \mathrm{Eig}(f, \lambda_1)$ and set $v_1 = (f - \lambda \cdot \mathrm{id}_V)(v_2)$. Moreover, let $v_3$ be any eigenvector of $f$ w. r. t. $\lambda_3$. Then, by the choice of $v_2$,

$$0 \neq v_1 \in \mathrm{Ker}\big(f - \lambda_1 \cdot \mathrm{id}_V\big) = \mathrm{Eig}(f, \lambda_1),$$

and $(v_1, v_2)$ are linearly independent. Since $\mathrm{Ker}\big((f - \lambda_1 \mathrm{id}_V)^2\big) \cap \mathrm{Eig}(f, \lambda_3) = \{0\}$ by Exercise 6 on Assignment Set 6, also $B = (v_1, v_2, v_3)$ is linearly independent and thus a basis of $V$. Moreover, we have $f(v_1) = \lambda_1 v_1$, since $v_1$ is an eigenvector w. r. t. $\lambda_1$; $f(v_2) = v_1 + \lambda_1 v_2$; and finally $f(v_3) = \lambda_3 v_3$. This leads to the following matrix representation

$$M_B^B(f) = \left(\begin{array}{cc|c} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ \hline 0 & 0 & \lambda_3 \end{array}\right).$$

***4th Case*:** $\lambda_1 = \lambda_2 = \lambda_3$ and $\dim_K \big(\mathrm{Eig}(f, \lambda_1)\big) = 2$.

***Claim*:** $\mathrm{Ker}\big((f - \lambda_1 \mathrm{id}_V)^2\big) = V$.

Let $g = f - \lambda_1 \mathrm{id}_V$. The Theorem of Cayley-Hamilton gives $0 = \chi_f(f) = g^3$. Hence

$$\mathrm{Ker}\left(g^3\right) = V.$$

By Lemma 1.16 we know that the ascending chain

$$\mathrm{Ker}(g) \subseteq \mathrm{Ker}\left(g^2\right) \subseteq \mathrm{Ker}\left(g^3\right) \subseteq \dots$$

will be strictly ascending until the moment where it becomes stationary for good. Since $\mathrm{Ker}(g) = \mathrm{Eig}(f, \lambda_1)$ has dimension 2 this implies

$$\mathrm{Ker}(g) \subsetneq \mathrm{Ker}\left(g^2\right) = \mathrm{Ker}\left(g^3\right),$$

and thus the claim follows.
Choose now any $v_2 \in \mathrm{Ker}\left((f - \lambda_1 \cdot \mathrm{id}_V)^2\right) \setminus \mathrm{Eig}(f, \lambda_1)$ and set $v_1 = (f - \lambda \cdot \mathrm{id}_V)(v_2) \in \mathrm{Eig}(f, \lambda_1)$. Moreover, since $\mathrm{Eig}(f, \lambda_1)$ has dimension 2, we may choose $v_3 \in \mathrm{Eig}(f, \lambda_1)$ linearly independent of $v_1$. Then $B = (v_1, v_2, v_3)$ is linearly independent and thus a basis of $V$. Moreover, we have $f(v_1) = \lambda_1 v_1$, since $v_1$ is an eigenvector w. r. t. $\lambda_1$; $f(v_2) = v_1 + \lambda_1 v_2$; and finally $f(v_3) = \lambda_1 v_1$. This leads to the following matrix representation

$$M_B^B(f) = \left(\begin{array}{cc|c} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ \hline 0 & 0 & \lambda_1 \end{array}\right).$$

***5th Case:*** $\lambda_1 = \lambda_2 = \lambda_3$ and $\dim_K\left(\mathrm{Eig}(f, \lambda_1)\right) = 1$.

***Claim:*** $\{0\} \subsetneq \mathrm{Ker}(f - \lambda_1 \mathrm{id}_V) \subsetneq \mathrm{Ker}\left((f - \lambda_1 \mathrm{id}_V)^2\right) \subsetneq \mathrm{Ker}\left((f - \lambda_1 \mathrm{id}_V)^3\right) = V$.

As in Case 4 we see that $g = f - \lambda_1 \mathrm{id}_V$ satisfies

$$\mathrm{Ker}\left(g^3\right) = V,$$

and since $\mathrm{Ker}(g) = \mathrm{Eig}(f, \lambda_1)$ has only dimension 1, $\mathrm{Ker}\left((f - \lambda_1 \mathrm{id}_V)^2\right)$ must have at least dimension 2 in view of Lemma 1.16. Suppose its dimension was 3. Then

$$g\big(g(v)\big) = g^2(v) = 0 \qquad \text{for all } v \in V.$$

Hence

$$\mathrm{Im}(g) \subseteq \mathrm{Ker}(g) = \mathrm{Eig}(f, \lambda_1).$$

However, by the dimension formula we have

$$\dim_K\left(\mathrm{Im}(g)\right) = \dim_K(V) - \dim_K\left(\mathrm{Ker}(g)\right) = 3 - 1 = 2,$$

in contradiction to the fact that the eigenspace has only dimension 1. Thus $\mathrm{Ker}\left(g^2\right)$ lies strictly between $\mathrm{Ker}(g)$ and $\mathrm{Ker}\left(g^3\right)$.
Choose now $v_3 \in V \setminus \mathrm{Ker}\left((f - \lambda_1 \mathrm{id}_V)^2\right)$ arbitrary, and set $v_2 = (f - \lambda_1 \mathrm{id}_V)(v_3) \in \mathrm{Ker}\left((f - \lambda_1 \mathrm{id}_V)^2\right)$ and $v_1 = (f - \lambda_1 \mathrm{id}_V)(v_2) \in \mathrm{Ker}(f - \lambda_1 \mathrm{id}_V) = \mathrm{Eig}(f, \lambda_1)$. In view of the above claim these vectors form a basis $B = (v_1, v_2, v_3)$ and since $f(v_1) = \lambda_1 v_1$, $f(v_2) = v_1 + \lambda_1 v_2$ and $f(v_3) = v_2 + \lambda_1 v_3$, we get the following matrix representation

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 1 \\ 0 & 0 & \lambda_1 \end{pmatrix}.$$

□

In the same way as for the $2 \times 2$-case the above proof provides an algorithm to calculate the Jordan normal form of an endomorphism resp. a square matrix and the corresponding basis respectively the transformation matrix.

**1.20 Example**

Let us consider the matrix

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 3 & 1 \\ -1 & -4 & -1 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, K).$$

The characteristic polynomial is

$$\chi_A = \det(A - t \cdot \mathbb{1}) = -t^3 + 5t^2 - 8t + 4 = (2-t)^2 \cdot (1-t).$$

The eigenvalues are thus $\lambda_1 = \lambda_2 = 2$ and $\lambda_3 = 1$, and we are either in Case 2 or in Case 3. To decide which of the cases it is, we have to calculate the eigenspace $\mathrm{Eig}(A, 2)$, i. e. we have to solve the following system of linear equations:

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & -4 & -3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (A - 2\mathbb{1}) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Using the Algorithms of Gauß and the fact that the third line of the above matrix is equal to the negative of the sum of the first two lines, we find that the eigenspace has dimension 1 and is

$$\mathrm{Eig}(A, 2) = \left\langle (1, -1, 1)^t \right\rangle.$$

We are therefore in Case 3 and the Jordan normal form of $A$ is

$$J(A) = \left(\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 2 & 0 \\ \hline 0 & 0 & 1 \end{array}\right).$$

In order to find the transformation matrix $T$ we also have to calculate $\mathrm{Ker}\left((A-2\mathbb{1})^2\right)$ and $\mathrm{Eig}(A, 1)$.

The system of linear equations

$$\begin{pmatrix} 0 & 0 & 0 \\ -1 & -3 & -2 \\ 2 & 6 & 4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (A - 2\mathbb{1})^2 \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

leads with the aid of the Gauß algorithm to

$$\mathrm{Ker}\left((A - 2\mathbb{1})^2\right) = \left\langle (1, -1, 1)^t, (2, 0, -1)^t \right\rangle$$

and we may therefore choose $v_2 = (2, 0, -1)^t \in \mathrm{Ker}\left((A - 2\mathbb{1})^2\right) \setminus \mathrm{Eig}(A, 2)$ and set $v_1 = (A - 2\mathbb{1}) \cdot v_2 = (1, -1, 1)^t$.

The corresponding system of linear equations for $\mathrm{Eig}(A, 1)$

$$\begin{pmatrix} 2 & 2 & 1 \\ 0 & 2 & 1 \\ -1 & -4 & -2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (A - \mathbb{1}) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gives the one-dimensional solution space

$$\mathrm{Eig}(A, 1) = \left\langle (0, 1, -2)^t \right\rangle$$

and we may set $v_3 = (0, 1, -2)^t$.

Hence the transformation matrix T has these vectors as column vectors

$$T = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \\ 1 & -1 & -2 \end{pmatrix}.$$

**1.21 Remark (Jordan Normal Form)**
Let $f \in \mathrm{End}_K(V)$ such that $\chi_f = (\lambda_1 - t)^{n_1} \cdots (\lambda_r - t)^{n_r}$ with pairwise distinct $\lambda_i$. Then there is a basis B of V such that[9]

$$J(f) := M_B^B(f) = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i)$$

$$= \left(\begin{array}{c|c|c|c} J_*(\lambda_*) & 0 & \cdots & 0 \\ \hline 0 & \boxed{J_*(\lambda_*)} & & \vdots \\ \hline \vdots & & \ddots & \vdots \\ \hline 0 & \cdots & \cdots & \boxed{J_*(\lambda_*)} \end{array}\right),$$

where

- $m_i = \min\left\{m \geq 1 \;\middle|\; \mathrm{Ker}\left((f - \lambda_i \mathrm{id}_V)^m\right) = \mathrm{Ker}\left((f - \lambda_i \mathrm{id}_V)^{m+1}\right)\right\}$,
- $t_{ij} = \mathrm{rank}\left((f - \lambda_i \mathrm{id}_V)^{j-1}\right) - 2 \cdot \mathrm{rank}\left((f - \lambda_i \mathrm{id}_V)^{j}\right) + \mathrm{rank}\left((f - \lambda_i \mathrm{id}_V)^{j+1}\right)$,
- $n_i = \sum_{j=1}^{m_i} t_{ij}$, and
- $J_j(\lambda_i) = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 & \lambda_i \end{pmatrix} \in \mathrm{Mat}(j \times j, K).$

We call $J(f)$ a *Jordan normal form* of f.
Of course, an analogous statement for square matrices holds as well.
The proof of these will be a major issue in the course Algebra II in Term 2.
Note, that the above description allows to calculate the Jordan normal form once we know a factorisation of the characteristic polynomial, just by calculating certain ranks of endomorphisms resp. matrices, which can easily be done with the aid of the Gauß algorithm.
Note also, that over so called *algebraically closed fields*, such as the complex numbers, every polynomial factorises, so that every square matrix has a Jordan normal form as representative of its similarity class!

[9] Recall that if we have two matrices $A \in \mathrm{Mat}(m \times n, K)$ and $B \in \mathrm{Mat}(p \times q, K)$, then $A \oplus B \in \mathrm{Mat}\left((m+p) \times (n+q), K\right)$ denotes the block diagonal matrix

$$A \oplus B = \left(\begin{array}{c|c} A & 0_{m \times q} \\ \hline 0_{p \times n} & B \end{array}\right).$$

# 2 Normal Forms of Symmetric Bilinear Forms & Matrices and Quadratic Forms

**2.0 General Assumptions** Throughout this section $K$ will be a *field* of $\operatorname{char}(K) \neq 2$, and $V$ will be a finite-dimensional $K$-*vector space*.

**2.1 Definition**

a. *A map*

$$b : V \times V \to K$$

*is called a* bilinear form *if and only if*[10] *for all* $v, w, u \in V$ *and* $\lambda, \mu \in K$

$$b(\lambda v + \mu w, u) = \lambda \cdot b(v, u) + \mu \cdot b(w, u)$$

*and*

$$b(u, \lambda v + \mu w) = \lambda \cdot b(u, v) + \mu \cdot b(u, w).$$

*We denote by* $\mathrm{Bil}_K(V)$ *the set of all bilinear forms on* $V$.

b. *A bilinear form* $b \in \mathrm{Bil}_K(V)$ *is called* symmetric *if and only if for all* $v, w \in V$

$$b(v, w) = b(w, v).$$

**2.2 Example**

a. **(Determinant)** Let $V = K^2$. The determinant map

$$\det : K^2 \times K^2 \to K : \left( \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \right) \mapsto \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$

is a bilinear form on $K^2$ which is *not* symmetric.

b. **(Scalar Product)** Let $V = \mathbb{R}^n$. The standard scalar product on $\mathbb{R}^n$

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} : (\underline{x}, \underline{y}) \mapsto \underline{x}^t \cdot \underline{y} = \sum_{i=1}^{n} x_i y_i$$

is a symmetric bilinear form.

c. **(Standard Example)** Let $V = K^n$ and let $A \in \mathrm{Mat}(n \times n, K)$ be fixed. The map

$$b_A : K^n \times K^n \to K : (\underline{x}, \underline{y}) \mapsto \underline{x}^t \cdot A \cdot \underline{y} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i a_{ij} y_j$$

is a bilinear form.
Moreover, $b_A$ is symmetric if and only if $A^t = A$, i. e. if $A$ is symmetric.

**Proof:** If $b_A$ is symmetric, then $a_{ij} = b_A(e_i, e_j) = b_A(e_j, e_i) = a_{ji}$ and $A$ is symmetric.
On the other hand, if $A = A^t$, then

$$b_A(\underline{x}, \underline{y}) = \underline{x}^t \cdot A \cdot \underline{y} = \underline{x}^t \cdot A^t \cdot \underline{y} = \left(\underline{x}^t \cdot A^t \cdot \underline{y}\right)^t = \underline{y}^t \cdot A \cdot \underline{x} = b_A(\underline{y}, \underline{x})$$

and thus $b_A$ is symmetric. □

[10] I. e., if $b$ is linear in the first component and linear in the second component.

**2.3 Definition**
Let $B = (v_1, \dots, v_n)$ be a basis of $V$. We call the matrix

$$M_B(b) = \big(b(v_i, v_j)\big)_{i,j=1,\dots,n} \in \mathrm{Mat}(n \times n, K)$$

the *matrix associated to* $b$ or the *matrix representation* of $b$ with respect to $B$.

**2.4 Example**
Let $V = \mathbb{R}^2$ and $b = \langle \cdot, \cdot \rangle$ be the standard scalar product. Let $E$ be the standard basis of $\mathbb{R}^2$ and $B = \big((1,1)^t, (1,0)^t\big)$ another basis. Then

$$M_E(b) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad M_B(b) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

**2.5 Proposition**
*Let $B = (v_1, \dots, v_n)$ be a basis of $V$. The map*

$$M_B : \mathrm{Bil}_K(V) \to \mathrm{Mat}(n \times n, K) : b \mapsto M_B(b)$$

*is a bijection.*
*Moreover, $b$ is symmetric if and only if $M_B(b)$ is symmetric.*
*Note also, for $v, w \in V$ we have*

$$b(v, w) = M_B(v)^t \cdot M_B(b) \cdot M_B(w) = b_{M_B(b)}\big(M_B(v), M_B(w)\big).$$

**Proof:**

***Claim*:** $M_B$ is injective.
Let $b, b' \in \mathrm{Bil}_K(V)$ such that $M_B(b) = M_B(b')$. Then for all $i, j = 1, \dots, n$ we have

$$b(v_i, v_j) = b'(v_i, v_j).$$

Let $v = \sum_{i=1}^n \lambda_i v_i \in V$ and $w = \sum_{j=1}^n \mu_j v_j \in V$ be given, then

$$b(v, w) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j b(v_i, v_j) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j b'(v_i, v_j) = b'(v, w).$$

Thus $b = b'$, and $M_B$ is injective.

***Claim*:** $M_B$ is surjective.
Let $A \in \mathrm{Mat}(n \times n, K)$ be given. Define

$$b : V \times V \to K : (v, w) \mapsto b_A\big(M_B(v), M_B(w)\big) = M_B(v)^t \cdot A \cdot M_B(w). \qquad (10)$$

Since $M_B(\lambda v + \lambda' v') = \lambda M_B(v) + \lambda' M_B(v')$, the map $b$ is actually bilinear, and since

$$b(v_i, v_j) = b_A(e_i, e_j) = a_{ij},$$

its matrix representation is $M_B(b) = A$. Thus $M_B$ is surjective.

***Claim*:** $b$ is symmetric if and only if $M_B(b)$ is symmetric.
By Equation (10) we know

$$b(v, w) = b_{M_B(b)}\big(M_B(v), M_B(w)\big).$$

However, applying Example 2.2 c., we have $b$ is symmetric if and only if $b_{M_B(b)}$ is symmetric if and only if $M_B(b)$ is symmetric. □

**2.6 Remark**
We have just shown that once we have fixed a basis of $V$ then symmetric bilinear forms and symmetric $n \times n$-matrices are virtually the same thing!

**Aim:** Given $b \in \mathrm{Bil}_K(V)$ symmetric, find a basis $B$ of $V$ such that $M_B(b)$ has a simple form!

**2.7 Proposition (Base Change)**
*Let* $B$ *and* $B'$ *be two bases of* $V$ *and let* $b \in \mathrm{Bil}_K(V)$. *Then*

$$M_{B'}(b) = \left(T_{B'}^{B}\right)^t \cdot M_B(b) \cdot T_{B'}^{B}.$$

*This allows us to define the* rank *of the bilinear form to be* $\mathrm{rank}(b) = \mathrm{rank}\left(M_B(b)\right)$, *and this number is independent of the chosen basis* $B$

**Proof:** Let's denote the matrix on the right hand side by $(a_{ij})_{i,j=1,\dots,n}$.
Let $B = (v_1, \dots, v_n)$, $B' = (v_1', \dots, v_n')$ and suppose $v_j' = \sum_{i=1}^{n} t_{ij} v_i$, i. e. $T_{B'}^{B} = (t_{ij})_{i,j=1,\dots,n}$. Then

$$b(v_i', v_j') = b\left(\sum_{k=1}^{n} t_{ki} v_k, \sum_{l=1}^{n} t_{lj} v_l\right) = \sum_{k=1}^{n}\sum_{l=1}^{n} t_{ki} t_{lj} b(v_k, v_l)$$

$$= (t_{1i} \dots t_{ni}) \cdot \left(b(v_k, v_l)\right)_{k,l=1,\dots,n} \cdot (t_{1j} \dots t_{nj})^t = a_{ij},$$

since $(t_{1i} \dots t_{ni})$ is the $i$-th row of $\left(T_{B'}^{B}\right)^t$ and $(t_{1j} \dots t_{nj})^t$ is the $j$-th column of $T_{B'}^{B}$.
Note, that the rank of a matrix is not changed when the matrix is multiplied by invertible matrices. Thus the rank of $b$ as defined does not depend on the chosen basis $B$. □

**2.8 Example (Example 2.4 continued)**
We have

$$(1,0)^t = 0 \cdot (1,1)^t + 1 \cdot (1,0)^t \text{ and } (0,1)^t = 1 \cdot (1,1)^t + (-1) \cdot (1,0)^t,$$

hence the base change matrix $T_E^B$ is

$$T_E^B = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Using the results in Example 2.4 we may verify the result of Proposition 2.7 in this example:

$$M_E(b) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \left(T_E^B\right)^t \cdot M_B(b) \cdot T_E^B.$$

**2.9 Remark**

a. If we define for $A, B \in \mathrm{Mat}(n \times n, K)$ symmetric

$$A \sim B \quad :\Longleftrightarrow \quad \exists\, T \in \mathrm{Gl}_n(K) \;:\; T^t \cdot A \cdot T,$$

then we have defined an equivalence relation on the set of all symmetric $n \times n$-matrices.

**Aim:** Find in each equivalence class a representative of a simple form!

Due to the above remarks this is the same thing as finding for a given symmetric matrix $A$ a basis $B$ of $K^n$ such that $M_B(b_A)$ has a simple form.

b. If $T \in Gl_n(K)$ and $\underline{t}_i$ denotes the $i$-th column of $T$ and if the matrix $T^t \cdot A \cdot T = (c_{ij})_{i,j=1\ldots,n}$, then

$$b_A(\underline{t}_i, \underline{t}_j) = \underline{t}_i^t \cdot A \cdot \underline{t}_j = c_{ij},$$

that is, the $ij$-th entry of $T^t \cdot A \cdot T$ is just the bilinear form $b_A$ evaluated at the $i$-th and $j$-th column of $T$!
In particular $b_A(e_i, e_j) = a_{ij}$, when $A = (a_{ij})_{i,j=1,\ldots,n}$.

