Combinatorial Group Theory

Lecture notes summer semester 2022

Benjamin Sambale
Leibniz University Hannover

November 4, 2026

$$M_{11} = \langle x, y, z \mid x^{11} = y^5 = z^4 = (xz)^3 = 1,\ yxy^{-1} = x^4,\ zyz^{-1} = y^2 \rangle$$

x	x^{-1}	y	y^{-1}	z	z^{-1}
2	3	1	1	4	4
5	1	6	7	8	9
1	10	11	12	11	13
14	15	4	4	1	1
6	2	13	11	5	5
7	5	10	2	16	16
17	6	2	17	14	18
19	9	9	18	10	2
8	16	14	8	2	10
3	11	17	6	9	8
10	12	5	3	12	3
11	13	3	13	13	11
12	17	12	5	3	12
20	4	16	9	17	7
4	18	21	22	21	19
9	20	18	14	6	6
13	7	7	10	18	14
15	21	8	16	7	17
22	8	22	20	15	22
16	14	19	21	20	20
18	22	20	15	22	15
21	19	15	19	19	21

Contents

Preface 2

1 Free Groups 4

2 Subgroups of Free Groups 11

3 Automorphisms of Free Groups 18

4 Group Extensions 21

5 Central Extensions 32

6 Extensions of the Alternating Groups 46

7 Symplectic Groups 52

8 Unitary Groups 58

9 Orthogonal Groups 67

10 Sporadic Groups 71

11 Coxeter Groups 81

12 Free Products and Amalgamated Free Products 100

13 The Burnside Problem 106

14 Group Classes and Varieties 119

15 p-Groups 125

16 Decision Problems 132

Exercises 134

Index 140

Preface

Disclaimer. This English version is an automatic translation of the German original, produced by the AI models Claude Opus 5.5 and Claude Sonnet 5.5 (Anthropic). It may contain errors. In case of doubt, the German original on the author's homepage is authoritative.

In combinatorial group theory, groups are studied mainly by means of generators and relations. Naturally, one therefore has to deal mostly with infinite objects. One of the most important motivations is the *Burnside problem*: Is every finitely generated group of finite exponent finite? Since geometric arguments are occasionally used (for example for Coxeter groups), one alternatively speaks of *geometric group theory*.

These lecture notes originated from a $3 + 1$ lecture course in the summer semester 2022 at Leibniz University Hannover. The course follows my Group Theory course in the winter semester 2020/21 and presupposes corresponding knowledge (references to it are marked with GT). At some points there is overlap between the two courses. The reasons are:

- Recalling knowledge.
- Presenting alternative proof methods.
- Topics that are in the Group Theory lecture notes but were not covered there for lack of time (example: Schur extensions).

Chapters 3, 5, 8, and 11–15 were not covered (so more errors are to be expected there). To make computing with (realistic) examples practical, at many points we give commands for the free computer algebra system GAP. At this point, many thanks to Thomas Breuer for useful advice. I also thank Annika Bartelt, Luca Blaas, Adrian Homma, Scheima Obeidi, Claude Opus, and Tim Wittenberg for pointing out some errors. The German original is available on my homepage. This English translation was produced by the AI models Claude Opus 5.5 and Claude Sonnet 5.5.

Literature:

- T. Camps, V. Rebel, G. Rosenberger, *Einführung in die kombinatorische und die geometrische Gruppentheorie*, Heldermann Verlag, Lemgo, 2008 (in German)[1]
- M. Hall, *The Theory of Groups*, 4th printing, The Macmillan Company, New York, 1963[2]
- D. J. S. Robinson, *A Course in the Theory of Groups*, 2nd edition, Springer, New York, 1996[3]
- W. Magnus, A. Karrass, D. Solitar, *Combinatorial Group Theory*, 2nd edition, Dover, Mineola, 2004[4]
- R. C. Lyndon, P. E. Schupp, *Combinatorial Group Theory*, Springer, Berlin, 1977
- J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge University Press, Cambridge, 1994
- D. L. Johnson, *Presentations of Groups*, 2nd edition, Cambridge University Press, Cambridge, 1997
- H. S. M. Coxeter, W. O. J. Moser, *Generators and Relations for Discrete Groups*, 4th edition, Springer, Berlin, 1980
- The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.11.1*; 2021, (`https://www.gap-system.org`).

[1] Focus on the geometric aspect.
[2] Treats finite and infinite groups on an equal footing. Despite its age, it is easy to read.
[3] Somewhat more modern and more extensive than Hall.
[4] Unchanged reprint of the 1975 edition. Small print, a lot of running text, hard to read.