**2.10 Theorem (Normal Form of Symmetric Bilinear Forms & Matrices)**

a. *Let $b \in \mathrm{Bil}_K(V)$ be symmetric, then there is basis $B = (v_1, \ldots, v_n)$ of $V$ such that $M_B(b)$ is a diagonal matrix, i. e. $b(v_i, v_j) = 0$ if $i \neq j$.*

b. *Let $A \in \mathrm{Mat}(n \times n, K)$ be symmetric, then there is a $T \in Gl_n(K)$ such that $T^t \cdot A \cdot T$ is a diagonal matrix.*

*We call such a diagonal matrix then a* normal form *for $b$ resp. $A$.*

**Proof:** a. We do the proof by induction on $n = \dim_K(V)$, where in the case $n = 1$ there is nothing to show, since $1 \times 1$-matrices are by default diagonal.
Let's now assume that $n > 1$ and that we have already proved the result for $n-1$-dimensional vector spaces.
If $b(v, v) = 0$ for all $v \in V$, then for arbitrary $v, w \in V$ we have

$$0 = b(v+w, v+w) - b(v, v) - b(w, w) = b(v, w) + b(w, v) = 2 \cdot b(v, w), \quad (11)$$

and hence $b(v, w) = 0$ for all $v, w \in V$, since $\mathrm{char}(K) \neq 2$. Then, however, $M_B(b)$ is the zero matrix for any basis, and in particular it is diagonal.
We may therefore assume that there is some $v \in V$ such that $b(v, v) \neq 0$. Set $U = \langle v \rangle$ and $U^\perp = \{u \in V \mid b(v, u) = 0\}$. By Exercise 3 on Assignment Set 7 we know that $U^\perp$ is a subspace such that $V = U + U^\perp$.

***Claim*:** $U \cap U^\perp = \{0\}$.

Let $u \in U \cap U^\perp$. Then there is a $\lambda \in K$ such that $u = \lambda v$ and

$$0 = b(v, u) = \lambda \cdot b(v, v).$$

However, since $b(v, v) \neq 0$, $\lambda$ must be zero, and hence $u = 0$.
This shows in particular that $\dim_K\left(U^\perp\right) = n - 1$, and therefore we may apply induction to the bilinear form

$$b_| : U^\perp \times U^\perp \to K : (u, w) \mapsto b(u, w).$$

Hence, there is a basis $B' = (v_1, \ldots, v_{n-1})$ of $U^\perp$ such that $b(v_i, v_j) = 0$ for all $i \neq j$. But then $B = (v_1, \ldots, v_n)$ with $v_n = v$ is a basis of $V$ and we have for all $i, j = 1, \ldots, n$ with $i \neq j$

$$b(v_i, v_j) = 0.$$

b. By Part a. there is a basis $B$ of $K^n$ such that $M_B(b_A)$ is a diagonal matrix. Set $T = T_B^E \in Gl_n(K)$ where $E$ is the standard basis of $K^n$, in other words let the vectors in $B$ be the columns of $T$, then

$$M_B(b_A) = \left(T_B^E\right)^t \cdot M_E(b) \cdot T_B^E = T^t \cdot A \cdot T.$$

□

Note that this proof gives a recursive algorithm for finding a basis of $V$ resp. a transformation matrix $T$ which diagonalises $b$ resp. $A$!

**2.11 Corollary (Theorem of Sylvester)**
*Let* $b \in \mathrm{Bil}_K(V)$ *and* $A \in \mathrm{Mat}(n \times n, K)$ *both be symmetric and of rank* $r$.

**a.** $\underline{K = \mathbb{C}}$**:** *There exists a basis* $B$ *of* $V$ *such that*

$$M_B(b) = \mathbb{1}_r \oplus 0_{n-r} = \left(\begin{array}{c|c} \mathbb{1}_r & 0_{r\times n-r} \\ \hline 0_{n-r\times r} & 0_{n-r} \end{array}\right)$$

*and there is some* $T \in Gl_n(\mathbb{C})$ *such that*

$$T^t \cdot A \cdot T = \mathbb{1}_r \oplus 0_{n-r} = \left(\begin{array}{c|c} \mathbb{1}_r & 0_{r\times n-r} \\ \hline 0_{n-r\times r} & 0_{n-r} \end{array}\right).$$

**b.** $\underline{K = \mathbb{R}}$**:** *There exists a basis* $B$ *of* $V$ *such that*

$$M_B(b) = \mathbb{1}_s \oplus -\mathbb{1}_t \oplus 0_{n-r} = \left(\begin{array}{c|c|c} \mathbb{1}_s & 0_{s\times t} & 0_{s\times n-r} \\ \hline 0_{t\times s} & -\mathbb{1}_t & 0_{s\times n-r} \\ \hline 0_{n-r\times s} & 0_{n-r\times t} & 0_{n-r} \end{array}\right)$$

*and there is some* $T \in Gl_n(\mathbb{R})$ *such that*

$$T^t \cdot A \cdot T = \mathbb{1}_s \oplus -\mathbb{1}_t \oplus 0_{n-r} = \left(\begin{array}{c|c|c} \mathbb{1}_s & 0_{s\times t} & 0_{s\times n-r} \\ \hline 0_{t\times s} & -\mathbb{1}_t & 0_{s\times n-r} \\ \hline 0_{n-r\times s} & 0_{n-r\times t} & 0_{n-r} \end{array}\right),$$

*where* $r = s + t = \mathrm{rank}\left(M_B(b)\right)$ *resp.* $r = s + t = \mathrm{rank}(A)$.
$s$ *is called the* index *of* $b$ *resp. of* $A$, $s - t$ *its* signature. *Both* $s$ *and* $t$ *are uniquely determined by* $b$ *resp. by* $A$.

**Proof:** It suffices to consider bilinear forms, since the result for matrices follows as in Theorem 2.10.

a. By Theorem 2.10 there is a basis $B' = (v_1', \ldots, v_n')$ of $V$ such that $b(v_i', v_j') = 0$ for all $i \neq j$. W. l. o. g. we may assume

$$b(v_i', v_i') \quad \begin{cases} \neq 0, & i = 1, \ldots, r, \\ = 0, & i = r+1, \ldots, n. \end{cases}$$

Choose for $i = 1, \ldots, r$ some square root $\sqrt{b(v_i', v_i')} \in \mathbb{C}$ of $b(v_i', v_i')$ and define

$$v_i = \begin{cases} \frac{1}{\sqrt{b(v_i', v_i')}} \cdot v_i', & i = 1, \ldots, r, \\ v_i', & i = r+1, \ldots, n. \end{cases}$$

Then $B = (v_1, \ldots, v_n)$ is a basis of $V$ such that $M_B(b)$ has the desired form.

b. By Theorem 2.10 there is a basis $B' = (v'_1, \dots, v'_n)$ of $V$ such that $b(v'_i, v'_j) = 0$ for all $i \neq j$. W. l. o. g. we may assume

$$b(v'_i, v'_i) \begin{cases} > 0, & i = 1, \dots, s, \\ < 0, & i = s+1, \dots, s+t, \\ = 0, & i = s+t+1, \dots, n. \end{cases}$$

Let's define

$$v_i = \begin{cases} \frac{1}{\sqrt{b(v'_i, v'_i)}} \cdot v'_i, & i = 1, \dots, s, \\ \frac{1}{\sqrt{-b(v'_i, v'_i)}} \cdot v'_i, & i = s+1, \dots, s+t, \\ v'_i, & i = s+t+1, \dots, n. \end{cases}$$

Then $B = (v_1, \dots, v_n)$ is a basis of $V$ such that $M_B(b)$ has the desired form. It remains to show that $s$ and $t$ are uniquely determined. Note first, that obviously $r = s + t = \operatorname{rank}\big(M_B(b)\big)$ and this rank is independent of the chosen basis by Proposition 2.7.

**Claim:** $s = \max\left\{ \dim_K(U) \;\middle|\; U \leq V, b(v,v) > 0 \;\forall\, 0 \neq v \in U \right\}$, and thus in particular $s$ depends only on $b$, and not on the chosen basis $B$.

By choice, the subspace $U = \langle v_1, \dots, v_s \rangle \leq V$ satisfies for $0 \neq \sum_{i=1}^{s} \lambda_i v_i \in U$

$$b\left( \sum_{i=1}^{s} \lambda_i v_i, \sum_{j=1}^{s} \lambda_j v_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j b(v_i, v_j) = \sum_{i=1}^{n} \lambda_i^2 b(v_i, v_i) > 0.$$

Thus $s$ is at most the maximum on the right hand side.
Set $W = \langle v_{s+1}, \dots, v_n \rangle$. In the same way as above we see that for $w \in W$

$$b(w, w) \leq 0.$$

Let $U' \leq V$ such that $b(v,v) > 0$ for all $0 \neq v \in U'$. Then $U' \cap W = \{0\}$, and hence

$$\begin{aligned} \dim_K(U') &= \dim_K(U' + W) - \dim_K(W) + \dim_K(U' \cap W) \\ &\leq n - \big(t + (n - t - s)\big) = s. \end{aligned}$$

Thus $s$ is also at least the maximum on the right hand side, which proves the claim.
If however $s$ and $r$ only depend on $b$, then $t = r - s$ does so as well.

□

**2.12 Example (Symmetric Gauß Algorithm)**
Recall first that any invertible matrix $T$ is the product of elementary matrices

$$T = P_1 \cdots P_k,$$

where an elementary matrix $P_i$ corresponds to performing one of the elementary operations in the Gauß algorithm, i. e. permuting rows or columns resp. adding multiples of rows or columns to each other. Recall moreover, that multiplying with an elementary $P$ from the right is a column operation, while multiplying with the matrix $P^t$ from left is the *corresponding* row operation!

Thus the relation

$$\begin{aligned} B &= T^t \cdot A \cdot T \\ &= P_k^t \cdots P_1^t \cdot A \cdot P_1 \cdots P_k \end{aligned}$$

says that $A$ can be transformed into $B$ by successively performing row/column operations, where each row operation is immediately also performed as column operation! This gives an algorithm to find a normal form for a symmetric $A$.
Let for example $A = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right) \in \mathrm{Mat}(2, \mathbb{R})$. Then

$$A = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right) \xrightarrow{R:II \mapsto II - \frac{1}{2}I} \left(\begin{smallmatrix} 2 & 1 \\ 0 & \frac{1}{2} \end{smallmatrix}\right) \xrightarrow{C:II \mapsto II - \frac{1}{2}I} \left(\begin{smallmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{smallmatrix}\right) \xrightarrow{R/C:I \mapsto \frac{1}{\sqrt{2}}I} \left(\begin{smallmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{smallmatrix}\right) \xrightarrow{R/C:II \mapsto \sqrt{2}II} \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right).$$

Thus the signature of $A$, its index and its rank are all 2.

**2.13 Definition**

a. *A homogeneous polynomial of degree* 2

$$q = \sum_{i=1}^{n} q_{ii} x_i^2 + 2 \cdot \sum_{i<j} q_{ij} x_i x_j \in K[x_1, \ldots, x_n]$$

*is called a* quadratic form. *By* $K[x_1, \ldots, x_n]_2$ *we denote the set of all such polynomials.*
*Note, since* $\mathrm{char}(K) \neq 2$, *every homogeneous polynomial of degree* 2 *has this form!*

b. *Let* $b \in \mathrm{Bil}_K(V)$ *be symmetric. We call the map*

$$q_b : V \to K : v \mapsto b(v, v)$$

*the* quadratic form associated to $b$.
*Note, if* $M_B(b) = (a_{ij})_{i,j=1,\ldots,n}$ *and* $M_B(v) = (y_1, \ldots, y_n)^t$ *for some basis* $B$ *of* $V$, *then*

$$q_b(v) = (y_1, \ldots, y_n) \cdot (a_{ij})_{i,j=1,\ldots,n} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} a_{ii} y_i^2 + 2 \cdot \sum_{i<j} a_{ij} y_i y_j.$$

*Thus, once we have fixed a basis* $B$ *of* $V$, $q_b$ *is a homogeneous polynomial function of degree* 2 *in the coordinates w. r. t.* $B$.
*We define for* $B = (v_1, \ldots, v_n)$

$$M_B(q_b) = \sum_{i=1}^{n} \sum_{j=1}^{n} b(v_i, v_j) \cdot x_i x_j = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \cdot \sum_{i<j} a_{ij} x_i x_j \in K[x_1, \ldots, x_n]$$

*and call this the* basis representation *of* $q_b$ *with respect to* $B$.

**2.14 Example (Example 2.8 continued)**
We have calculated $M_B(b) = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, and hence

$$M_B(q_b) = 2x_1^2 + x_2^2 + 2x_1 x_2.$$

**2.15 Proposition**

a. *Let* $b \in \mathrm{Bil}_K(V)$ *be symmetric. Then for* $v, w \in V$

$$b(v,w) = \tfrac{1}{2} \cdot \big(q_b(v+w) - q_b(v) - q_b(w)\big).$$

*In particular, the bilinear form* $b$ *is uniquely determined by its associated quadratic form.*

b. *Let* $q \in K[x_1, \ldots, x_n]$ *be a quadratic form, and let* $B$ *be a basis of* $V$. *Then there is a symmetric* $b \in \mathrm{Bil}_K(V)$ *such that*

$$q = M_B(q_b).$$

c. *The map*

$$\{b \in \mathrm{Bil}_K(V) \mid b \textit{ symmetric}\} \longrightarrow K[x_1, \ldots, x_n]_2 : b \mapsto M_B(q_b)$$

*is bijective.*

**Proof:** a. This is just Equation (11) in the Proof of Theorem 2.10.

b. Let $q = \sum_{i=1}^n q_{ii}x_i^2 + 2 \cdot \sum_{i<j} q_{ij}x_ix_j$ be given, and set $A = (q_{ij})_{i,j=1,\ldots,n}$ with $q_{ji} := q_{ij}$ for $i < j$. Then $A \in \mathrm{Mat}(n \times n, K)$ is symmetric and by Proposition 2.5 there is a (unique) symmetric bilinear form $b \in \mathrm{Bil}_K(V)$ such that $M_B(b) = A$, which implies

$$M_B(q_b) = q.$$

c. Part a. gives the injectivity, and Part b. the surjectivity.

□

**2.16 Corollary (Normal Forms of Quadratic Forms)**
*Let* $b \in \mathrm{Bil}_K(V)$ *be symmetric.*

a. *There is a basis* $B$ *of* $V$ *such that* $M_B(q_b) = \sum_{i=1}^n a_i x_i^2$ *with* $a_i = q_b(v_i)$.

b. *If* $K = \mathbb{C}$, *then there is a basis* $B$ *of* $V$ *such that* $M_B(q_b) = \sum_{i=1}^r x_i^2$, *where* $r = \mathrm{rank}(b)$.

c. *If* $K = \mathbb{R}$, *then there is a basis* $B$ *of* $V$ *such that* $M_B(q_b) = \sum_{i=1}^s x_i^2 - \sum_{i=s+1}^r x_i^2$, *where* $s = \mathrm{index}(b)$ *and* $r = \mathrm{rank}(b)$.

**Proof:** This follows right away from Theorem 2.10 and Corollary 2.11. □

# 3 Normal Forms of Orthogonal, Unitary and Self-Adjoint Endomorphisms and Matrices

**3.0 General Assumptions** Throughout this section $\mathbb{K} = \mathbb{R}$, the field of real numbers, or $\mathbb{K} = \mathbb{C}$, the field of complex numbers. By

$$\bar{\cdot} : \mathbb{K} \to \mathbb{K} : \lambda \mapsto \overline{\lambda}$$

we denote the complex conjugation, and if $\lambda \in \mathbb{R}$, then of course $\overline{\lambda} = \lambda$. $V$ will be a finite-dimensional $\mathbb{K}$-*vector space*.

**3.1 Definition**

A *scalar product* on $V$ is a map $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$ such that

(i) for $v, w, u \in V$ and $\lambda, \mu \in \mathbb{K}$ we have

$$\langle \lambda v + \mu w, u\rangle = \lambda \cdot \langle v, u\rangle + \mu \cdot \langle w, u\rangle$$

and

$$\langle u, \lambda v + \mu w\rangle = \overline{\lambda} \cdot \langle u, v\rangle + \overline{\mu} \cdot \langle u, w\rangle.$$

(ii) $\langle v, w\rangle = \overline{\langle w, v\rangle}$ for $v, w \in V$.

(iii) $\langle v, v\rangle > 0$ for $0 \neq v \in V$.

The first property is called the *sesqui-linearity* of the scalar product, the second property is called its *anti-symmetry* and due to the third property it is said to be *definite*.

**3.2 Remark**

If $\mathbb{K} = \mathbb{R}$, then a scalar product is just a definite symmetric bilinear form.

**3.3 Example**

a. **(Standard Scalar Product)** Let $V = \mathbb{K}^n$, then the map

$$\langle\cdot,\cdot\rangle : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K} : (\underline{x}, \underline{y}) \mapsto \underline{x}^t \cdot \overline{\underline{y}} = \sum_{i=1}^{n} x_i \cdot \overline{y_i}$$

is a scalar product, the so called *standard scalar product*.

b. Let $V = \mathbb{R}[x]_{<n} = \{p \in \mathbb{R}[x] \mid \deg(p) < n\}$. The map

$$\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R} : (p, q) \mapsto \int_0^1 p(x) \cdot q(x)\,dx$$

defines a scalar product on $V$, due to the rules for integrals.

**3.4 Definition**

a. *A tuple* $\big(V, \langle\cdot,\cdot\rangle\big)$ *consisting of a finite-dimensional* $\mathbb{K}$*-vector space* $V$ *and a scalar product* $\langle\cdot,\cdot\rangle$ *is called a (finite-dimensional)* Hilbert space. *If* $\mathbb{K} = \mathbb{R}$*, one calls it also* Euclidean space.

b. *If* $\big(V, \langle\cdot,\cdot\rangle\big)$ *is a Hilbert space and* $B = (v_1, \ldots, v_n)$ *is a basis of* $V$ *such that*

$$\langle v_i, v_j\rangle = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

*then* $B$ *is called an* orthonormal basis *(ONB) of* $V$.

c. *If* $\big(V, \langle\cdot,\cdot\rangle\big)$ *is a Hilbert space and* $v \in V$*, then we define*

$$\|v\| = \sqrt{\langle v, v\rangle}$$

*and call this the* length *or the* norm *of* $v$.

d. *If* $\big(V, \langle\cdot,\cdot\rangle\big)$ *is a Hilbert space and* $U \leq V$ *a subspace, then we call*

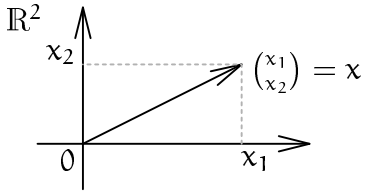
$$U^\perp = \{v \in V \mid \langle v, u\rangle = 0 \ \forall\, u \in U\}$$

*the* orthogonal complement *of* $U$.

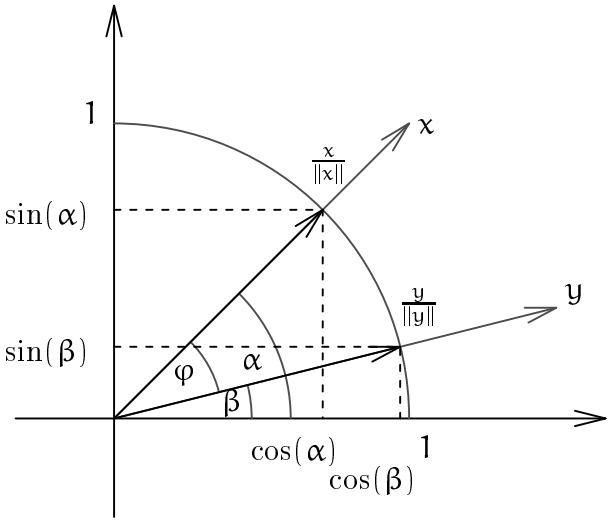
**3.5 Example (Explanation for the Notion ONB)**

Let $V = \mathbb{R}^2$ and $\langle \cdot, \cdot \rangle$ be the standard scalar product.

If $x \in \mathbb{R}^2$ is some vector, then by the Theorem of Pythagoras the length of $x$ is indeed just $\|x\| = \sqrt{x_1^2 + x_2^2}$.




And if $x \in \mathbb{R}^2$ and $y \in \mathbb{R}^2$ are two vectors in the plane, then some geometrical observations lead to an algorithm for calculating the angle $\measuredangle(x, y)$ between these two vectors.




For this we scale the vectors so that they have length one by dividing them by their length, i. e. we consider the vectors $\frac{x}{\|x\|}$ and $\frac{y}{\|y\|}$. Using the notation in the above plan we have

$$\measuredangle(x, y) = \measuredangle\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right) = \alpha - \beta = \varphi.$$

Using the theorems from trigonometry we have

$$\begin{aligned} \cos(\varphi) &= \cos(\alpha - \beta) \\ &= \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) \\ &= \frac{x_1 y_1 + x_2 y_2}{\|x\| \cdot \|y\|} \\ &= \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \end{aligned}$$

or alternatively

$$\measuredangle(x, y) = \varphi = \arccos\left(\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}\right).$$

In particular, $x$ and $y$ are orthogonal to each other if and only if $\cos(\varphi) = 0$ if and only if $\langle x, y \rangle = 0$.

We have thus seen, that the standard scalar product determines angles, lengths and thus distances in $\mathbb{R}^2$, and we may therefore use scalar products in general in order to generalise these properties.

*From now on we will assume that $\big(V, \langle\cdot,\cdot\rangle\big)$ is a Hilbert space, i. e. that $V$ is endowed with a fixed scalar product.*

The following Lemma tells us how to find the base representation of a vector with respect to an ONB without having to solve a system of linear equations.

**3.6 Lemma (Parseval-Equation)**

*Let $B = (v_1, \ldots, v_n)$ be an ONB of the Hilbert space $\big(V, \langle\cdot,\cdot\rangle\big)$ and let $v \in V$, then*

$$v = \sum_{i=1}^{n} \langle v, v_i\rangle \cdot v_i.$$

**Proof:** Since $B$ is a basis, there are unique elements $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ such that $v = \sum_{i=1}^{n} \lambda_i v_i$. Using the scalar product and the fact that $B$ is an ONB we find

$$\langle v, v_j\rangle = \left\langle \sum_{i=1}^{n} \lambda_i v_i, v_j \right\rangle = \sum_{i=1}^{n} \lambda_i \cdot \langle v_i, v_j\rangle = \lambda_j.$$

□

**3.7 Lemma**

*If $v_1, \ldots, v_r \in V$ such that $\langle v_i, v_j\rangle = \delta_{ij}$ for $i, j = 1, \ldots, r$, then $(v_1, \ldots, v_r)$ is linearly independent.*

**Proof:** Let $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ such that $\sum_{i=1}^{r} \lambda_i v_i = 0$. Then for $j = 1, \ldots, r$

$$0 = \left\langle \sum_{i=1}^{r} \lambda_i v_i, v_j \right\rangle = \sum_{i=1}^{r} \lambda_i \cdot \langle v_i, v_j\rangle = \lambda_j.$$

Hence, $(v_1, \ldots, v_r)$ is linearly independent. □

**3.8 Theorem (Gram-Schmidt)**

*Let $U \leq V$ be a subspace of the Hilbert space $\big(V, \langle\cdot,\cdot\rangle\big)$, then any ONB of $U$ can be extended to an ONB of $V$. In particular, every Hilbert space has an ONB.*
*Moreover, $V = U \oplus U^{\perp}$.*

**Proof:** Let $B = (v_1, \ldots, v_r)$ be an ONB of $U$, $n = \dim_{\mathbb{K}}(V)$ and $m = n - r$. We do the proof by induction on $m$. If $m = 0$, then $n = r$ and hence $U = V$, so that $B$ is already an ONB of $V$.

Suppose now that $m > 0$ and that the statement holds true for $m - 1$, i. e. for subspaces of dimension $r + 1$. By assumption $r < n$, and hence there is some $v \in V \setminus U$. We set

$$v' = v - \sum_{i=1}^{r} \langle v, v_i\rangle \cdot v_i \neq 0$$

and

$$v_{r+1} = \frac{1}{||v'||} \cdot v'.$$

Then $\langle v_{r+1}, v_{r+1}\rangle = \frac{\langle v', v'\rangle}{||v||^2} = 1$ and

$$\langle v_{r+1}, v_i\rangle = \frac{\langle v, v_i\rangle - \sum_{j=1}^{r} \langle v, v_j\rangle \cdot \langle v_i, v_j\rangle}{||v'||} = \frac{\langle v, v_i\rangle - \langle v, v_i\rangle}{||v'||} = 0.$$

Hence, by Lemma 3.7 $(v_1, \dots, v_{r+1})$ is an ONB of the subspace $U' = \langle v_1, \dots, v_{r+1} \rangle$, which has dimension $r+1$. So by induction $(v_1, \dots, v_{r+1})$ can be extended to an ONB of $V$. Finally, it is Exercise 4 on Assignment Set 8 to show $V = U \oplus U^\perp$. □

**3.9 Example**

Let $V = \mathbb{K}^3$ and $\langle \cdot, \cdot \rangle$ be the standard scalar product.

a. The standard basis $E = (e_1, e_2, e_3)$ fulfils $\langle e_i, e_j \rangle = \delta_{ij}$, and is thus an ONB of $\big(\mathbb{K}^3, \langle \cdot, \cdot \rangle\big)$.

b. Let $U = \big\langle (2,1,2)^t, (3,1,1)^t \big\rangle \leq \mathbb{K}^3$, $u_1 = (2,1,2)^t$ and $u_2 = (3,1,1)^t$.
Let's first of all find an ONB of $U$, using the algorithm of Gram-Schmidt.

***Step 1*:** $v_1 = \frac{1}{\|u_1\|} \cdot u_1 = \frac{1}{3} \cdot (2,1,2)^t$.

***Step 2*:** Set

$$v' = u_2 - \langle u_2, v_1 \rangle \cdot v_1 = (3,1,1)^t - \tfrac{9}{3} \cdot \tfrac{1}{3} \cdot (2,1,2)^t = (1,0,-1)^t,$$

and then

$$v_2 = \frac{1}{\|v_2\|} \cdot v_2 = \frac{1}{\sqrt{2}} \cdot (1,0,-1)^t.$$

Then $B = (v_1, v_2)$ is an ONB of $U$.
Let us now extend $B$ to an ONB of $V$.

***Step 3*:** For this we choose some vector $u_3 = (1,0,0)^t \notin U$. We then set

$$\begin{aligned} v' &= u_3 - \langle u_3, v_1 \rangle \cdot v_1 - \langle u_3, v_2 \rangle \cdot v_2 \\ &= (1,0,0)^t - \tfrac{2}{9} \cdot (2,1,2)^t - \tfrac{1}{2} \cdot (1,0,-1)^t \\ &= \tfrac{1}{18} \cdot (1,-4,1), \end{aligned}$$

and hence

$$v_3 = \tfrac{1}{\|v'\|} \cdot v' = \tfrac{1}{3\sqrt{2}} \cdot (1,-4,1)^t.$$

Then $(v_1, v_2, v_3)$ is an ONB of $\mathbb{K}^3$ which extends an ONB of $U$.

**3.10 Idea**

When we consider Hilbert spaces, that is vector spaces together with the additional structure of a scalar product, then we should like to restrict our attention to maps from $V$ to $V$ which respect the structure, i. e. maps $f : V \to V$ which are $\mathbb{K}$-linear and which respect the scalar product.
However, what does it mean that an endomorphism respects the scalar product?
We will give two different interpretations of this in the following definition, both of which make sense and lead to interesting classes of endomorphisms.
As always, we will treat the case of square matrices at the same time.

**3.11 Definition**

Let $f \in \mathrm{End}_{\mathbb{K}}(V)$ and $A \in \mathrm{Mat}(n \times n, \mathbb{K})$.

a. If $\langle f(v), f(w) \rangle = \langle v, w \rangle$ for all $v, w \in V$, then $f$ is said to be *orthogonal* (if $\mathbb{K} = \mathbb{R}$) or *unitary* (if $\mathbb{K} = \mathbb{C}$).

b. If $\langle f(v), w \rangle = \langle v, f(w) \rangle$ for all $v, w \in V$, then $f$ is said to be *self-adjoint*. If $\mathbb{K} = \mathbb{R}$ we say also $f$ is *symmetric*, and if $\mathbb{K} = \mathbb{C}$ we say likewise $f$ is *hermitian*.

c. If $A \in \mathrm{Gl}_n(K)$ and $A^{-1} = \overline{A}^t$, then $A$ is called *orthogonal* (if $\mathbb{K} = \mathbb{R}$) or *unitary* (if $\mathbb{K} = \mathbb{C}$). We set $\mathcal{O}(n) = \{B \in \mathrm{Mat}(n \times n, \mathbb{R}) \mid A \text{ is orthogonal}\}$ and $\mathcal{U}(n) = \{B \in \mathrm{Mat}(n \times n, \mathbb{C}) \mid A \text{ is unitary}\}$. These are subgroups of $\mathrm{Gl}_n(\mathbb{R})$ resp. of $\mathrm{Gl}_n(\mathbb{C})$, as one easily verifies.

d. If $A = \overline{A}^t$, then $A$ is called *self-adjoint*, in the real case also *symmetric* and in the complex case also *hermitian*.

**3.12 Example**

Let $V = \mathbb{K}^n$ and let $\langle \cdot, \cdot \rangle$ be the standard scalar product.

a. Let $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ and let $\underline{a}_1, \ldots, \underline{a}_n$ denote the columns of $A$. Then $A^{-1} = \overline{A}^t$ if and only if $\overline{A}^t \cdot A = \mathbb{1}$ if and only if $A^t \cdot \overline{A} = \mathbb{1}$ if and only if $\underline{a}_i^t \cdot \underline{a}_j = \delta_{ij}$ for all $i, j$ if and only if $(\underline{a}_1, \ldots, \underline{a}_n)$ is an ONB of $\mathbb{K}^n$.
This shows e. g. that the following matrix is orthogonal and self-adjoint

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

b. Let $A \in \mathrm{Gl}_n(\mathbb{K})$ be such that $A^{-1} = \overline{A}^t$. Then

$$f_A : \mathbb{K}^n \to \mathbb{K}^n : \underline{x} \mapsto A \cdot \underline{x}$$

is an orthogonal resp. unitary endomorphism.

**Proof:** Let $\underline{x}, \underline{y} \in \mathbb{K}^n$. Then

$$\langle f_A(\underline{x}), f_A(\underline{y}) \rangle = \left(A \cdot \underline{x}\right)^t \cdot \overline{A \cdot \underline{y}}$$
$$= \underline{x}^t \cdot A^t \cdot \overline{A} \cdot \underline{y} = \underline{x}^t \cdot \overline{\overline{A}^t \cdot A} \cdot \underline{y} = \underline{x}^t \cdot \underline{y} = \langle \underline{x}, \underline{y} \rangle.$$

□

c. Let $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ be self-adjoint, then $f_A$ is self-adjoint.

**Proof:** Let $\underline{x}, \underline{y} \in \mathbb{K}^n$. Then

$$\langle f_A(\underline{x}), \underline{y} \rangle = \left(A \cdot \underline{x}\right)^t \cdot \underline{y} = \underline{x}^t \cdot A^t \cdot \underline{y} = \underline{x}^t \cdot \left(\overline{A}^t\right)^t \cdot \underline{y} = \underline{x}^t \cdot \overline{A} \cdot \underline{y} = \langle \underline{x}, f_A(\underline{y}) \rangle.$$

□

d. Let $\left(V, \langle \cdot, \cdot \rangle\right)$ be any Hilbert space and let $B = (v_1, \ldots, v_n)$ and $B' = (v'_1, \ldots, v'_n)$ be two ONB. Then the base change matrix $T_{B'}^B$ is orthogonal resp. unitary.

**Proof:** Let $T_{B'}^B = (t_{ij})_{i,j=1,\ldots,n}$ and denote by $\underline{t}_i$ the $i$-th column of this matrix. Then $v'_j = \sum_{i=1}^n t_{ij} v_i$ and therefore

$$\langle \underline{t}_k, \underline{t}_l \rangle = \sum\nolimits_{i=1}^n t_{ik} \cdot \overline{t_{il}} = \sum\nolimits_{i=1}^n \sum\nolimits_{j=1}^n t_{ik} \cdot \overline{t_{jl}} \cdot \langle v_i, v_j \rangle$$
$$= \left\langle \sum\nolimits_{i=1}^n t_{ik} v_i, \sum\nolimits_{j=1}^n t_{jl} v_j \right\rangle = \langle v'_k, v'_l \rangle = \delta_{kl}.$$

□

**3.13 Proposition**

*Let* $g \in \mathrm{End}_{\mathbb{K}}(V)$ *be orthogonal resp. unitary. Then:*

a. $\|f(v)\| = \|v\|$ *for all* $v \in V$.

b. $\frac{\langle f(v), f(w)\rangle}{\|f(v)\| \cdot \|f(w)\|} = \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|}$ *for all* $v, w \in V$.

*Hence,* $f$ *preserves lengths, distances and angles.*

**Proof:** a. For $v \in V$ we have $\|f(v)\| = \sqrt{\langle f(v), f(v)\rangle} = \sqrt{\langle v, v\rangle} = \|v\|$.

b. This follows from Part a. and the definition of orthogonal resp. unitary.

□

**3.14 Proposition**

*Let* $B$ *be an* ONB *of* $V$, $f \in \mathrm{End}_{\mathbb{K}}(V)$. *Then:*

a. $f$ *is orthogonal resp. unitary if and only* $M_B^B(f)$ *is so.*

b. $f$ *is self-adjoint if and only if* $M_B^B(f)$ *is so.*

**Proof:** Let $B = (v_1, \ldots, v_n)$, then the Parseval-Equation 3.6 gives

$$f(v_j) = \sum_{i=1}^{n} \langle f(v_j), v_i\rangle \cdot v_i.$$

Hence $M_B^B(f) = (a_{ij})_{i,j=1,\ldots,n}$ with $a_{ij} = \langle f(v_j), v_i\rangle$. Let's denote the columns of $M_B^B(f)$ by $\underline{a}_1, \ldots, \underline{a}_n$.

a. We then find

$$\langle f(v_j), f(v_l)\rangle = \left\langle \sum_{i=1}^{n} a_{ij} v_i, \sum_{k=1}^{n} a_{kl} v_k \right\rangle$$
$$= \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ij} \cdot \overline{a_{kl}} \cdot \langle v_i, v_k\rangle = \sum_{i=1}^{n} a_{ij} \cdot \overline{a_{il}} = \langle \underline{a}_j, \underline{a}_l\rangle.$$

Taking Example 3.12 a. into account, we have $M_B^B(f)$ is orthogonal/unitary if and only if $\langle \underline{a}_j, \underline{a}_l\rangle = \delta_{jl}\ \forall\ j, l$ if and only if $\langle f(v_j), f(v_l)\rangle = \delta_{jl} = \langle v_j, v_l\rangle\ \forall\ j, l$. If $f$ is orthogonal/unitary, then the last condition is obviously satisfied and therefore $M_B^B(f)$ is orthogonal/unitary.
If conversely $M_B^B(f)$ is orthogonal/unitary, and $v = \sum_{j=1}^{n} \lambda_j v_j, w = \sum_{l=1}^{n} \mu_l v_l \in V$, then by the above equivalence we get

$$\langle f(v), f(w)\rangle = \sum_{j=1}^{n} \sum_{l=1}^{n} \lambda_j \cdot \overline{\mu_l} \cdot \langle f(v_j), f(v_l)\rangle = \sum_{j=1}^{n} \sum_{l=1}^{n} \lambda_j \cdot \overline{\mu_l} \cdot \langle v_j, v_l\rangle = \langle v, w\rangle.$$

b. Let $v = \sum_{i=1}^{n} \lambda_i v_i, w = \sum_{j=1}^{n} \mu_j v_j \in V$. Then

$$\langle f(v), w\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \cdot \overline{\mu_j} \cdot \langle f(v_i), v_j\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \cdot \overline{\mu_j} \cdot a_{ji}$$

and

$$\langle v, f(w)\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \cdot \overline{\mu_j} \cdot \langle v_i, f(v_j)\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \cdot \overline{\mu_j} \cdot \overline{a_{ij}}.$$

If, now, $M_B^B(f)$ is self-adjoint, then $\overline{a_{ij}} = a_{ji}$, and thus $f$ is self-adjoint.
If $f$ is self-adjoint, then we may apply the above inequalities to $v = v_i$ and $w = v_j$ for $i, j$ arbitrary, in order to find $a_{ji} = \overline{a_{ij}}$ for all $i, j$. Thus $M_B^B(f)$ is self-adjoint.

□

**3.15 Theorem (Normal Forms for Unitary Endomorphisms & Matrices)**

a. *If $f \in \mathrm{End}_{\mathbb{C}}(V)$ is unitary, then there is an ONB $B = (v_1, \ldots, v_n)$ of $V$ such that $f(v_i) = \lambda_i v_i$ and $|\lambda_i| = 1$, i. e.*

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}.$$

b. *If $A \in \mathcal{U}(n)$, then there is a $T \in \mathcal{U}(n)$ such that for some $\lambda_i \in \mathbb{C}$ with $|\lambda_i| = 1$*

$$T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}.$$

**Proof:** a.

***Claim*:** If $\lambda$ is an eigenvalue of $f$, then $|\lambda| = 1$.
By assumption there is some $0 \neq v \in V$ such that $f(v) = \lambda v$. Thus

$$\lambda \cdot \overline{\lambda} \cdot \langle v, v \rangle = \langle \lambda v, \lambda v \rangle = \langle f(v), f(v) \rangle = \langle v, v \rangle.$$

Since the scalar product is definite $\langle v, v \rangle \neq 0$, and hence $|\lambda|^2 = \lambda \cdot \overline{\lambda} = 1$.

***Claim*:** $f$ has an ONB of eigenvectors.

We do the proof by induction on $n = \dim_K(V)$, where for the case $n = 1$ there is nothing to show.
We may therefore assume that $n > 1$ and that unitary endomorphisms on Hilbert spaces of dimension $n - 1$ are diagonalisable w. r. t. an ONB.
Since $\mathbb{K} = \mathbb{C}$, the characteristic polynomial of $f$ factorises and, hence, $f$ has an eigenvalue $\lambda_n$ with corresponding eigenvector $0 \neq v_n \in V$ of length $\|v_n\| = 1$ and by the above claim $\lambda_n \cdot \overline{\lambda_n} = 1$. We set $U = \langle v_n \rangle$, and we show

$$f\big(U^{\perp}\big) \subseteq U^{\perp}.$$

For this let $u \in U^{\perp}$. Then

$$\begin{aligned} \langle f(u), v_n \rangle &= \lambda_n \cdot \overline{\lambda_n} \cdot \langle f(u), v_n \rangle \\ &= \lambda_n \cdot \langle f(u), \lambda_n v_n \rangle = \lambda_n \cdot \langle f(u), f(v_n) \rangle = \lambda_n \cdot \langle u, v_n \rangle = 0. \end{aligned}$$

Thus $f(u) \perp v_n$, and hence $f(u) \in U^{\perp}$.

We may therefore consider the endomorphism $f$ restricted to $U^\perp$

$$f_| : U^\perp \to U^\perp : u \mapsto f(u),$$

which by default is unitary again.
Thus by induction, since $\dim_K\left(U^\perp\right) = n-1$, there is an ONB $(v_1, \ldots, v_{n-1})$ of $U^\perp$ such that

$$f(v_i) = \lambda_i v_i$$

for $i = 1, \ldots, n-1$ and some $\lambda_i \in \mathbb{C}$ with $|\lambda_i| = 1$.
However, $\langle v_i, v_j\rangle = \delta_{ij}$ for all $i, j = 1, \ldots, n$, since $v_i \in \langle v_n\rangle^\perp$. Thus by Lemma 3.7 $B = (v_1, \ldots, v_n)$ is an ONB of $V$ and $f(v_i) = \lambda_i v_i$ with $|\lambda_i| = 1$ for all $i = 1, \ldots, n$.

b. We may apply Part a. to $V = \mathbb{C}^n$ with the standard scalar product and the endomorphism $f = f_A$. Then $T = T_B^E$ will do, where $B$ is the ONB which Part a. gives us, and $E$ is the standard basis. Note that by Proposition 3.14 $f$ is unitary and by Example 3.12 d. $T \in U(n)$.