1 Free Groups

Remark 1.1. A basis B of a vector space V can be characterized by the following property: For every vector space W, every map $B \to W$ has exactly one linear extension $V \to W$. In this way one can define bases for arbitrary (nonabelian) groups. In contrast to vector spaces, however, most groups do not have bases.

Definition 1.2 (Universal property)**.** A group F is called *free* with respect to a subset $X \subseteq F$ if for every group G and every map $\sigma\colon X \to G$ there exists exactly one homomorphism $\widehat{\sigma}\colon F \to G$ with $\widehat{\sigma}(x) = \sigma(x)$ for all $x \in X$ (i. e., $\widehat{\sigma}$ is an extension of σ).

Example 1.3.

(i) The trivial group $F = 1$ is the only free group with respect to $X = \varnothing$, since every nontrivial group F has at least two endomorphisms.

(ii) The group $\mathbb{Z}$ is free with respect to $X = \{1\}$, since every map $\sigma\colon X \to G$ has the unique extension $\mathbb{Z} \to G$, $n \mapsto \sigma(1)^n$.

Definition 1.4. Let X be a set, which we call an *alphabet*. The elements of X are called *letters*. Let W be the set of all *words* of the form $w = x_1^{\epsilon_1} \dots x_n^{\epsilon_n}$ with $n \in \mathbb{N}_0$, $x_1, \dots, x_n \in X$ and $\epsilon_1, \dots, \epsilon_n \in \{\pm 1\}$. Here $|w| := n$ is called the *length* of w. For $n = 0$ we obtain the *empty* word $w = 1$. If $x_i \neq x_{i+1}$ or $\epsilon_i = \epsilon_{i+1}$ for $i = 1, \dots, n-1$, then w is called *reduced*. Clearly, every word w can be transformed into a reduced word by successively deleting parts of the form xx^{-1} or $x^{-1}x$. Two words $v, w \in W$ are called *equivalent*, if they can be transformed into the same reduced word. This is an equivalence relation on W. The set of equivalence classes $F_X := \{[w] : w \in W\}$ then forms a group with respect to concatenation, i. e.,

$$[w][v] := [wv] \qquad [w], [v] \in F_X.$$

The identity element is the equivalence class of the empty word [1]. The inverse of $[x_1^{\epsilon_1} \dots x_n^{\epsilon_n}]$ is $[x_n^{-\epsilon_n} \dots x_1^{-\epsilon_1}]$. By identifying $x \in X$ with $[x] \in F_X$, we may assume $X \subseteq F_X$.

Theorem 1.5.

(i) F_X is free with respect to X.

(ii) Every free group with respect to X is isomorphic to F_X.

(iii) We have $F_X \cong F_Y$ if and only if X and Y have the same cardinality.

Proof.

(i) Let G be a group and $\sigma\colon X \to G$. For $w = x_1^{\epsilon_1} \dots x_n^{\epsilon_n} \in W$ we define

$$\widehat{\sigma}(w) := \sigma(x_1)^{\epsilon_1} \dots \sigma(x_n)^{\epsilon_n} \in G.$$

For equivalent words $v, w \in W$ clearly $\widehat{\sigma}(v) = \widehat{\sigma}(w)$ holds. Hence $\widehat{\sigma}$ induces a well-defined map $F_X \to G$, which we also denote by $\widehat{\sigma}$. Since $\widehat{\sigma}(wv) = \widehat{\sigma}(w)\widehat{\sigma}(v)$ for $w, v \in W$, $\widehat{\sigma}$ is a homomorphism. Since $F_X = \langle X \rangle$, $\widehat{\sigma}$ is uniquely determined by σ.

(iii) Let $\sigma\colon X \to Y$ be a bijection. Then there exist homomorphisms $\alpha\colon F_X \to F_Y$ and $\beta\colon F_Y \to F_X$ with $\alpha_{|X} = \sigma$ and $\beta_{|Y} = \sigma^{-1}$. Hence $\alpha\beta\colon F_Y \to F_Y$ is an extension of id_Y. The universal property implies $\alpha\beta = \mathrm{id}_{F_Y}$. Similarly, one shows $\beta\alpha = \mathrm{id}_{F_X}$. Hence α is an isomorphism between F_X and F_Y.