□

**3.16 Remark (Normal Forms of Orthogonal Endomorphisms & Matrices)**
The case of orthogonal endomorphisms and matrices is considerably harder, due to the fact, that over $\mathbb{R}$ the characteristic polynomial need not factorise. However, one can show the following generalisation.
If $f \in \mathrm{End}_{\mathbb{R}}(V)$ is orthogonal resp. $A \in \mathcal{O}(n)$, then there is an ONB $B$ of $V$ resp. a $T \in \mathcal{O}(n)$ such that

$$M_B^B(f) = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & A_r \end{pmatrix} \quad \text{resp.} \quad T^{-1} \cdot A \cdot T = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & A_r \end{pmatrix}$$

where either $A_i = (1) \in \mathrm{Mat}(1\times 1, \mathbb{R})$ or $A_i = (-1) \in \mathrm{Mat}(1\times 1, \mathbb{R})$ or

$$A_i = \begin{pmatrix} \cos(\alpha_i) & \sin(\alpha_i) \\ -\sin(\alpha_i) & \cos(\alpha_i) \end{pmatrix}$$

for some $\alpha_i \in [0, 2\pi)$.

**3.17 Theorem (Normal Forms of Self-Adjoint Endomorphisms & Matrices)**

a. *Let $f \in \mathrm{End}_{\mathbb{K}}(V)$ be self-adjoint, then there is an ONB $B = (v_1, \ldots, v_n)$ of $V$ such that $f(v_i) = \lambda_i v_i$ with $\lambda_i \in \mathbb{R}$ for $i = 1, \ldots, n$, i. e.*

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \lambda_n \end{pmatrix}.$$

b. *If* $A \in \mathrm{Mat}(n \times n, \mathbb{K})$ *is self-adjoint, then there is a* $T \in \mathcal{O}(n)$ *resp.* $T \in \mathcal{U}(n)$ *such that*

$$T^{-1} \cdot A \cdot T = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}$$

*and* $\lambda_i \in \mathbb{R}$ *for* $i = 1, \dots, n$.

**Proof:**

***Claim*:** If $\lambda$ is an eigenvalue of $f$, then $\lambda \in \mathbb{R}$.
By assumption there is some $0 \neq v \in V$ such that $f(v) = \lambda v$. Thus

$$\lambda \cdot \langle v, v \rangle = \langle \lambda v, v \rangle = \langle f(v), v \rangle = \langle v, f(v) \rangle = \langle v, \lambda v \rangle = \overline{\lambda} \cdot \langle v, v \rangle.$$

Since the scalar product is definite $\langle v, v \rangle \neq 0$, and hence $\lambda = \overline{\lambda}$, i. e. $\lambda \in \mathbb{R}$.

***Claim*:** If $\lambda$ is an eigenvalue of $A$, then $\lambda \in \mathbb{R}$.

By assumption $\lambda$ is an eigenvalue of the self-adjoint endomorphism $f_A$, and thus by the above claim $\lambda \in \mathbb{R}$.

a.

***Claim*:** $f$ has some eigenvalue!
Let $B$ be any ONB of $V$ and $M = M_B^B(f)$, then by Proposition 3.14 $M = \overline{M}^t \in \mathrm{Mat}(n \times n, \mathbb{K}) \subseteq \mathrm{Mat}(n \times n, \mathbb{C})$. We may thus consider $M$ as a complex matrix, no matter whether its entries are real or complex. Therefore

$$\chi_f = \chi_M \in \mathbb{K}[t] \subseteq \mathbb{C}[t],$$

and considered as a complex polynomial it must have a zero $\lambda \in \mathbb{C}$ in the complex numbers. This, however, is an eigenvalue of the matrix $M$ considered as complex matrix, and by the above claim it is therefore a real number, with the property $\chi_f(\lambda) = \chi_M(\lambda) = 0$. That is, it is an eigenvalue of $f$!

***Claim*:** $f$ has an ONB of eigenvectors.

We do the proof by induction on $n = \dim_{\mathbb{K}}(V)$, where for the case $n = 1$ there is nothing to show.
We may therefore assume that $n > 1$ and that unitary endomorphisms on Hilbert spaces of dimension $n - 1$ are diagonalisable w. r. t. an ONB.
We have just shown that $f$ has an eigenvalue $\lambda_n$ with corresponding eigenvector $0 \neq v_n \in V$ of length $\|v_n\| = 1$ and $\lambda_n \in \mathbb{R}$. We set $U = \langle v_n \rangle$, and we show

$$f(U^\perp) \subseteq U^\perp.$$

For this let $u \in U^\perp$. Then

$$\langle f(u), v_n \rangle = \langle u, f(v_n) \rangle = \langle u, \lambda_n v_n \rangle = \overline{\lambda_n} \cdot \langle u, v_n \rangle = 0.$$

Thus $f(u) \perp v_n$, and hence $f(u) \in U^\perp$.

We may therefore consider the endomorphism $f$ restricted to $U^\perp$

$$f_| : U^\perp \to U^\perp : u \mapsto f(u),$$

which by default is self-adjoint again.
Thus by induction, since $\dim_K \left(U^\perp\right) = n-1$, there is an ONB $(v_1, \ldots, v_{n-1})$ of $U^\perp$ such that

$$f(v_i) = \lambda_i v_i$$

for $i = 1, \ldots, n-1$ and some $\lambda_i \in \mathbb{R}$.
However, $\langle v_i, v_j \rangle = \delta_{ij}$ for all $i, j = 1, \ldots, n$, since $v_i \in \langle v_n \rangle^\perp$. Thus by Lemma 3.7 $B = (v_1, \ldots, v_n)$ is an ONB of $V$ and $f(v_i) = \lambda_i v_i$ with $\lambda_i \in \mathbb{R}$ for all $i = 1, \ldots, n$.

b. We may apply Part a. to $V = \mathbb{K}^n$ with the standard scalar product and the endomorphism $f = f_A$. Then $T = T_B^E$ will do, where $B$ is the ONB which Part a. gives us, and $E$ is the standard basis. Note that by Proposition 3.14 $f$ is self-adjoint and by Example 3.12 $T \in O(n)$ resp. $T \in \mathcal{U}(n)$.

□

**3.18 Example**
Consider the matrix

$$A = \begin{pmatrix} 0 & -1 & i \\ -1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, \mathbb{C}).$$

Since $A = \overline{A}^t$, the matrix $A$ is self-adjoint. It is our aim to diagonalise $A$ w. r. t. an ONB, so first of all we have to find the eigenvalues of $A$.

$$\chi_A = \det(A - t \cdot \mathbb{1}) = -t^3 + 3t - 2 = (1+t)^2 \cdot (2-t),$$

which implies that the eigenvalues are $-1$ and $2$.
We next have to find ONB's of the eigenspaces of $A$, using the Gauß algorithm and the algorithm of Gram-Schmidt.
In order to find $\mathrm{Eig}(A, -1)$ we solve the linear system of equations

$$\begin{pmatrix} 1 & -1 & i \\ -1 & 1 & -i \\ -i & i & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (A + \mathbb{1}) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The algorithm of Gauß gives $\mathrm{Eig}(A, -1) = \left\langle (1,1,0)^t, (0,1,-i)^t \right\rangle$. We use the algorithm of Gram-Schmidt to transform these vectors into an ONB of the eigenspace, and we get

$$v_1 = \tfrac{1}{\sqrt{2}} \cdot (1,1,0)^t, \quad v_2 = \tfrac{1}{\sqrt{6}} \cdot (-1,1,-2i)^t.$$

We then calculate the eigenspace $\mathrm{Eig}(A, 2)$ with the aid of

$$\begin{pmatrix} -2 & -1 & i \\ -1 & -2 & -i \\ -i & i & -2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (A + \mathbb{1}) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

and get $\mathrm{Eig}(A,2) = \langle (1,-1,-i)^t \rangle$. Gram-Schmidt tells us to cut this vector down to length 1 in oder to have an ONB of $\mathrm{Eig}(A,2)$

$$v_3 = \tfrac{1}{\sqrt{3}} \cdot (1,-1,-i)^t.$$

Thus the matrix having these vectors $v_1, v_2, v_3$ as columns is the wanted transformation matrix,

$$T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ 0 & -\frac{2i}{\sqrt{6}} & -\frac{i}{\sqrt{3}} \end{pmatrix} \in \mathcal{U}(3),$$

and

$$T^{-1} \cdot A \cdot T = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

The above results on normal forms of self-adjoint matrices allow a classification of real symmetric bilinear forms with respect to base changes which respects distances and angles, i. e. with respect to ONB's. This is desirable when we consider geometric interpretations of symmetric bilinear forms respectively quadratic forms.

**3.19 Corollary (Normal Forms of Quadratic Forms by ONB's)**

a. *Let* $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ *be a symmetric matrix, then there is a* $T \in \mathcal{O}(n)$ *such that*

$$T^t \cdot A \cdot T = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}.$$

*where* $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ *are the eigenvalues of* $A$.

b. *Let* $b \in \mathrm{Bil}_{\mathbb{R}}(V)$ *be symmetric, then there is an ONB of* $V$ *such that*

$$M_B(b) = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}$$

*and*

$$M_B(q_b) = \sum_{i=1}^{n} \lambda_i x_i^2.$$

**Proof:** Part a. follows from Theorem 3.17 and the fact that for orthogonal matrices we have $T^{-1} = T^t$!

Part b. then is an immediate consequence of Part a. and the correspondence between symmetric matrices and symmetric bilinear forms studied in Section 2. □

# 4 Normal Forms of Cone Sections

**4.0 General Assumptions** We consider $\mathbb{R}^2$ endowed with the standard scalar product $\langle \cdot, \cdot \rangle$. By

$$\mathrm{I}(\mathbb{R}^2) = \{\varphi : \mathbb{R}^2 \to \mathbb{R}^2 \mid \varphi \text{ is an isometry}\}$$

we denote the group of isometries of the real plane. From the course on "From Groups to Geometry" it is known that

$$\mathrm{I}(\mathbb{R}^2) = \{\tau_v \circ f \mid v \in \mathbb{R}^2, f \in \mathcal{O}(2)\},$$

where $\tau_v : \mathbb{R}^2 \to \mathbb{R}^2 : \underline{x} \mapsto v + \underline{x}$ is the translation by $v$. I. e. every isometry can be decomposed as an orthogonal endomorphism followed by a translation.

**4.1 Definition**
Let $Q = \{p \in \mathbb{R}[x_1, x_2] \mid \deg(p) = 2\}$. For $p, q \in Q$ we define

$$p \sim q \iff \exists\, \varphi \in \mathrm{I}(\mathbb{R}^2), 0 \neq \lambda \in \mathbb{R} \;:\; q = \lambda \cdot (p \circ \varphi).$$

One easily checks that this defines an equivalence relation on $Q$, since $(\mathrm{I}(\mathbb{R}^2), \circ)$ and $(\mathbb{R} \setminus \{0\}, \cdot)$ are groups. We call the elements of $Q$ *conics*.

As always, we are interested in finding *simple* representatives for the equivalence classes of this relation, and we call them *normal forms*.

**4.2 Remark**
A polynomial $p \in Q$ has the form

$$\begin{aligned} p &= \alpha_{11}x_1^2 + 2\alpha_{12}x_1x_2 + \alpha_{22}x_2^2 + \alpha_1x_1 + \alpha_2x_2 + \alpha \\ &= (x_1, x_2) \cdot \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (\alpha_1, \alpha_2) \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \alpha \\ &= \langle \underline{x}, S \cdot \underline{x} \rangle + \langle \underline{a}, \underline{x} \rangle + \alpha, \end{aligned}$$

where $\underline{x} = (x_1, x_2)^t$, $\alpha_{21} := \alpha_{12}$, $\underline{a} = (\alpha_1, \alpha_2)^t$ and $S = \left(\begin{smallmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{smallmatrix}\right)$ is symmetric.
We are actually interested in the zero set of $p$, i. e. in

$$Z(p) := \{(x, y)^t \in \mathbb{R}^2 \mid p(x, y) = 0\}.$$

E. g. $p = x_2 - x_1^2$, then $Z(p)$ is the standard parabola in $\mathbb{R}^2$.
Note that multiplying $p$ with a non-zero constant $\lambda$ does *not* change $Z(p)$, and changing the coordinates by an isometry preserves distances and angles, that is, $Z(p)$ will be changed by a rotation or reflection followed by a translation.
E. g. let $p = x_2 - x_1^2$, $q = -x_2^2 + x_1 + 2x_2 + 1$, $f = f_A$ with $A = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $v = (1, 2)^t$ and $\varphi = \tau_v \circ f$. We claim that $q = p \circ \varphi$, in particular $p \sim q$. For this just note

$$\begin{aligned} p \circ \varphi &= p\big(\tau_v(f(x_1, x_2))\big) = \\ &\quad p\big(\tau_v(-x_2, x_1)\big) = p(-x_2 + 1, x_1 + 2) = (x_1 + 2) - (-x_2 + 1)^2 = q. \end{aligned}$$

Note that $Z(q)$ can be derived from $Z(p)$ by applying $\varphi^{-1}$ to it!

In order to be able to find the normal forms of

**4.3 Lemma**

*Let* $S \in \mathrm{Mat}(2 \times 2, \mathbb{R})$ *be symmetric.*

a. $\mathrm{Ker}\left(S^2\right) = \mathrm{Ker}(S)$ *and* $\mathrm{Im}\left(S^2\right) = \mathrm{Im}(S)$.

b. *For all* $\underline{a} \in \mathbb{R}^2$ *there is some* $\underline{c} \in \mathbb{R}^2$ *such that* $S^2 \cdot \underline{c} + S \cdot \underline{a} = \underline{0}$.

**Proof:** a. It is clear, that $\mathrm{Ker}(S) \subseteq \mathrm{Ker}\left(S^2\right)$. Let now $\underline{x} \in \mathrm{Ker}\left(S^2\right)$. Then

$$0 = \langle \underline{x}, S^2\underline{x} \rangle = \langle S\underline{x}, S\underline{x} \rangle.$$

This, however, implies $S\underline{x} = \underline{0}$, and thus $\underline{x} \in \mathrm{Ker}(S)$.
Again it is clear that $\mathrm{Im}\left(S^2\right) \subseteq \mathrm{Im}(S)$. But then the dimension formula gives

$$\dim_{\mathbb{R}} \mathrm{Im}\left(S^2\right) = 2 - \dim_{\mathbb{R}} \mathrm{Ker}\left(S^2\right) = 2 - \dim_{\mathbb{R}} \mathrm{Ker}(S) = \dim_{\mathbb{R}} \mathrm{Im}(S).$$

Hence $\mathrm{Im}\left(S^2\right) = \mathrm{Im}(S)$.

b. We have $S \cdot (-\underline{a}) \in \mathrm{Im}(S) = \mathrm{Im}\left(S^2\right)$. Thus there is a $\underline{c} \in \mathbb{R}^2$ such that $S^2 \cdot \underline{c} = S \cdot (-\underline{a})$, which proves the claim.

□

**4.4 Theorem (Classification of Cone Sections)**

*Let* $p = \langle \underline{x}, S\underline{x} \rangle + \langle \underline{a}, \underline{x} \rangle + \alpha \in Q$ *be arbitrary with* $S = \left(\begin{smallmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{smallmatrix}\right) \in \mathrm{Mat}(2 \times 2, \mathbb{R})$ *symmetric and* $\underline{a} = (\alpha_1, \alpha_2)^t \in \mathbb{R}^2$.
*Then* $p$ *is equivalent to one of the following* normal forms*:*

I: $\det(S) > 0$.

I.1: $\alpha \neq 0$ *and* $\alpha_{11} > 0$. $p \sim (\lambda_1 x_1)^2 + (\lambda_2 x_2)^2 - 1$ *and* $Z(p)$ *is an* ellipse.

I.2: $\alpha \neq 0$ *and* $\alpha_{11} < 0$. $p \sim (\lambda_1 x_1)^2 + (\lambda_2 x_2)^2 + 1$ *and* $Z(p)$ *is the empty set.*

I.3: $\alpha = 0$. $p \sim (\lambda_1 x_1)^2 + (\lambda_2 x_2)^2$ *and* $Z(p)$ *is a single point.*

II: $\det(S) < 0$.

II.1: $\alpha \neq 0$. $p \sim (\lambda_1 x_1)^2 - (\lambda_2 x_2)^2 - 1$ *and* $Z(p)$ *is a* hyperbola.

II.2: $\alpha = 0$. $p \sim (\lambda_1 x_1)^2 - (\lambda_2 x_2)^2$ *and* $Z(p)$ *consists of two different lines through the origin.*

III: $\det(S) = 0$, $a \neq (0,0)^t$. $p \sim x_1^2 - \lambda x_2$ *and* $Z(p)$ *is a* parabola.

IV: $\det(S) = 0$, $a = (0,0)^t$.

IV.1: $\alpha \neq 0$ *and* $S$ *has a positive eigenvalue.* $p \sim x_1^2 - \lambda$, $\lambda > 0$, *and* $Z(p)$ *consists of two parallel lines.*

IV.2: $\alpha \neq 0$ *and* $S$ *has a negative eigenvalue.* $p \sim x_1^2 + \lambda$, $\lambda > 0$, *and* $Z(p)$ *is the empty set.*

IV.3: $\alpha = 0$. $p \sim x_1^2$ *and* $Z(p)$ *consists of a line counted twice.*

**Proof:** 1st Case: $\underline{a} = (0,0)^t$: Let's first consider the case $\underline{a} = (0,0)^t$.
By Corollary 3.19 there is a $T \in \mathcal{O}(2)$, such that

$$T^t \cdot S \cdot T = T^{-1} \cdot S \cdot T = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

Note that not both eigenvalues $\mu_1$ and $\mu_2$ can be zero, since $S \neq 0$. Hence we may assume $\mu_1 \neq 0$ and $\mu_1 \geq \mu_2$ if $\mu_2 \neq 0$.
The endomorphism $f_T : \mathbb{R}^2 \to \mathbb{R}^2 : \underline{x} \mapsto T\underline{x}$ is a rotation or a reflection and we have

$$\begin{aligned} p(T\underline{x}) &= \langle T\underline{x}, (S \cdot T)\underline{x}\rangle + \alpha \\ &= \langle \underline{x}, \left(T^t \cdot S \cdot T\right)\underline{x}\rangle + \alpha \\ &= \mu_1 x_1^2 + \mu_2 x_2^2 + \alpha. \end{aligned}$$

Multiplying with a suitable constant, we may assume that either $\alpha = 0$ or $\alpha = -1$. Define $\lambda_i = \sqrt{|\mu_i|}$, then we have to distinguish the following cases.
Case 1.1: $\mu_1, \mu_2 > 0$: This is equivalent to the fact that $S$ is positive definite and hence that $\det(S) > 0$ and $\alpha_{11} > 0$. If $\alpha = -1$, we are in Case I.1, and if $\alpha = 0$, we are in Case I.3.
Case 1.2: $\mu_1, \mu_2 < 0$: This is equivalent to the fact that $-S$ positive definite, hence that $\det(S) = \det(-S) > 0$ and $-\alpha_{11} > 0$. Is $\alpha = -1$, we are in Case I.2, and for $\alpha = 0$ its again Case I.3, since we may multiply the polynomial once more by $-1$.
Case 1.3: $\mu_1 > 0, \mu_2 < 0$: This is equivalent to $\mu_1 \cdot \mu_2 = \det(S) < 0$. $\alpha = -1$ leads to Case II.1 and $\alpha = 0$ to Case II.2.
Case 1.4: $\mu_1 > 0, \mu_2 = 0$ or $\mu_1 < 0, \mu_2 = 0$: This is Equivalent to $\det(S) = 0$. If $\mu_1 > 0$ and $\alpha = -1$ we are in Case IV.1, for $\mu_1 < 0$ and $\alpha = -1$ we get Case IV.2, and for $\alpha = 0$ it is Case IV.3.
2nd Case: $\underline{a} \neq (0,0)^t$: In Case $\underline{a} = (0,0)^t$ we got around without applying any translations. This will now be different.
For $\underline{c} \in \mathbb{R}^2$ the translation $t_{\underline{c}} : \mathbb{R}^2 \to \mathbb{R}^2 : \underline{x} \mapsto \underline{x} + \underline{c}$ leads to the following coordinate transformation for $p$

$$\begin{aligned} p(\underline{x} + \underline{c}) &= \langle \underline{x} + \underline{c}, S\underline{x} + S\underline{c}\rangle + 2\langle \underline{a}, \underline{x} + \underline{c}\rangle + \alpha \\ &= \langle \underline{x}, S\underline{x}\rangle + 2\langle \underline{a} + S\underline{c}, \underline{x}\rangle + \langle \underline{c}, S\underline{c}\rangle + 2\langle \underline{a}, \underline{c}\rangle + \alpha \qquad (12) \\ &= \langle \underline{x}, S\underline{x}\rangle + 2\langle \underline{b}, \underline{x}\rangle + \beta, \end{aligned}$$

where $\underline{b} = \underline{a} + S\underline{c}$ and $\beta = \langle \underline{c}, S\underline{c}\rangle + 2\langle \underline{a}, \underline{c}\rangle + \alpha$.
Case 2.1: $\exists\, \underline{c} \in \mathbb{R}^2 : \underline{b} = \underline{a} + S\underline{c} = (0,0)^t$: The transformation $p \mapsto p\left(t_{\underline{c}}(\underline{x})\right)$ reduces to the first Case "$\underline{a} = (0,0)^t$". Hence $p$ is equivalent one of the Cases I, II or IV.
Case 2.2: $\forall\, \underline{c} \in \mathbb{R}^2 : \underline{b} = \underline{a} + S\underline{c} \neq (0,0)^t$: By Lemma 4.3 there is a $\underline{c} \in \mathbb{R}^2$ such that $S\underline{b} = S^2\underline{c} + S\underline{a} = 0$. If we define $\delta := -\frac{\beta}{2\langle \underline{b}, \underline{b}\rangle}$, then the translation $t_{\underline{c}+\delta\underline{b}}$ leads to

$$\begin{aligned} p(\underline{x} + \underline{c} + \delta\underline{b}) &= \langle \underline{x}, S\underline{x}\rangle + 2\langle \underline{a} + S(\underline{c} + \delta\underline{b}), \underline{x}\rangle + \langle \underline{c} + \delta\underline{b}, S(\underline{c} + \delta\underline{b})\rangle + 2\langle \underline{a}, \underline{c} + \delta\underline{b}\rangle + \alpha \\ &= \langle \underline{x}, S\underline{x}\rangle + 2\langle \underline{b} + \delta S\underline{b}, \underline{x}\rangle + \delta^2\langle \underline{b}, S\underline{b}\rangle + 2\delta\langle \underline{b}, \underline{b}\rangle + \beta \\ &= \langle \underline{x}, S\underline{x}\rangle + 2\langle \underline{b}, \underline{x}\rangle + 2\delta\langle \underline{b}, \underline{b}\rangle + \beta \\ &= \langle \underline{x}, S\underline{x}\rangle + 2\langle \underline{b}, \underline{x}\rangle. \end{aligned}$$

Taking into account that $0$ is certainly an eigenvalue of $S$, since $S\underline{b} = 0$, and since $S \neq 0$, Corollary 3.19 implies the existence of a $T \in \mathcal{O}(2)$, such that

$$D := T^t \cdot S \cdot T = T^{-1} \cdot S \cdot T = \begin{pmatrix} \mu_1 & 0 \\ 0 & 0 \end{pmatrix},$$

where $\mu_1 \neq 0$. In particular we are in the case $\det(S) = 0$.
Moreover, for $T^t\underline{b} =: (\mu, \lambda)^t$ we have, taking $T^t = T^{-1}$ into account,

$$(\mu_1\mu, 0) = \big(T^t \cdot S \cdot T\big) \cdot (T^t\underline{b}) = T^t \cdot (S\underline{b}) = 0,$$

and hence $T^t\underline{b} = (0, \lambda)^t$, where $\lambda \neq 0$, since $T^t$ is invertible and $\underline{b} \neq (0,0)^t$. But then the map $\underline{x} \mapsto T\underline{x}$ transforms the polynomial $\langle \underline{x}, S\underline{x}\rangle + 2\langle \underline{b}, \underline{x}\rangle$ into

$$\langle T\underline{x}, (S \cdot T)\underline{x}\rangle + 2\langle \underline{b}, T\underline{x}\rangle = \langle \underline{x}^t, D\underline{x}\rangle + 2\langle T^t\underline{b}, \underline{x}\rangle = \mu_1 x_1^2 + 2\lambda x_2.$$

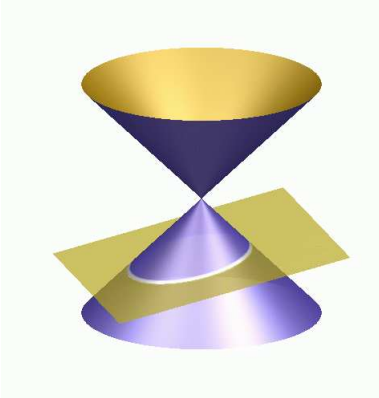
I. e.

$$q := p\big((t_{\underline{c}+\delta\underline{b}} \circ f_T)(\underline{x})\big) = \mu_1 x_1^2 + 2\lambda x_2,$$

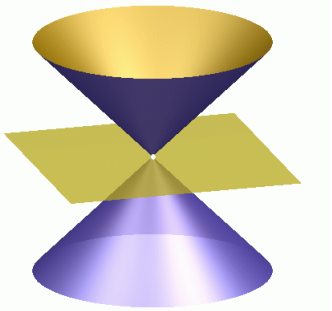
and we are in Case III. □

**4.5 Remark**

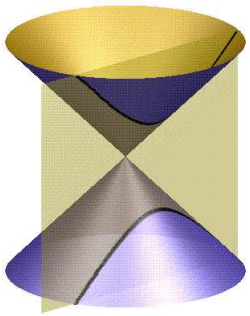
The sets $Z(p)$ with $p \in Q$ are called *cone sections* since all of them, except for the cases I.2, IV.1 and IV.2, can be realised as intersections of the double cone $C = \{x_3^2 - x_2^2 - x_1^2 = 0\}$ in $\mathbb{R}^3$ with a suitable plane.



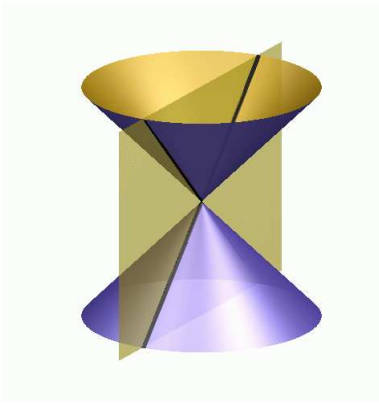
I.1: Ellipse



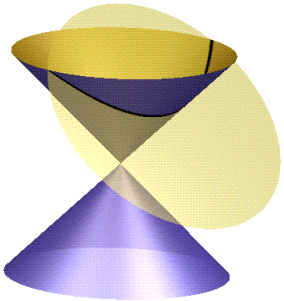
I.3: Point



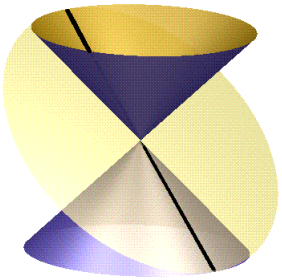
II.1: Hyperbola



II.2: Two Lines with Intersection



III: Parabola



IV.3: Double Line

APPENDIX A

# Assignments and Solutions

## Assignment Set 1

Exercise 2 should be handed in for marking. Exercises with an asterisque (*) are considered challenging, and you should not spend too much time on trying to solve them.

**Exercise 1:** Let $n \geq 2$ be an integer. Denote by $\sigma = (1\ 2\ \ldots\ n) \in \mathbb{S}_n$ the $n$-cycle with $\sigma(i) = i + 1$ for $i < n$ and $\sigma(n) = 1$, and by $\tau \in \mathbb{S}_n$ the permutation with $\tau(i) = n + 1 - i$ for $i = 1, \ldots, n$.

a. Show that $\sigma^n = (1)$, $\tau^2 = (1)$, and $\sigma^\tau = \sigma^{-1}$.

b. Show that $\langle \sigma, \tau \rangle = \{\sigma^i, \tau \circ \sigma^i \mid i = 0, \ldots, n-1\}$ is a group of order $2n$. We denote this group by $\mathbb{D}_{2n}$ and call it the dihedral group of order $2n$.

c. Define a permutation $\pi : \mathbb{D}_{2n} \to \mathbb{D}_{2n}$ on $\mathbb{D}_{2n}$ by $\pi(\sigma^i) = \sigma^{n-1-i}$ and $\pi(\tau \circ \sigma^i) = \tau \circ \sigma^i$. Setting $\pi_i := \pi^i \in \mathrm{Sym}(\mathbb{D}_{2n})$ for $i = 1, \ldots, r$, $r \geq 3$ any integer, show that the check digit code $C_{\mathbb{D}_{2n}}(\pi_1, \ldots, \pi_r, (1))$ detects errors of type II.

Note: If you have problems dealing with the general case, you may just replace $n$ by any of the numbers $n = 3$, 4 or 5.

**Exercise 2:** Let $(G, \cdot)$ be a group such that $g^2 = e_G$ for all $g \in G$. Show that $G$ is abelian.

**Exercise 3:** Let $p$ be a prime number and set $\mathbb{Z}\left[\frac{1}{p}\right] = \left\{\frac{z}{p^n} \in \mathbb{Q} \mid n, z \in \mathbb{Z}, n \geq 0\right\}$.

a. Show that $\mathbb{Z}\left[\frac{1}{p}\right]$ is a subgroup of $(\mathbb{Q}, +)$, i. e. of the rational numbers with respect to addition.

b.* Find all the subgroups of $\left(\mathbb{Z}\left[\frac{1}{p}\right], +\right)$ and of $(\mathbb{Z}_{p^\infty}, +)$, where $\mathbb{Z}_{p^\infty} = \mathbb{Z}\left[\frac{1}{p}\right]/\mathbb{Z}$ is the factor group of $\mathbb{Z}\left[\frac{1}{p}\right]$ by the normal subgroup $\mathbb{Z}$.

Note: It will turn out that every strict subgroup of $\mathbb{Z}_{p^\infty}$ is finite and cyclic, while the $\mathbb{Z}_{p^\infty}$ itself is not even finitely generated!

**Exercise* 4:** Let $(\mathbb{C}, +, \cdot)$ denote the field of complex numbers, and denote by $i$ the imaginary unit with $i^2 = -1$. Consider the subgroup

$$\mathbb{Q}_8 := \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \right\rangle < \mathrm{Gl}_2(\mathbb{C})$$

of the group of invertible $2 \times 2$-matrices over $\mathbb{C}$ with respect to matrix multiplication. Find all the subgroups of $\mathbb{Q}_8$.

## Assignment Set 2

**Exercise 1:** Let $(G, \cdot)$ is a group, $g \in G$ and $n = \min\{m \in \mathbb{Z} \mid m > 0, g^m = e_G\} < \infty$. Show that $\langle g \rangle = \{e_G = g^0, g^1, g^2, \ldots, g^{n-1}\}$.

**Exercise 2:** Which of the sets $A = \{z \in \mathbb{Z} \mid z > 0\}$, $B = \{\frac{1}{z} \mid z \in A\}$ and $A \cup B$ is a group with respect to the multplication of integers?

**Exercise 3:** Show that $\mathbb{E} = \{z \in \mathbb{C} \mid |z| = 1\}$ is a subgroup of $(\mathbb{C} \setminus \{0\}, \cdot)$, where $\mathbb{C}$ denotes the complex numbers.

**Exercise 4:** Let $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$. Define

$$\circ : \mathbb{R}_{>0} \times \mathbb{R}_{>0} \to \mathbb{R}_{>0} : (x, y) \mapsto x \circ y := x^y.$$

Is $(\mathbb{R}_{>0}, \circ)$ a group?

## Assignment Set 3

Exercises 1 and 3 should be handed in for marking.

**Exercise 1:** Let $(G, \cdot)$ be a finite group and let $U, V \leq G$. Use the Theorem of Lagrange to prove the following statements.

a. If $V \subseteq U$, then $|G : V| = |G : U| \cdot |U : V|$.

b. If $\gcd\big(|G : U|, |G : V|\big) = 1$, then $G = U \cdot V$.

**Exercise 2:** Let $(G, \cdot)$ be a group, $U_i \leq G$ for $i \in I$. Show that $\bigcap_{i \in I} U_i \leq G$.

**Exercise 3:** Let $(G, \cdot)$ be a group, $N, N_1, N_2 \trianglelefteq G$, $U \leq G$.

a. $N \cap U \trianglelefteq U$.

b. $N_1 \cap N_2 \trianglelefteq G$.

**Exercise 4:**

a.* Let $\pi = (a_1\ a_2\ \ldots\ a_k) \in \mathbb{S}_n$ be a $k$-cycle and $\sigma = (b_1\ b_2\ \ldots\ b_l) \in \mathbb{S}_n$ be an $l$-cycle. Show that that $\pi$ and $\sigma$ are conjugate (i. e. $\exists\ \zeta \in \mathbb{S}_n$ s. t. $\zeta \circ \pi \circ \zeta^{-1} = \sigma$) if and only if $k = l$.

Hint, if $k = l$ then it is easy just to give $\zeta$, for the opposite direction I recommend to have a look at $\pi^k$ and $\sigma^k$.

b. We know that any permutation in $\mathbb{S}_n$ has a unique representation as a product of disjoint cycles. Suppose that $\pi = \zeta_1 \circ \ldots \circ \zeta_r$ is such a representation for $\pi \in \mathbb{S}_n$ and suppose that $\zeta_i$ is a $k_i$-cycle with $k_1 \geq k_2 \geq \ldots \geq k_r$. We then call $(k_1, \ldots, k_r)$ the *cycle type* of $\pi$.
Use part a. in order to show that two permutations are conjugate if and only if they have the same cycle type.

c. Use these results to show that

$$\mathbb{K}_4 = \{(1), (1\,2)(3\,4), (1\,3)(2\,4), (1\,4)(2\,3)\}$$

is a normal subgroup of $\mathbb{S}_4$, the so called Kleinian group of order 4.

Hint, in oder to see that it is a subgroup of $\mathbb{S}_4$ it is best to write down the group table, which shows the closedness with respect to the group operation and with respect to taking inverses.

**Exercise 5:** Find all the subgroups of $\mathbb{D}_8 = \langle (1\ 2\ 3\ 4), (1\ 4)(2\ 3)\rangle$, which was introduced on the assingment set 2. Which of the subgroups are normal subgroups?

**Solution to Exercise 1** a. By the Theorem of Lagrange we know

$$|G : V| = \frac{|G|}{|V|}, |G : U| = \frac{|G|}{|U|}, \text{ and } |U : V| = \frac{|U|}{|V|}.$$

This proves the claim.

b. We set $H = U \cap V$. Then by Part a. we have

$$|G : V| \,\big|\, |G : H| \text{ and } |G : U| \,\big|\, |G : H|.$$

Thus also the least common multiple divides $|G : H|$, i. e.

$$|G : V| \cdot |G : U| = \frac{|G : V| \cdot |G : U|}{\gcd\left(|G : V|, |G : U|\right)} = \operatorname{lcm}\left(|G : V|, |G : U|\right) \,\big|\, |G : H|,$$

and there is a number $m > 0$ such that

$$m \cdot \frac{|G|}{|V|} \cdot \frac{|G|}{|U|} = m \cdot |G : V| \cdot |G : U| = |G : H| = \frac{|G|}{|H|}.$$

From this equation we deduce with the product formula

$$|U \cdot V| = \frac{|U| \cdot |V|}{|H|} = |G| \cdot m \geq |G|.$$

Being a subset of $G$, this implies $U \cdot V = G$.

**Solution to Exercise 2**
Since $e_G \in U_i$ for all $i \in I$, $e_G \in \bigcap_{i\in I} U_i$, so that this set is non-empty. Let $u, v \in \bigcap_{i\in I} U_i$. We have to show that $u \cdot v, u^{-1} \in \bigcap_{i\in I} U_i$. By assumption $u, v \in U_i$ for all $i \in I$, and thus $u \cdot v, u^{-1} \in U_i$ for all $i \in I$, since the $U_i$ are subgroups. But then $u \cdot v, u^{-1} \in \bigcap_{i\in I} U_i$, and $\bigcap_{i\in I} U_i \leq G$.

**Solution to Exercise 3** a. Being the intersection of subgroups, $N \cap U$ is a subgroup of $G$, which is contained in $U$. Hence, it's a subgroup of $U$. It remains to check the normality condition. Let $u \in U$ and $n \in N \cap U$. Then

$$u \cdot n \cdot u^{-1} \in N \cap U,$$

since $N$ is a normal subgroup of $G$ and $U$ is closed under multiplication. Thus, $N \cap U \trianglelefteq U$.

b. Being the intersection of subgroups $N_1 \cap N_2$ is a subgroup of $G$. It remains to check a normality condition. Let $g \in G$ and $n \in N_1 \cap N_2$. Since both, $N_1$ and $N_2$, are normal subgroups of $G$, we get

$$g \cdot n \cdot g^{-1} \in N_i$$

for $i = 1, 2$, and thus $g \cdot n \cdot g^{-1} \in N_1 \cap N_2$.

**Solution to Exercise 4** a. Let $\{1, \ldots, n\} = \{a_1, \ldots, a_n\} = \{b_1, \ldots, b_n\}$.

"$\Leftarrow$" Suppose $k = l$. We define a permutation $\zeta \in \mathbb{S}_n$ by

$$\zeta = \begin{pmatrix} a_1 & \ldots & a_n \\ b_1 & \ldots & b_n \end{pmatrix}.$$

The inverse of $\zeta$ is then just

$$\zeta^{-1} = \begin{pmatrix} b_1 & \ldots & b_n \\ a_1 & \ldots & a_n \end{pmatrix}.$$

We claim that $\zeta \circ \pi \circ \zeta^{-1} = \sigma$. For this we apply both maps to $b_i$, $i = 1, \ldots, n$:

$$\big(\zeta \circ \pi \circ \zeta^{-1}\big)(b_i) = \zeta\big(\pi(a_i)\big) = \begin{cases} \zeta(a_{i+1}) = b_{i+1} = \sigma(b_i), & \text{if } 1 \le i < k = l, \\ \zeta(a_1) = b_1 = \sigma(b_l), & \text{if } i = k = l, \\ \zeta(b_i) = a_i = \sigma(b_i), & \text{if } k+1 = l+1 \le i \le n. \end{cases}$$

"$\Rightarrow$" We may assume that $k \le l$. Let $\zeta \in \mathbb{S}_n$ be such that $\zeta \circ \pi \circ \zeta^{-1} = \sigma$. Then

$$\sigma^k = \big(\zeta \circ \pi \circ \zeta^{-1}\big)^k = \zeta \circ \pi^k \circ \zeta^{-1} = \zeta \circ (1) \circ \zeta^{-1} = (1).$$

Hence, $l = o(\sigma) \le k$, and thus $k = l$.

b. Let $\pi = \zeta_1 \circ \ldots \circ \zeta_r$ and $\sigma = \tau_1 \circ \ldots \circ \tau_s$, where $\zeta_i$ is a $k_i$-cycle with $k_1 \ge k_2 \ge \ldots \ge k_r$ and $\tau_i$ is an $l_i$-cycle with $l_1 \ge l_2 \ge \ldots \ge l_s$.