Conversely, let an isomorphism $\alpha\colon F_X \to F_Y$ be given. We consider

$$N_X := \langle g^2, [g,h] : g,h \in F_X \rangle \trianglelefteq F_X$$

and $\overline{F_X} := F_X/N_X$. Since $\alpha(N_X) = \langle \alpha(g), [\alpha(g),\alpha(h)] : g,h \in F_X\rangle = N_Y$, we have $\overline{F_X} \cong \overline{F_Y}$. By construction, $\overline{F_X}$ is an abelian group with $\overline{g}^2 = 1$ for all $\overline{g} \in \overline{F_X}$. By the scalar multiplication $\lambda\overline{g} := \overline{g}^\lambda$ for $\lambda \in \mathbb{F}_2$, $\overline{F_X}$ becomes an $\mathbb{F}_2$-vector space (an infinite version of the elementary abelian groups from GT). The elements $\overline{x} := xN_X$ with $x \in X$ form a generating set of $\overline{F_X}$. Suppose there exist pairwise distinct $x_1,\dots,x_n \in X$ with $x_1\dots x_n \in N_X$. Let $\sigma\colon X \to \mathbb{F}_2$ with $\sigma(x_1) = 1$ and $\sigma(x_i) = 0$ for $i = 2,\dots,n$. Let $\widehat{\sigma}\colon F_X \to \mathbb{F}_2$ be the extension of σ. Since $\widehat{\sigma}(g^2) = 2\widehat{\sigma}(g) = 0$ and $\widehat{\sigma}([g,h]) = \widehat{\sigma}(g)+\widehat{\sigma}(h)-\widehat{\sigma}(g)-\widehat{\sigma}(h) = 0$, we have $x_1\dots x_n \in N_X \subseteq \mathrm{Ker}(\widehat{\sigma})$. This contradicts $\widehat{\sigma}(x_1\dots x_n) = \sigma(x_1) = 1$. Hence $x_1\dots x_n \notin N_X$. This shows that $\{\overline{x} : x \in X\}$ is a basis of $\overline{F_X}$. Thus $|X| = \dim \overline{F_X} = \dim \overline{F_Y} = |Y|$.[5]

(ii) Follows as in (iii) (only the universal property is used). □

Definition 1.6. If F is free with respect to X, then $\mathrm{rk}\, F := |X|$ is called the *rank* of F. By Theorem 1.5, up to isomorphism there is only one free group of rank $r \in \mathbb{N}$. We denote it by F_r.

Lemma 1.7 (cf. GT, Exercise 60)**.** *Every word $w \in W$ is equivalent to exactly one reduced word $\tilde{w} \in W$.*

Proof (Van der Waerden). We already know that w is equivalent to at least one reduced word. For uniqueness, let $R \subseteq W$ be the set of all reduced words. For $r = x_1^{\epsilon_1}\dots x_n^{\epsilon_n} \in R$ and $x \in X$ let

$${}^x r := \begin{cases} x x_1^{\epsilon_1}\dots x_n^{\epsilon_n} & \text{if } x \neq x_1^{-\epsilon_1}, \\ x_2^{\epsilon_2}\dots x_n^{\epsilon_n} & \text{if } x = x_1^{-\epsilon_1}. \end{cases}$$

Clearly, x induces a permutation $\sigma(x) \in \mathrm{Sym}(R)$ with inverse map $\sigma(x^{-1})$. By the universal property, σ extends to an action $F_X \to \mathrm{Sym}(R)$. For equivalent reduced words $v, w \in R$ we have

$$v = {}^{[v]}1 = {}^{[w]}1 = w. \qquad \square$$

Remark 1.8.

(i) For $r \geq 2$, F_r is nonabelian, since $xyx^{-1}y^{-1} \neq 1$ is reduced for $x \neq y$.

(ii) Let $w = x_1^{\epsilon_1}\dots x_n^{\epsilon_n} \in W$ be reduced with finite order k in F_X. By conjugating with $x_1^{-\epsilon}$ if necessary, we may assume $x_1^{\epsilon_1} \neq x_n^{-\epsilon_n}$. Then w^k is also reduced and $w = 1$ follows. Hence F_r is torsion-free.