"$\Rightarrow$" Suppose there is a $\xi \in \mathbb{S}_n$ such that $\xi \circ \pi \circ \xi^{-1} = \sigma$. By Part a. we then have that

$$\omega_i := \xi \circ \zeta_i \circ \xi^{-1}$$

also is a $k_i$-cycle. We deduce thus that

$$\begin{aligned} \omega_1 \circ \omega_2 \circ \ldots \circ \omega_r &= \big(\xi \circ \zeta_1 \circ \xi^{-1}\big) \circ \big(\xi \circ \zeta_2 \circ \xi^{-1}\big) \circ \ldots \circ \big(\xi \circ \zeta_r \circ \xi^{-1}\big) \\ &= \xi \circ \big(\zeta_1 \circ \zeta_2 \circ \ldots \circ \zeta_r\big) \circ \xi^{-1} = \xi \circ \pi \circ \xi^{-1} = \sigma. \end{aligned}$$

Thus the cycle type of $\sigma$ must be $(k_1, \ldots, k_r)$. However, it is also $(l_1, \ldots, l_s)$, which implies that

$$(k_1, \ldots, k_r) = (l_1, \ldots, l_s).$$

"$\Leftarrow$" Let's now suppose that $r = s$ and $(k_1, \ldots, k_r) = (l_1, \ldots, l_s)$. Moreover, suppose that $\zeta_i = (a_{i,1} \ldots a_{i,k_i})$ and $\tau_i = (b_{i,1} \ldots b_{i,k_i})$ for $i = 1, \ldots, r$. Then

$$\{1, \ldots, n\} = \{a_{i,j} \mid i = 1, \ldots, r, j = 1, \ldots, k_i\} = \{b_{i,j} \mid i = 1, \ldots, r, j = 1, \ldots, k_i\}.$$

As in the proof of Part a. we define a permutation

$$\xi = \begin{pmatrix} a_{1,1} & \dots & a_{r,k_r} \\ b_{1,1} & \dots & b_{r,k_r} \end{pmatrix},$$

and it follows for $b_{i,j}$

$$\big(\xi \circ \pi \circ \xi^{-1}\big)(b_{i,j}) = \xi\big(\pi(a_{i,j})\big) = \begin{cases} \xi(a_{i,j+1}) = b_{i,j+1} = \sigma(b_{i,j}), & \text{if } 1 \le j < k_i, \\ \xi(a_{i,1}) = b_{i,1} = \sigma(b_{i,j}), & \text{if } j = k_i. \end{cases}$$

This proves that $\xi \circ \pi \circ \xi^{-1} = \sigma$.

c. We set $e = (1)$, $a = (1\ 2)(3\ 4)$, $b = (1\ 3)(2\ 4)$ and $c = (1\ 4)(2\ 3)$. Then the multiplication table of $\mathbb{K}_4$ looks like

| | e | a | b | c |
|---|---|---|---|---|
| e | e | a | b | c |
| a | a | e | c | b |
| b | b | c | e | a |
| c | c | b | a | e |

This shows that $\mathbb{K}_4$ is closed under the multiplication and that every element has an inverse. Thus $\mathbb{K}_4 \le \mathbb{S}_4$ is a subgroup.
Let now $\zeta \in \mathbb{S}_n$ be given and $\pi \in \mathbb{K}_4$. We have to show that $\zeta \circ \pi \circ \zeta^{-1} \in \mathbb{K}_4$ in order to see that $\mathbb{K}_4$ is a normal subgroup. If $\pi = e$, then the product is again $e$ and belongs to $\mathbb{K}_4$. If $\pi \in \{a, b, c\}$, then by Part b. the product $\zeta \circ \pi \circ \zeta^{-1}$ has cycle type $(2, 2)$. However, all elements of this cycle type in $\mathbb{S}_4$ belong to $\mathbb{K}_4$. Thus the product does as well.

**Solution to Exercise 5**

Note, that with $\sigma = (1\ 2\ 3\ 4)$ and $\tau = (1\ 4)(2\ 3)$ we have by Exercise 1 on Assignment Set 1

$$\begin{aligned} \mathbb{D}_8 &= \big\{\mathrm{id}, \sigma, \sigma^2, \sigma^3, \tau, \tau\circ\sigma, \tau\circ\sigma^2, \tau\circ\sigma^3\big\} \\ &= \big\{(1), (1\ 2\ 3\ 4), (1\ 3)(2\ 4), (1\ 4\ 3\ 2), \\ &\qquad (1\ 4)(2\ 3), (1\ 3), (1\ 2)(3\ 4), (2\ 4)\big\}. \end{aligned}$$

If $U \le \mathbb{D}_8$, then by the Theorem of Lagrange $|U| \in \{1, 2, 4, 8\}$.
If $|U| = 1$, then of course $U = \mathbb{1}$.
If $|U| = 2$, then $U$ is cyclic and generated by one element of order 2. Thus $U$ is one of the groups $\langle\tau\rangle$, $\langle\tau\circ\sigma\rangle$, $\langle\tau\circ\sigma^2\rangle$, $\langle\tau\circ\sigma^3\rangle$, or $\langle\sigma^2\rangle$.
If $|U| = 4$, then $U$ may be cyclic or it only contains elements of order at most 2. If it is cyclic, then it must contain two elements of order 4, which are inverse to each other, thus $U = \langle\sigma\rangle$. Otherwise $U$ is one of the two groups $\langle\tau, \tau\circ\sigma^2\rangle$ or $\langle\tau\circ\sigma, \tau\circ\sigma^3\rangle$.
If $|U| = 8$, then of course $U = \mathbb{D}_8$.

## Assignment Set 4

Exercises 1 and 2 should be handed in for marking.

**Exercise 1:** Let $\alpha \in \mathrm{Hom}(G, H)$, where $(G, \cdot)$ and $(H, *)$ are groups.

a. $\mathrm{Im}(\alpha) := \alpha(G) \leq G$ and is called the *image* of $\alpha$.

b. If $\alpha$ is bijective, then $\alpha^{-1} \in \mathrm{Hom}(H, G)$.
In particular, $\big(\mathrm{Aut}(G), \circ\big)$ is a subgroup of $\big(\mathrm{Sym}(G), \circ\big)$.

**Exercise 2:** Let $(G, \cdot)$ be a group and let $N, N' \trianglelefteq G$ be two normal subgroups. Prove the Isomorphism Theorem $(N \cdot N')/N' \cong N/(N \cap N')$.

**Exercise 3:** Suppose that $A$ and $B$ are finite sets with the same number of elements. Show that the groups $\big(\mathrm{Sym}(A), \circ\big)$ and $\big(\mathrm{Sym}(B), \circ\big)$ are isomorphic.

**Exercise 4:** Let $(G, \cdot)$ be a group. We call $Z(G) = \{h \in G \mid gh = hg \ \forall\ g \in G\}$ the centre of $G$, i. e. the set of elements in $G$ which commute with all other elements.[1]

a. $Z(G) \trianglelefteq G$.

b. If $G/Z(G)$ is cyclic, then $G$ is abelian.[2]

**Exercise* 5: [Generators and Relations]**

Let $x$ and $y$ be two different symbols. Consider the set of *words*

$$W = \left\{x^{\alpha_1} y^{\beta_1} \cdots x^{\alpha_r} y^{\beta_r} \;\middle|\; \alpha_i, \beta_i \in \mathbb{Z}, r \geq 1\right\} \cup \{e\},$$

where $e$ is just a symbol defining the so called empty word. We use the common exponential laws in order to simplify such words and we consider words which become the same that way to be the same, e. g.

$$x^3y^5y^{-3}x^0y^{-2}x^{-3} = x^3y^5y^{-3}y^{-2}x^{-3} = x^3y^{5-3-2}x^{-3} = x^3x^{-3} = x^{3-3} = e.$$

There is then a natural way to multiply words just by putting them together, and having the empty word $e$ operate as the identity. This way $W$ becomes a group, and obviously $W = \langle x, y\rangle$ is generated by the elements $x, y$.[3]
Moreover, if $M = \{w_1 = w_1', \ldots, w_r = w_r'\}$ is a set of equations of words in $W$ - called *relations*-, then we consider the smallest normal subgroups which contains the set $M' = \left\{w_1^{-1} \cdot w_1', \ldots, w_r^{-1} \cdot w_r'\right\}$

$$N(M) = \bigcap_{M' \subseteq N \trianglelefteq W(X)} N,$$

and we denote by $\langle x, y \mid w_1 = w_1', \ldots, w_r = w_r'\rangle$ the quotient group of $W$ by the normal subgroup $N(M)$. By abuse of notation we will denote the *generators* of this quotient group still by $x$ and $y$ rather than by $xN(M)$ and $yN(M)$.

[1] Note, that obviously $G$ is abelian if and only if $G = Z(G)$.

[2] This then implies $G = Z(G)$!

[3] You are not required to prove these facts! Their proof is quite tedious and can be found - in a more general setting - in many textbooks, e. g. Michael Weinstein, Examples of Groups, pp. 52ff.

a. Show that $W$ has the following universal property: given any group $(G, \cdot)$ such that $G = \langle g, h\rangle$, then there is a *unique* epimorphism $\alpha : W \to G$ such that $\alpha(x) = g$ and $\alpha(y) = h$.
Moreover, if the relations $w_i = w_i'$ still hold when you replace $x$ by $g$ and $y$ by $h$, then $\alpha$ induces an epimorphism

$$\overline{\alpha} : \langle x, y \mid \ w_1 = w_1', \ldots, w_r = w_r'\rangle \longrightarrow G : w \mapsto \alpha(w).$$

with $\overline{\alpha}(x) = g$ and $\overline{\alpha}(y) = h$.[4]

b. Show that the group $\left\langle x, y \;\middle|\; x^n = e, y^2 = e, yxy^{-1} = x^{-1}\right\rangle$ is isomorphic to $\mathbb{D}_{2n}$.

c. Show that the group $\left\langle x, y \;\middle|\; x^4 = e, y^4 = e, yxy^{-1} = x^{-1}\right\rangle$ is isomorphic to $\mathbb{Q}_8$.

d. Show that a non-abelian group of order 8 is either isomorphic to $\mathbb{D}_8$ or to $\mathbb{Q}_8$.

**Solution to Exercise 1** a. Since $e_H = \alpha(e_G) \in \mathrm{Im}(\alpha)$, the set is non-empty. Moreover, for $g, g' \in G$ we have $\alpha(g) * \alpha\big(g'\big) = \alpha\big(g \cdot g'\big) \in \mathrm{Im}(\alpha)$ and $\alpha(g)^{-1} = \alpha\big(g^{-1}\big) \in \mathrm{Im}(\alpha)$. Hence, $\mathrm{Im}(\alpha) \leq H$.

b. Let $h, h' \in H$ be given. Since $\alpha$ is bijective, there are elements $g, g' \in G$ such that $\alpha(g) = h$ and $\alpha\big(g'\big) = h'$. We thus have for the inverse mapping $\alpha^{-1}$

$$\alpha^{-1}\big(h * h'\big) = \alpha^{-1}\big(\alpha(g) * \alpha(g')\big) = \alpha^{-1}\big(\alpha(g \cdot g')\big) = g \cdot g' = \alpha^{-1}(h) \cdot \alpha^{-1}\big(h'\big).$$

For the "in particular part" just note that we have shown in the lecture that $\mathrm{Aut}(G)$ is closed under composition of maps and that we have just proved that it is also closed under taking inverses. Moreover, since $\mathrm{id}_G \in \mathrm{Aut}(G)$, it is a non-empty subset of $\mathrm{Sym}(G)$, and hence a subgroup thereof.

**Solution to Exercise 2**
Note that $N \cdot N'$ is actually a group, and $N'$ is a normal subgroup thereof. We have also seen that $N \cap N' \trianglelefteq N$, so that the quotient groups in this statement actually exist! Let's now define a map by

$$\alpha : N \to N \cdot N'/N' : n \mapsto nN'.$$

We are going to show that this map is an epimorphism with kernel $N \cap N'$, and then we apply the Homomorphism Theorem.

***Step 1*:** $\alpha$ is a homomorphism.

Let $n, m \in N$, then $\alpha(n \cdot m) = nmN' = nN' \cdot mN' = \alpha(n) \cdot \alpha(m)$.

***Step 2*:** $\alpha$ is a surjective.

Let $nn'N' \in NN'/N'$ be arbitrary with $n \in N$ and $n' \in N'$. Then $\alpha(n) = nN' = nn'N'$, thus $\alpha$ is surjective.

***Step 3*:** $\mathrm{Ker}(\alpha) = N \cap N'$.

[4] That the relations $w_i = w_i'$ are satisfied when replacing $x$ by $g$ and $y$ by $h$ is the same as saying that $N(M)$ is contained in the kernel of $\alpha$.

We have $n \in \mathrm{Ker}(\alpha)$ if and only if $nN' = N'$ if and only if $n \in N' \cap N$.
Applying now the Homomorphism Theorem we get

$$N/N \cap N' = N/\mathrm{Ker}(\alpha) \cong \mathrm{Im}(\alpha) = NN'/N'.$$

**Solution to Exercise 3**
Since $A$ and $B$ have the same order, there exists a bijection

$$\phi : A \to B.$$

We use this to define a map

$$\alpha : \mathrm{Sym}(A) \to \mathrm{Sym}(B) : \pi \mapsto \phi \circ \pi \circ \phi^{-1}.$$

This map is obviously bijective with inverse

$$\beta : \mathrm{Sym}(B) \to \mathrm{Sym}(A) : \pi \mapsto \phi^{-1} \circ \pi \circ \phi.$$

Moreover, for $\pi, \sigma \in \mathrm{Sym}(A)$ we have

$$\alpha(\pi \circ \sigma) = \phi \circ \pi \circ \sigma \circ \phi^{-1} = \phi \circ \pi \circ \phi^{-1} \circ \phi \circ \sigma \circ \phi^{-1} = \alpha(\pi) \circ \alpha(\sigma).$$

Thus $\alpha$ is also a homomorphism, hence it is an isomorphism.

**Solution to Exercise 4** a. Let's show first that $Z(G)$ is actually a subgroup of $G$. Since $e_G$ commutes with any element in $G$, it belongs to $Z(G)$, so that the set is non-empty. Let $h, h' \in Z(G)$ and $g \in G$ arbitrary. Then

$$hh'g = hgh' = ghh' \quad \text{and} \quad h^{-1}g = \left(g^{-1}h\right)^{-1} = \left(hg^{-1}\right)^{-1} = gh^{-1},$$

hence $hh', h^{-1} \in Z(G)$ and $Z(G) \leq G$.
It remains to check the normality condition. Let for this $g \in G$ and $h \in Z(G)$. Then

$$ghg^{-1} = hgg^{-1} = he_G = h \in Z(G).$$

b. By assumption there is some $g \in G$ such that $G/Z(G) = \left\langle gZ(G) \right\rangle$ is generated by the coset $gZ(G)$. This, however, implies that

$$G = \bigcup_{k \in \mathbb{Z}} g^k Z(G).$$

Let now $h, h' \in G$ be arbitrary. We then find $k, k' \in Z$ and $u, u' \in Z(G)$ such that $h = g^k u$ and $h' = g^{k'} u'$. Thus, using the exponential laws and the fact that $u$ and $u'$ commute with any element in $G$, we get

$$hh' = g^k u g^{k'} u' = g^k g^{k'} u u' = g^{k'} g^k u' u = g^{k'} u' g^k u = h'h.$$

Thus $G$ is commutative.

**Solution to Exercise 5** a. We define a map

$$\alpha : W \to G : x^{\alpha_1} \cdots y^{\beta_r} \mapsto g^{\alpha_1} \cdots h^{\beta_r}.$$

Note that the representation of an element in $W$ as a word is not unique; we used the exponential laws to identify certain words! We therefore have to check that the definition of $\alpha$ does not depend on the given representation. However,

since the exponential laws also apply in $G$, two representations of the same word will lead to the same image. Hence, $\alpha$ is welldefined.
Moreover, it is indeed clear that $\alpha$ is a homomorphism mapping $x$ to $g$ and $y$ to $h$, and since $G$ is generated by $g$ and $h$ the homomorphism is also surjective.
Let's now show the uniqueness. Let $\beta \in \mathrm{Hom}(W, G)$ be such that $\beta(x) = g$ and $\beta(y) = h$, and let $x^{\alpha_1} \cdots y^{\beta_r} \in W$ be any word. Using the rules for homomorphisms

$$\beta\big(x^{\alpha_1} \cdots y^{\beta_r}\big) = \beta(x)^{\alpha_1} \cdots \beta(y)^{\beta_r} = g^{\alpha_1} \cdots h^{\beta_r} = \alpha\big(x^{\alpha_1} \cdots y^{\beta_r}\big).$$

It remains to show that $\alpha$ induces an epimorphism from $\langle x, y \mid w_1 = w_1', \ldots, w_r = w_r' \rangle$ to $G$, if $g$ and $h$ satisfy the relations $w_i = w_i'$ for $i = 1, \ldots, r$. Note that the latter is the same as saying that

$$w_i^{-1} w_i' \in \mathrm{Ker}(\alpha),$$

and since $\mathrm{Ker}(\alpha)$ is a normal subgroup of $W$ we thus have $N(M) \subseteq \mathrm{Ker}(\alpha)$. We only have to show that the above map $\overline{\alpha}$ is welldefined, then it's clear that it is an epimorphism. Let $w$ and $w'$ be two words which coincide in $\langle x, y \mid w_1 = w_1', \ldots, w_r = w_r' \rangle$, i. e. $wN(M) = w'N(M)$. Thus

$$w^{-1} w' \in N(M) \subseteq \mathrm{Ker}(\alpha),$$

and hence $e_G = \alpha\big(w^{-1}w'\big) = \alpha(w)^{-1} \cdot \alpha(w')$, which implies $\alpha(w) = \alpha(w')$. The morphism $\overline{\alpha}$ is therefore welldefined.

b. Recall from Exercise 1, Set 1, that

$$\mathbb{D}_{2n} = \langle \sigma, \tau \rangle = \big\{ \sigma^i, \tau\sigma^i \;\big|\; i = 0, \ldots, n-1 \big\}$$

and that the generators $\sigma$ and $\tau$ satisfy the relations

$$\sigma^n = (1), \tau^2 = (1) \quad \text{and} \quad \tau\sigma\tau^{-1} = \sigma^{-1}.$$

Hence by Part a. there is an epimorphism

$$\overline{\alpha} : \big\langle x, y \;\big|\; x^n = e, y^2 = e, yxy^{-1} = x^{-1} \big\rangle \longrightarrow \mathbb{D}_{2n}.$$

Once we know that the group $\big\langle x, y \;\big|\; x^n = e, y^2 = e, yxy^{-1} = x^{-1} \big\rangle$ has at most $2n$ elements, we are therefore done, since then the map must be bijective.
The same proof as in Exercise 1, Set 2, applies in order to show

$$\big\langle x, y \;\big|\; x^n = e, y^2 = e, yxy^{-1} = x^{-1} \big\rangle = \big\{ x^i y^j \;\big|\; i = 0, 1; j = 0, \ldots, n-1 \big\}.$$

c. Recall from Exercise 5, Set 2, that

$$\mathbb{Q}_8 = \langle A, B \rangle = \big\{ A^i B^j \mid i = 0, \ldots, 3; j = 0, 1 \big\}$$

where $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & i \\ i & 0 \end{smallmatrix}\right)$. Moreover, we have shown there

$$A^4 = \mathbb{1}, B^4 = \mathbb{1} \quad \text{and} \quad BAB^{-1} = A^{-1}.$$

We may therefore once more apply Part a. in order to get an epimorphism

$$\overline{\alpha} : \big\langle x, y \;\big|\; x^4 = e, y^4 = e, yxy^{-1} = x^{-1} \big\rangle \longrightarrow \mathbb{Q}_8.$$

And the same proof as in Exercise 5, Set 2, shows that

$$\langle x, y \mid x^4 = e, y^4 = e, yxy^{-1} = x^{-1} \rangle = \{x^i y^j \mid i = 0, \ldots, 3; j = 0, 1\}.$$

Hence the group has at most 8 elements, and since $\overline{\alpha}$ is a surjection on a set with 8 elements, this must be a bijection.

d. Since $G$ is not abelian, it is not cyclic, and hence it does not contain any element of order 8. By the Theorem of Lagrange the elements of $G$ must therefore be of order 1, 2 or 4.
Suppose that $G$ does not contain any element of order 4. Then $g^2 = e_G$ for all $g \in G$, and hence $g = g^{-1}$ for all $g \in G$. Let $g, h \in G$ be given. Then

$$gh = (gh)^{-1} = h^{-1}g^{-1} = hg.$$

This means that $G$ is abelian in contradiction to our assumption.
Hence there is some $g \in G$ of order 4. Then $N = \langle g \rangle$ is a subgroup of index 2, and is therefore a normal subgroup. Let $u \in G \setminus N$ and $U = \langle u \rangle$. Since $N \trianglelefteq G$, the set $UN$ is a subgroup of $G$ with more than 4 elements. By the Theorem of Lagrange it must therefore be equal to $G$. That is

$$G = UN = \langle u, g \rangle.$$

Moreover, since $N$ is a normal subgroup, we have $ugu^{-1} \in N$ and this element has the same order as the element $g$, which is 4. There are only two choices for this in $N$, namely $g$ and $g^{-1}$. If $ugu^{-1} = g$, then $ug = gu$. However, if the two generators of $G$ commute, then $G$ is abelian, which it is not by assumption. Therefore

$$ugu^{-1} = g^{-1}.$$

We now have to distinguish two cases. $u$ could be of order 2 or it could be of order 4.
If $o(u) = 2$, then $g$ and $u$ satisfy the relations

$$g^4 = e, u^2 = e \quad \text{and} \quad ugu^{-1} = g^{-1}.$$

Hence by Part a. we get – in the same way as in Part b. – an isomorphism

$$\langle x, y \mid x^4 = e, y^2 = e, yxy^{-1} = x^{-1} \rangle \cong G.$$

But by Part b. this group is also isomorphic to $\mathbb{D}_8$.
If $o(u) = 4$, then $g$ and $u$ satisfy the relations

$$g^4 = e, u^4 = e \quad \text{and} \quad ugu^{-1} = g^{-1}.$$

Hence by Part a. we get – in the same way as in Part c. – an isomorphism

$$\langle x, y \mid x^4 = e, y^4 = e, yxy^{-1} = x^{-1} \rangle \cong G.$$

But by Part c. this group is also isomorphic to $\mathbb{Q}_8$.

# Assignment Set 5

*One* of the Exercises 1 or 3 should be handed in for marking.

**Exercise 1:** Let $(G, \cdot)$ be a group and $N \leq G$.
Show that $N$ is a normal subgroup of $G$ if and only if $G = N_G(N)$.

**Exercise 2:** **[Class Equation]** Let $(G, \cdot)$ be a finite group. We call for $g \in G$ the set $C_G(g) = \{h \in G \mid hg = gh\}$ the *centraliser* of $g$ in $G$.

a. Show that the group $G$ acts on the set $G$ by conjugation, i. e. show that

$$\alpha : G \to \mathrm{Sym}(G) : g \mapsto \alpha_g$$

is a homomorphism, where $\alpha_g : G \to G : h \mapsto h^g = ghg^{-1}$.

b. Show that there are $g_1, \ldots, g_n \in G$ such that

$$|G| = \sum_{i=1}^{n} |G : C_G(g_i)|.$$

c. Show that there $g_1, \ldots, g_r \in G$ such that $|G : C_G(g_i)| > 1$ and

$$|G| = |Z(G)| + \sum_{i=1}^{r} |G : C_G(g_i)|.$$

Hint, show that $C_G(g)$ is just the stabiliser $\mathrm{Stab}_G(g)$ of $g$ under the group operation in a. and use the Orbit Stabilser Theorem. By $Z(G)$ we mean the centre of $G$ introduced in Exercise 4, Set 4.

**Exercise 3:** Let $(G, \cdot)$ be a group of order $p^n$ for some prime $p$. Show that $|Z(G)| > 1$.

Hint, use the class equation.

**Exercise 4:** Show that any group of order $p^2$, $p$ some prime, is abelian.

Hint, use Exercise 3 above and Exercise 4, Set 4.

**Exercise* 5:** Calculate $Z(\mathbb{D}_8)$ and $Z(\mathbb{Q}_8)$.

Hint, you may use Exercise 3 above and Exercise 4, Set 4, in order save many calculations.

**Exercise 6:** Use Corollary 4.7 to show that $\mathbb{D}_8$ is not a normal subgroup of $\mathbb{S}_4$.

**Solution to Exercise 1**
By definition $N_G(N) = \left\{ g \in G \;\middle|\; N = N^g = gNg^{-1} \right\}$. However again by definition, $N$ is normal if and only if $N^g = N$ for all $g \in G$, and this is then the case if and only if $g \in N_G(N)$ for all $g \in G$, i. e. $G = N_G(N)$.

**Solution to Exercise 2** a. Note, we have already shown in the lecture that the maps $\alpha_g$ are automorphisms of $G$, hence they are in particular bijective and belong to $\mathrm{Sym}(G)$.
We have to show that $\alpha(g \cdot g') = \alpha(g) \circ \alpha(g')$, or with the above notation $\alpha_{gg'} = \alpha_g \circ \alpha_{g'}$. Let $h \in G$. Then

$$\alpha_{gg'}(h) = (gg')\cdot h \cdot (gg')^{-1} = gg'hg'^{-1}g^{-1} = \alpha_g\big(g'hg'^{-1}\big) = \alpha_g\big(\alpha_{g'}(h)\big) = (\alpha_g \circ \alpha_{g'})(h).$$

b. We note that $\mathrm{Stab}_G(g) = \{h \in G \mid g^h = g\} = C_G(g)$, and by the Orbit Stabiliser Theorem we therefore have $|\mathrm{orb}_G(g)| = |G : C_G(g)|$.
Since $G$ operates on $G$, there are $g_1, \ldots, g_n \in G$ such that $G = \coprod_{i=1}^{n} \mathrm{orb}_G(g_i)$ is the disjoint union of the orbits of the $g_i$. Thus

$$|G| = \sum_{i=1}^{n} |\mathrm{orb}_G(g_i)| = \sum_{i=1}^{n} |G : C_G(g_i)|.$$

c. Note that $g \in Z(G)$ if and only if $gh = hg$ for all $h \in G$ if and only if $g = hgh^{-1} = g^h$ for all $h \in G$ if and only if $\mathrm{orb}_G(g) = \left\{g^h \;\middle|\; h \in G\right\} = \{g\}$ consists only of one element.
Let $g_1, \ldots, g_n \in G$ be as in Part b. and suppose that they have been ordered in such a way that the orbits of $g_1, \ldots, g_r$ consist of more than one element and the orbits of $g_{r+1}, \ldots, g_n$ all contain only one element. We have just proved that then

$$Z(G) = \bigcup_{i=r+1}^{n} \mathrm{orb}_G(g_i).$$

The result therefore follows from Part b.

**Solution to Exercise 3**

By the Class Equation we know there are $g_1, \ldots, g_r \in G$ such that

$$p^n = |G| = |Z(G)| + \sum_{i=1}^{r} |G : C_G(g_i)|, \tag{13}$$

and $|G : C_G(g_i)| > 1$ for all $i = 1, \ldots, r$. Since this index $|G : C_G(g_i)|$ is a divisor of $|G| = p^n$, it must be divisible by $p$. Considering the Equation (13) modulo $p$ we get

$$|Z(G)| \equiv 0 (\mathrm{mod}\ p).$$

Thus the number must be divisible by $p$ as well, in particular it is greater than $1$.

**Solution to Exercise 4**

Let $G$ be a group of order $p^2$. By Exercise 3 the centre $Z(G)$ has order greater than $1$, and by the Theorem of Lagrange its order must then be $p$ or $p^2$. If the order is $p^2$, then $G = Z(G)$ and $G$ is abelian.
Suppose therefore the order of $Z(G)$ was $p$, then however $G/Z(G)$ has also order $p$ and is therefore cyclic. Hence by Exercise 4, Set 4, $G$ is again abelian. (However, this implies $Z(G) = G$ and its order is $p^2$ in contradiction to our assumption! That is, this case will not occur.)

**Solution to Exercise 5**

For both groups $\mathbb{D}_8$ and $\mathbb{Q}_8$ the centre must have order $1$, $2$, $4$ or $8$ by the Theorem Lagrange. By Exercise 3 above it cannot be $1$.
If the order was $8$, that is, if the centre was the whole group, the group would be abelian, which both are not.
If it had order 4, then the quotient group by the centre would be of order $2$ and hence cyclic. But then again, by Exercise 4, Set 4, the group itself would be abelian, which it is not.

Thus it must have order 2 in both cases. It therefore suffices in both cases to find one element of order two which commutes with all the other elements, and this one will then generate the centre.

Using the notation from Exercise 1, Set 1, and Exercise 5, Set 2, we see that $\sigma^2 \in \mathbb{D}_8$ and $A^2 = B^2 \in \mathbb{Q}_8$ will do, i. e.

$$Z(\mathbb{D}_8) = \langle \sigma^2 \rangle \quad \text{and } Z(\mathbb{Q}_8) = \langle A^2 \rangle.$$

**Solution to Exercise 6**

Besides $\mathbb{D}_8 = \langle (1\ 2\ 3\ 4), (1\ 4)(2\ 3) \rangle$ the group $\mathbb{S}_4$ contains two other subgroups of order 8 – both of which are isomorphic to $\mathbb{D}_8$ – namely $\langle (1\ 2\ 4\ 3), (1\ 3)(2\ 4) \rangle$ and $\langle (1\ 3\ 2\ 4), (1\ 4)(2\ 3) \rangle$. Since $8 = 2^3$ is the maximal power of 2 which divides the order of $\mathbb{S}_4$, which is $4! = 24$, these three groups are 2-Sylow subgroups of $\mathbb{S}_4$. Having more than one 2-Sylow subgroup, none of them can be normal by Corollary 4.7.

## Assignment Set 6

*One* of the Exercises 1 or 3 should be handed in for marking.

**Exercise 1:** Calculate the eigenvalues and the eigenspaces of the following matrix and decide whether it is diagonalisable, triangulable or neither of the two:

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}(4 \times 4, \mathbb{R}).$$

**Exercise 2:** Let $V = \left\{ \sum_{i=0}^{2} a_i x^i \;\middle|\; a_0, a_1, a_2 \in \mathbb{R} \right\}$ be the 3-dimensional vector space of polynomials of degree less than or equal to 2, and let $\lambda \in \mathbb{R}$ be fixed. Consider the map

$$f : V \to V : \sum_{i=0}^{2} a_i x^i \mapsto \sum_{i=0}^{2} a_i (x+\lambda)^i - \lambda \cdot \sum_{i=1}^{2} a_i \cdot i \cdot x^{i-1}.$$

Show the following:

a. $f$ is $\mathbb{R}$-linear.

b. Calculate $M_B^B(f)$, where $B = \left(1, x, x^2\right)$ is the canonical basis of $V$.

c. Calculate the characteristic polynomial $\chi_f$.

**Exercise 3:** Let $V$ be an $n$-dimensional $K$-vector space, and let $f \in \mathrm{End}_K(V)$ such that $f^{n-1} \neq 0$, but $f^n = 0$, where $0$ means the zero-map. Show:

a. There is a $v \in V$ such that $B = \left(f^{n-1}(v), f^{n-2}(v), \ldots, f(v), v\right)$ is a basis of $V$.

b. Find the matrix representation $M_B^B(f)$ w. r. t. the basis $B$ in Part a.

**Exercise 4:** Let $V$ be a finite-dimensional $K$-vector space and let $g \in \mathrm{End}_K(V)$. Show there is an $m \geq 1$ such that

$$\mathrm{Ker}(g) \subsetneq \mathrm{Ker}\left(g^2\right) \subsetneq \ldots \subsetneq \mathrm{Ker}\left(g^m\right) = \mathrm{Ker}\left(g^k\right) \quad \text{for all} \quad k \geq m.$$

**Exercise 5:** Use the Theorem of Cayley-Hamilton to show that for $A \in \mathrm{Gl}_n(K)$ there is a polynomial $g = \sum_{i=0}^{n-1} b_i t^i \in K[t]$ such that $A^{-1} = g(A) = \sum_{i=0}^{n-1} b_i A^i$.

**Exercise 6:** Let $f \in \mathrm{End}_K(V)$ and let $\lambda, \mu \in K$ be two different eigenvalues of $f$. Show that for any $m \geq 0$

$$\mathrm{Eig}(f, \lambda) \cap \mathrm{Ker}\left((f - \mu\,\mathrm{id}_V)^m\right) = \{0\}.$$

**Solution to Exercise 1**

$\chi_A = \det(A - t\mathbb{1}) = (3-t) \cdot (2-t)^3$, and hence $A$ is triangulable, since the characteristic polynomial factorises. $\mathrm{Eig}(A, 2) = \left\langle (1,0,0,1)^t, (0,0,1,0)^t \right\rangle$, $\mathrm{Eig}(A, 3) = \left\langle (1,1,1,1)^t \right\rangle$.

$A$ is not diagonalisable, since $K^4$ does not posses a basis of eigenvectors of $A$.

**Solution to Exercise 2** a. Let $p = a_2x^2 + a_1x + a_0, q = b_2x^2 + b_1x + b_0 \in V$ and $\nu, \mu \in \mathbb{R}$ be given. Then

$$f(\mu p+\nu q) = f\left(\sum_{i=0}^{2}(a_i + b_i)\cdot x^i\right) = \sum_{i=0}^{2}(\mu a_i+\nu b_i)\cdot(x+\lambda)^i-\lambda\cdot\sum_{i=1}^{2}(\mu a_i+\nu b_i)\cdot i\cdot x^{i-1}$$

$$= \mu\left(\sum_{i=0}^{2} a_i(x+\lambda)^i - \lambda\cdot\sum_{i=1}^{2} a_i\cdot i\cdot x^{i-1}\right)+\nu\left(\sum_{i=0}^{2} b_i(x+\lambda)^i - \lambda\cdot\sum_{i=1}^{2} b_i\cdot i\cdot x^{i-1}\right) = \mu f(p)+\nu f(q).$$

b. Note that $f(1) = 1$, $f(x) = (x+\lambda)-\lambda = x$ and $f\left(x^2\right) = (x+\lambda)^2-2\lambda x = x^2+\lambda^2$, hence

$$M_B^B(f) = \begin{pmatrix} 1 & 0 & \lambda^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

c. $\chi_f = \chi_{M_B^B(f)} = (1-t)^3$.

**Solution to Exercise 3** a. By assumption $f^{n-1} \neq 0$, so there is a $v \in V$ such that $f^{n-1}(v) \neq 0$. Define $B$ as in the claim using this vector $v$.

Since $V$ has dimension $n$ it suffices to show that $B$ is linearly independent. For this let $\lambda_1, \ldots, \lambda_n \in K$ such that $\sum_{i=1}^{n} \lambda_i \cdot f^{n-i}(v) = 0$. We have to show that $\lambda_1 = \ldots = \lambda_n = 0$.

Suppose this is not the case and let $m \in \{1, \ldots, n\}$ be minimal such that $\lambda_m \neq 0$. Then

$$0 = f^{m-1}(0) = f^{m-1}\left(\sum_{i=1}^{n}\lambda_i f^{n-i}(v)\right)$$

$$= \sum_{i=1}^{m-1}\lambda_i f^{n-i+m-1}(v) + \lambda_m f^{n-m+m-1}(v) + \sum_{i=m+1}^{n}\lambda_i f^{n-i+m-1}(v) = \lambda_m f^{n-1}(v),$$

where the first sum is zero since $\lambda_1 = \ldots = \lambda_{m-1} = 0$, and the last sum vanishes, since $f^n$ is the zero-map. However, by assumption neither $\lambda_m$ nor $f^{n-1}(v)$ vanish, which leads to a contradiction.

b. Since $f\left(f^k(v)\right) = f^{k+1}(v)$ and this is $0$, if $k = n-1$, we get

$$M_B^B(f) = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \ldots & \ldots & \ldots & 0 \end{pmatrix}.$$

**Solution to Exercise 4**

If $v \in \operatorname{Ker}\left(g^k\right)$, then $g^{k+1}(v) = g\left(g^k(v)\right) = g(0) = 0$. Thus for all $k \geq 1$ we have

$$\operatorname{Ker}\left(g^k\right) \subseteq \operatorname{Ker}\left(g^{k+1}\right). \tag{14}$$

Moreover, since the vector space $V$ is finite dimensional, the chain of kernels cannot ascend forever. Let therefore

$$m = \min\left\{k \geq 1 \;\middle|\; \operatorname{Ker}\left(g^k\right) = \operatorname{Ker}\left(g^{k+1}\right)\right\}.$$

We have to show that then $\operatorname{Ker}\left(g^m\right) = \operatorname{Ker}\left(g^k\right)$ for all $k \geq m$, and we do this by induction on $k$. We get the induction base $k = m$ for free. Let's now suppose that $k > m$ and that we have already shown $\operatorname{Ker}\left(g^m\right) = \operatorname{Ker}\left(g^{k-1}\right)$. By Equation (14) we thus get $\operatorname{Ker}\left(g^m\right) = \operatorname{Ker}\left(g^{k-1}\right) \subseteq \operatorname{Ker}\left(g^k\right)$.
It remains to prove the opposite inclusion. Let therefore $v \in \operatorname{Ker}\left(g^k\right)$ be given. Then $0 = g^k(v) = g^{k-1}\big(g(v)\big)$. Hence $g(v) \in \operatorname{Ker}\left(g^{k-1}\right) = \operatorname{Ker}\left(g^m\right)$, and thus

$$g^{m+1}(v) = g^m\big(g(v)\big) = 0.$$

However, by definition of $m$ we have $\operatorname{Ker}\left(g^m\right) = \operatorname{Ker}\left(g^{m+1}\right)$ and thus we have shown

$$v \in \operatorname{Ker}\left(g^{m+1}\right) = \operatorname{Ker}\left(g^m\right).$$

**Solution to Exercise 5**
Let $\chi_A = (-1)^n t^n + a_{n-1}t^{n-1} + \ldots + a_0 \in K[t]$ be the characteristic polynomial of $A$.
Since $A$ is invertible, the kernel of $A$ consists only of the zero-vector. Hence $\operatorname{Eig}(A, 0) = \operatorname{Ker}(A) = \{0\}$, which implies that $0$ is not an eigenvalue of $A$. Hence $a_0 = \chi_A(0) \neq 0$. Define

$$g = \tfrac{(-1)^n}{-a_0} \cdot t^{n-1} + \tfrac{a_{n-1}}{-a_0} \cdot t^{n-2} + \ldots + \tfrac{a_1}{-a_0} \in K[t].$$

Then

$$g(A) \cdot A = \tfrac{1}{-a_0} \cdot \chi_A(A) + \mathbb{1} = \mathbb{1},$$

where the latter equality is due to the Thm. of Cayley-Hamilton. Thus $g(A) = A^{-1}$.

**Solution to Exercise 6**
We do the proof by induction on $m \geq 0$. For $m = 0$ the kernel of $(f - \mu \operatorname{id}_V)^0 = \operatorname{id}_V$ consists only of the zero-vector, so there is nothing to show.
Let now $m > 0$ and suppose the claim has been proved for $m - 1$. Let then $v \in \operatorname{Eig}(f, \lambda) \cap \operatorname{Ker}\left((f - \mu \operatorname{id}_V)^m\right)$, we have to show the $v = 0$. By assumption

$$\begin{aligned} 0 = (f - \mu \operatorname{id}_V)^m(v) &= (f - \mu \operatorname{id}_V)^{m-1}\big((f - \mu \operatorname{id}_V)(v)\big) = (f - \mu \operatorname{id}_V)^{m-1}\big(f(v) - \mu v\big) \\ &= (f - \mu \operatorname{id}_V)^{m-1}\big((\lambda - \mu) \cdot v\big) = (\lambda - \mu) \cdot (f - \mu \operatorname{id}_V)^{m-1}(v). \end{aligned}$$

Since $\lambda \neq \mu$ this implies $(f - \mu \operatorname{id}_V)^{m-1}(v) = 0$, and therefore

$$v \in \operatorname{Eig}(f, \lambda) \cap \operatorname{Ker}\left((f - \mu \operatorname{id}_V)^{m-1}\right) = \{0\}.$$

# Assignment Set 7

Exercise 3 should be handed in for marking.

**Exercise 1:** Find a Jordan normal form and the corresponding transformation matrix $T$ for the following matrices:

$$A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 1 \\ -1 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 1 \\ 2 & 1 & -1 \\ -6 & -5 & -3 \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} -3 & -1 & 1 \\ -1 & -3 & 1 \\ -2 & -2 & 0 \end{pmatrix}.$$

**Exercise 2:** Find a Jordan normal form for the Endomorphism in Exercise 2 on Assignment Set 6.

**Exercise 3:** Let $V$ be a $K$-vector space, $U \leq V$ a subspace, and $b \in \mathrm{Bil}_K(V)$. Show

a. $U^\perp = \{v \in V \mid b(v,u) = 0 \ \forall\, u \in U\}$ is a subspace of $V$.

b.* If $U = \langle v \rangle$ and $b(v,v) \neq 0$, then $V = U + U^\perp$.

**Exercise 4:** Consider $b : K^2 \to K^2 : \big((x_1,x_2)^t, (y_1,y_2)^t\big) \mapsto 2 \cdot x_1 \cdot y_1 + x_1 \cdot y_2 + y_1 \cdot x_2 - x_2 \cdot y_2$. Let $E = (e_1, e_2)$ be the standard basis of $K^2$ and $B = (v_1, v_2)$ with $v_1 = (1,1)^t$ and $v_2 = (1,-1)^t$ some other basis.

a. Show that $b$ is a bilinear map.

b. Calculate the matrix representations $M_E(b)$ and $M_B(b)$.

c. Calculate the transformation matrix $T_B^E$ and verify

$$M_B(b) = \big(T_B^E\big)^t \cdot M_E(b) \cdot T_B^E.$$

**Exercise 5:** Let $b \in \mathrm{Bil}_K(V)$ be a bilinear form and $q = q_b : V \to K : v \mapsto b(v,v)$ its associated quadratic form. Show that for all $u, v, w \in V$

$$q(u+v+w) - q(u+v) - q(v+w) - q(u+w) + q(u) + q(v) + q(w) = 0.$$

**Solution to Exercise 1**

$$\chi_A = (1-t)\cdot(2-t), \quad \chi_B = (5-t)^2, \quad \chi_C = (2+t)^2\cdot(2-t), \quad \chi_D = (2+t)^3.$$

$$A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}, \quad J(A) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \quad T^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

$$B = \begin{pmatrix} 4 & 1 \\ -1 & 6 \end{pmatrix}, \quad J(B) = \begin{pmatrix} 5 & 1 \\ 0 & 5 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \quad T^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}.$$

$$C = \begin{pmatrix} 0 & 1 & 1 \\ 2 & 1 & -1 \\ -6 & -5 & -3 \end{pmatrix}, \quad J(C) = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & -1 & 0 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \quad T^{-1} = \begin{pmatrix} 1 & 1 & 1 \\ -2 & -1 & -1 \\ -1 & -1 & 0 \end{pmatrix}.$$

$$D = \begin{pmatrix} -3 & -1 & 1 \\ -1 & -3 & 1 \\ -2 & -2 & 0 \end{pmatrix}, \quad J(D) = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & 1 & -1 \\ -1 & 0 & 1 \\ -2 & 0 & 0 \end{pmatrix} \quad T^{-1} = \begin{pmatrix} 0 & 0 & -\frac{1}{2} \\ 1 & 1 & -1 \\ 0 & 1 & -\frac{1}{2} \end{pmatrix}.$$

**Solution to Exercise 2**
We showed in Exercise 2 on Assignment Set 6 that for $B = (1, x, x^2)$

$$M_B^B(f) = \begin{pmatrix} 1 & 0 & \lambda^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus $\dim_{\mathbb{R}}\big(\mathrm{Eig}(f,1)\big) = \dim_{\mathbb{R}}\big(\mathrm{Ker}(f - \mathrm{id}_V)\big) = 3 - \mathrm{rank}\,\big(M_B^B(f) - \mathbb{1}\big) = 2$, which implies that the following matrix is a Jordan normal form for $f$:

$$J(f) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Solution to Exercise 3** a. Since $b(0,u) = b(0+0,u) = b(0,u) + b(0,u)$ for any $u \in U$, we see that $b(0,u) = 0$ for any $u \in U$. Therefore $0 \in U^\perp$ and the latter is non-empty. Let now $v, w \in U^\perp$, $\lambda, \mu \in \mathbb{K}$ and $u \in U$. Then

$$b(\lambda v + \mu w, u) = \lambda b(v,u) + \mu b(w,u) = 0,$$

and hence $\lambda v + \mu w \in U^\perp$. This shows that $U^\perp$ is a subspace of $V$.

b. Let $w \in V$ be arbitrary. We have to show that $w$ is a sum of a vector in $U$ and one in $U^\perp$. Set $\lambda = \frac{b(v,w)}{b(v,v)} \in K$ and $u = w - \lambda \cdot v$. Then

$$b(v,u) = b(v, w - \lambda v) = b(v,w) - \lambda b(v,v) = 0.$$

Hence, $u \in U^\perp$, and $w = \lambda v + u \in U + U^\perp$. This proves the claim.

**Solution to Exercise 4** a. Note that $b\big((x_1,x_2)^t, (y_1,y_2)^t\big) = (x_1,x_2)\cdot A \cdot (y_1,y_2)^t$, where $A = \left(\begin{smallmatrix} 2 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. Since matrix multiplication is distributive, $b$ is a bilinear form, and since $A = A^t$, i. e. since $A$ is symmetric,

$$b(\underline{x},\underline{y}) = \underline{x}^t \cdot A \cdot \underline{y} = \underline{x}^t \cdot A^t \cdot \underline{y} = (A\cdot\underline{x})^t \cdot \underline{y} = \big((A\cdot\underline{x})^t \cdot \underline{y}\big)^t = \underline{y}^t \cdot (A\cdot\underline{x}) = b(\underline{y},\underline{x}).$$

Hence, $b$ is symmetric.

b. Just calculating $b(e_i, e_j)$ and $b(v_i, v_j)$ for all $i, j$ we get

$$M_E(b) = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} = A \quad \text{and} \quad M_B(b) = \begin{pmatrix} 3 & 3 \\ 3 & -1 \end{pmatrix}.$$

c. The base change $T_B^E$ has the vectors of $B$ as column vectors, since $E$ is the standard basis, thus

$$T_B^E = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad M_B(b) = \big(T_B^E\big)^t \cdot M_E(b) \cdot T_B^E.$$

**Solution to Exercise 5**
Let $u, v, w \in V$ be given. Note that

$$\begin{aligned} q(u+v+w) &= b(u+v+w, u+v+w) \\ &= b(u,u) + 2b(u, v+w) + b(v+w, v+w) = q(u) + 2b(u,v) + 2b(u,w) + q(v+w) \end{aligned}$$

and

$$q(u+v) = b(u+v, u+v) = b(u,u) + 2b(u,v) + b(v,v) = q(u) + 2b(u,v) + q(v)$$

and

$$q(u+w) = b(u+w, u+w) = b(u,u) + 2b(u,w) + b(w,w) = q(u) + 2b(u,w) + q(w)$$

Using these results we finally get

$$\begin{aligned} &q(u+v+w) - q(u+v) - q(v+w) - q(u+w) + q(u) + q(v) + q(w) \\ &\quad = \big(q(u) + 2b(u,v) + 2b(u,w) + q(v+w)\big) - \big(q(u) + 2b(u,v) + q(v)\big) \\ &\qquad - q(v+w) - \big(q(u) + 2b(u,w) + q(w)\big) + q(u) + q(v) + q(w) = 0. \end{aligned}$$

# Assignment Set 8

Exercise 1 should be handed in for marking.

**Exercise* 1:** We call a bilinear form $b \in \mathrm{Bil}_{\mathbb{R}}(V)$ *positive definite* if and only if $b(v,v) > 0$ for all $0 \neq v \in V$. Let $A = \left(\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right) \in \mathrm{Mat}(2\times 2, \mathbb{R})$ be a symmetric matrix. Show, the bilinear form $b_A$ is positive definite if and only if $a_{11} > 0$ and $\det(A) > 0$.

Hint, use Corollary 2.11 to find a $T \in \mathrm{Gl}_2(\mathbb{R})$ such that $T^t \cdot A \cdot T = \left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right)$ with $a, b \in \{-1, 0, 1\}$, and note that $T^t \cdot A \cdot T = \left(b_A(\underline{t_i}, \underline{t_j})\right)_{i,j=1,2}$ if $\underline{t_k}$ denotes the k-th column of $T$.

**Exercise 2:** Calculate the rank, the index and the signature of the bilinear form corresponding to the following symmetric matrices:

$$A = \begin{pmatrix} -1 & 4 \\ 4 & -16 \end{pmatrix} \in \mathrm{Mat}(2\times 2, \mathbb{R}) \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \in \mathrm{Mat}(4\times 4, \mathbb{R}).$$

**Exercise 3:** Use the Algorithm of Gram-Schmidt to calculate an ONB of the subspace $\left\langle (1,-1,1,-1)^t, (1,0,1,0)^t, (2,2,1,0)^t \right\rangle$ of $K^4$ w. r. t. the standard scalar product.

**Exercise 4:** Let $\left(V, \langle\cdot,\cdot\rangle\right)$ be a Hilbert space and $U \leq V$ a subspace of $V$. Show that

a. $U^\perp \leq V$ is a subspace of $V$.

b. $V = U \oplus U^\perp$, i. e. $V = U + U^\perp$ and $U \cap U^\perp = \{0\}$.

Hint, in b. show first that $U \cap U^\perp = \{0\}$ and calculate then the dimension of $U + U^\perp$, taking Gram-Schmidt into consideration.

**Exercise 5:** Find for the following matrix an orthogonal matrix $T$ which diagonalises it:

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 0 & -1 & 0 \\ -2 & 0 & -1 \end{pmatrix}.$$

**Exercise 6:** Let $f \in \mathrm{End}_K(V)$, $\left(V, \langle\cdot,\cdot\rangle\right)$ a finite-dimensional Hilbert space. Show there is a unique endomorphism[6] $f^* \in \mathrm{End}_K(V)$ such that for all $v, w \in W$

$$\langle f(v), w\rangle = \langle v, f^*(w)\rangle.$$

[6]Note, $f^*$ is called the *adjoint* of $f$, and $f$ is self-adjoint if and only if $f = f^*$!

**Solution to Exercise 1**

We start by collecting some useful remarks. By Corollary 2.11 there is an invertible matrix $T = \left(\begin{smallmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{smallmatrix}\right) \in Gl_2(\mathbb{R})$ such that

$$T^t \cdot A \cdot T = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \tag{15}$$

with $a, b \in \{-1, 0, 1\}$. This gives

$$\det(A) \cdot \det(T)^2 = \det\left(T^t \cdot A \cdot T\right) = a \cdot b \tag{16}$$

The columns $\underline{t_1} = (t_{11}, t_{21})^t$ and $\underline{t_2} = (t_{12}, t_{22})^t$ of $T$ form a basis of $K^2$, since $T$ is invertible, and we have

$$b_A(\underline{t_1}, \underline{t_1}) = a, \quad b_A(\underline{t_1}, \underline{t_2}) = b_A(\underline{t_2}, \underline{t_1}) = 0 \quad \text{and} \quad b_A(\underline{t_2}, \underline{t_2}) = b. \tag{17}$$

Let $v = \lambda_1 \underline{t_1} + \lambda_2 \underline{t_2}$. Then

$$b_A(v, v) = \lambda_1^2 a + \lambda_2^2 b. \tag{18}$$

If $b_A$ is positive definite, then

$$a_{11} = b_A(e_1, e_1) > 0$$

and in view of Equation (16) and (17) we have

$$\det(A) \cdot \det(T)^2 = b_A(\underline{t_1}, \underline{t_1}) \cdot b_A(\underline{t_2}, \underline{t_2}) > 0,$$

and hence also $\det(A) > 0$.

Suppose now, vice versa, that $a_{11} > 0$ and $\det(A) > 0$. We have to show that $b_A(v, v) > 0$ for all $v \in K^2$.

Note first of all, that $\det(A) > 0$ implies either $a = b = 1$ or $a = b = -1$ in view of Equation (15) and (16). Let $\lambda_1, \lambda_2 \in \mathbb{R}$ be such that $e_1 = \lambda_1 \underline{t_1} + \lambda_2 \underline{t_2}$, then

$$0 < a_{11} = b_A(e_1, e_2) = \lambda_1^2 a + \lambda_2^2 b.$$

Thus we must have $a = b = 1$.

Let now $v = \lambda_1 \underline{t_1} + \lambda_2 \underline{t_2} \in K^2$ be arbitrary, then by Equation (18) $b_A(v, v) = \lambda_1^2 + \lambda_2^2 > 0$, since $a = b = 1$. Hence $b$ is positive definite.

**Solution to Exercise 2**

The normal forms of $A$ and $B$ may be calculated using the symmetric Gauß-Algorithm[7] and turn out to be

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

Thus $\operatorname{rank}(A) = 1$, $\operatorname{index}(A) = 0$ and $\operatorname{signature}(A) = -1$, while $\operatorname{rank}(B) = 4$, $\operatorname{index}(B) = 2$ and $\operatorname{signature}(B) = 0$.

[7] For $B$ do the following row-column-operations (R/C): 1) $II \mapsto II - I$; 2) $III \mapsto III - II$; 3) $I \mapsto I + \frac{1}{2}II$; 4) $III \mapsto III - I$; 5) $II \mapsto II + \frac{1}{2}IV$; 6) $IV \mapsto IV - II$.

**Solution to Exercise 3**
Let $u_1 = (1,-1,1,-1)^t$, $u_2 = (1,0,1,0)^t$ and $u_3 = (2,2,1,0)^t$. Using the Algorithm of Gram-Schmidt, we get:

$$v_1 = \tfrac{1}{\|u_1\|} \cdot u_1 = \tfrac{1}{2} \cdot (1,-1,1,-1)^t.$$

We set

$$v_2' = u_2 - \langle u_2, v_1\rangle \cdot v_1 = (1,0,1,0)^t - \tfrac{1}{2} \cdot (1,-1,1,-1)^t = \tfrac{1}{2} \cdot (1,1,1,1)^t$$

and then again

$$v_2 = \tfrac{1}{\|v_2'\|} \cdot v_2' = \tfrac{1}{2} \cdot (1,1,1,1)^t.$$

And finally

$$v_3' = u_3 - \langle u_3, v_1\rangle \cdot v_1 - \langle u_3, v_2\rangle \cdot v_2 = (2,2,1,0)^t - \tfrac{1}{4} \cdot (1,-1,1,-1)^t - \tfrac{5}{4} \cdot (1,1,1,1)^t = \tfrac{1}{2} \cdot (1,2,-1,-2)^t.$$

Hence

$$v_3 = \tfrac{1}{\|v_3'\|} \cdot v_3' = \tfrac{1}{\sqrt{10}} \cdot (1,2,-1,-2)^t.$$

And $(v_1, v_2, v_3)$ is an ONB of $U = \langle u_1, u_2, u_3\rangle$.

**Solution to Exercise 4** a. Since $0$ is orthogonal to every vector, $0 \in U^\perp$ and the latter is non-empty. Let now $v, w \in U^\perp$, $\lambda, \mu \in \mathbb{K}$ and $u \in U$. Then

$$\langle \lambda v + \mu w, u\rangle = \lambda\langle v, u\rangle + \mu\langle w, u\rangle = 0,$$

and hence $\lambda v + \mu w \in U^\perp$. This shows that $U^\perp$ is a subspace of $V$.

b. We show first that $U \cap U^\perp = \{0\}$. If $v \in U \cap U^\perp$, then $\langle v, v\rangle = 0$. Since the scalar product is definite, this implies $v = 0$.
It remains to show that $V = U + U^\perp$. Extend an ONB $(v_1, \ldots, v_r)$ of $U$ to an ONB $(v_1, \ldots, v_n)$ of $V$. Then $v_{r+1}, \ldots, v_n \in U^\perp$ and hence

$$\dim_{\mathbb{K}}\left(U^\perp\right) \geq n - r = \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(U).$$

By the dimension formula for $U + U^\perp$ we therefore get

$$\dim_{\mathbb{K}}(V) \geq \dim_{\mathbb{K}}\left(U + U^\perp\right) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}\left(U^\perp\right) - \dim_{\mathbb{K}}\left(U \cap U^\perp\right) \geq \dim_{\mathbb{K}}(V).$$

This, however, implies $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}\left(U + U^\perp\right)$ and $V = U + U^\perp$.

**Solution to Exercise 5**
Note first, since $A$ is symmetric, hence self-adjoint over $\mathbb{R}$, there exists an orthogonal matrix $T$ such that $T^{-1} \cdot A \cdot T$ is a diagonal matrix, and the columns of $T$ are an ONB of eigenvectors of $\mathbb{R}^3$.
We have $\chi_A = (3-t) \cdot (1-t)^2$, so that

$$T^{-1} \cdot A \cdot T = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

In order to calculate $T$, we have to calculate the eigenspaces w. r. t. $-1$ and $3$, and then to use Gram-Schmidt to orthonormalize them. This gives

$$\mathrm{Eig}(A,-1) = \left\langle (0,1,0)^t, \tfrac{1}{\sqrt{2}} \cdot (1,0,1)^t \right\rangle \quad \text{and} \quad \mathrm{Eig}(A,3) = \left\langle \tfrac{1}{\sqrt{2}} \cdot (1,0,-1)^t \right\rangle.$$

Thus the following matrix will do:

$$T = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

**Solution to Exercise 6**

Let's first show the *existence* of $f^*$. For this we choose an ONB $B = (v_1, \ldots, v_n)$ of $V$. We define $f^*$ on the basis vectors

$$f^*(v_j) = \sum_{i=1}^{n} \langle v_j, f(v_i) \rangle \cdot v_i \tag{19}$$

for $i = 1, \ldots, n$. By linear continuation this defines an endomorphism $f^* \in \mathrm{End}_K(V)$, i. e. if $v = \sum_{i=1}^{n} \lambda_i v_i \in V$, then

$$f^*(v) = \sum_{j=1}^{n} \sum_{i=1}^{n} \lambda_j \cdot \langle v_j, f(v_i) \rangle \cdot v_i.$$

We have to show that $\langle f(v), w \rangle = \langle v, f^*(w) \rangle$ for all $v, w \in V$.
By the Parseval-Equation we have

$$f^*(v_j) = \sum_{i=1}^{n} \langle f^*(v_j), v_i \rangle \cdot v_i. \tag{20}$$

The uniqueness of the basis representation of a vector gives in view of Equation (19) and (20) therefore

$$\langle v_j, f(v_i) \rangle = \langle f^*(v_j), v_i \rangle$$

or equivalently

$$\langle f(v_i), v_j \rangle = \langle v_i, f^*(v_j) \rangle$$

for all $i, j = 1, \ldots, n$. Let now $v = \sum_{i=1}^{n} \lambda_i v_i$ and $w = \sum_{j=1}^{n} \mu_j v_j$ be given, then

$$\langle f(v), w \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \overline{\mu_j} \langle f(v_i), v_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \overline{\mu_j} \langle v_i, f^*(v_j) \rangle = \langle v, f^*(w) \rangle.$$

It remains to show the uniqueness. Let therefore $f' \in \mathrm{End}_K(V)$ be any endomorphism such that

$$\langle f(v), w \rangle = \langle v, f'(w) \rangle \tag{21}$$

for all $v, w \in V$. We have to show $f'(w) = f^*(w)$ for all $w \in V$. From Equation (21) if follows that

$$0 = \langle v, f'(w) \rangle - \langle v, f^*(w) \rangle = \langle v, f'(w) - f^*(w) \rangle$$

for all $v, w \in V$. Fix $w$ and choose now $v = f'(w) - f^*(w)$, then

$$\langle f'(w) - f^*(w), f'(w) - f^*(w) \rangle = 0,$$

which implies $f'(w) - f^*(w) = 0$, since the scalar product is definite.

# Bibliography


[Bu] Burnside, W.: *Theory of Groups of Finite Order.* Dover $^2$1911.

[Doe3] Doerk, Klaus und Trevor Hawkes: *Finite Soluble Groups.* (De Gruyter Expositions in Mathematics, 4) New York 1992.

[Go1] Gorenstein, Daniel: *Finite Groups.* New York $^2$1980.

[Go2] Gorenstein, Daniel: *Finite Simple Groups.* New York 1982.

[Go3] Gorenstein, Daniel: *The Classification of Finite Simple Groups.* Vol. 1. New York 1983.

[Go4] Gorenstein, Daniel: *The Classification of Finite Simple Groups.* Vol. 2. New York 1996.

[Hum] Humphreys, John F.: *A Course in Group Theory.* Oxford 1996.

[Su] Suzuki, Michio: *Group Theory I.* (Die Grundlehren der mathematischen Wissenschaften, Bd. 247) Berlin 1982.

[We] Weinstein, Michael: *Examples of Groups.* Passaic 1977.

[HK] Hoffman, Kenneth and Ray Kunze: *Linear Algebra.* Englewood Cliffs$^2$ 1971.