Corollary 1.9. *If F is free with respect to $X \subseteq F$, then every element of F can be written uniquely in the form $x_1^{a_1}\dots x_n^{a_n}$, where $n \in \mathbb{N}_0$, $x_1,\dots,x_n \in X$ with $x_i \neq x_{i+1}$ for $i = 1,\dots,n-1$ and $a_1,\dots,a_n \in \mathbb{Z}\setminus\{0\}$. In particular, $F = \langle X\rangle$.*

[5] requires the axiom of choice if $|X| = \infty$

Proof. W. l. o. g. we may assume $F = F_X$. By Lemma 1.7, every element can be written uniquely in reduced form. By combining equal letters, we obtain a representation of the desired form. □

Remark 1.10. By Cayley, every group is isomorphic to a subgroup of a symmetric group. Free groups have a dual property.

Theorem 1.11. *Every group G is isomorphic to a quotient group of a free group F. If G can be generated by n elements, then we can choose* $\operatorname{rk} F = n$.

Proof. Let X be a generating set of G (if necessary $X = G$). Then the inclusion $X \to G$ can be extended to an epimorphism $F_X \to G$. The claim follows from the homomorphism theorem. □

Example 1.12. Let

$$a := \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \qquad b := \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

We show that $G := \langle a, b\rangle \leq \mathrm{GL}(2, \mathbb{Q})$ is free with respect to $\{a, b\}$. To this end, let $X = \{x, y\}$ and let $\varphi\colon F_X \to G$ be the epimorphism with $\varphi(x) = a$ and $\varphi(y) = b$. Suppose there exists a nontrivial reduced word $w := z_1^{k_1} \dots z_n^{k_n} \in \operatorname{Ker}(\varphi)$ with $z_1, \dots, z_n \in X$ and $k_1, \dots, k_n \in \mathbb{Z} \setminus \{0\}$. After conjugation, we may assume $z_n = x$ and $k_n > 0$. Then $1 = \varphi(w) = \dots a^{k_{n-2}} b^{k_{n-1}} a^{k_n}$. For $k \in \mathbb{Z}$ we have

$$a^k := \begin{pmatrix} 1 & 2k \\ 0 & 1 \end{pmatrix}, \qquad b^k := \begin{pmatrix} 1 & 0 \\ 2k & 1 \end{pmatrix}$$

(induction on k). Let

$$V_> := \big\{(s,t) \in \mathbb{Q}^2 : |s| > |t|\big\}, \qquad V_< := \big\{(s,t) \in \mathbb{Q}^2 : |s| < |t|\big\}.$$

For $v := (1,1) \in \mathbb{Q}^2$ we have $a^{k_n} v = (2k_n + 1, 1) \in V_>$ since $k_n > 0$. For $v = (s,t) \in V_>$ we have $b^{k_{n-1}} v = (s, 2k_{n-1}s + t) \in V_<$, since $|2k_{n-1}s + t| \geq 2|k_{n-1}||s| - |t| > |s|$ (triangle inequality) because $k_{n-1} \neq 0$. Similarly, $a^{k_{n-2}} v \in V_>$ for $v \in V_<$ etc. [6] This yields the contradiction

$$(1,1) = \varphi(w)(1,1) = \dots a^{k_{n-2}} b^{k_{n-1}} a^{k_n} (1,1) \in V_> \cup V_<.$$

Hence φ is injective and $G \cong F_X \cong F_2$.

Remark 1.13.

(i) Let X be a generating set for G and $\sigma\colon F_X \to G$ with $\sigma(x) = x$ as in Theorem 1.11. The elements in $\operatorname{Ker}(\sigma)$ are called *relators* for G with respect to X. Thus for $x_1^{\epsilon_1} \dots x_n^{\epsilon_n} \in \operatorname{Ker}(\sigma)$ we have $x_1^{\epsilon_1} \dots x_n^{\epsilon_n} = 1$ in G. An equation of this form is called a *relation* for G with respect to X.

(ii) Conversely, let $R \subseteq F_X$. Let $N := \langle R\rangle^{F_X} := \langle gRg^{-1} : g \in F_X\rangle \trianglelefteq F_X$ be the normal closure of R in F_X. We set

$$G := \langle X \mid R\rangle = \langle X \mid \{r = 1 : r \in R\}\rangle := F_X/N.$$

One often identifies letters $x \in X$ with their cosets $xN \in G$ (in general not injective!). If $|X| + |R| < \infty$, then G is called *finitely presented*. In this way every group can be described by generators and relations (this corresponds to the statement that every vector space is the solution set of a system of linear equations). In general, however, it is difficult to read off the properties of G from X and R.

[6] The argument is called the *ping-pong lemma*.

Example 1.14.

(i) $\langle X \mid \varnothing \rangle \cong F_X$.

(ii) $\langle x \mid x^n \rangle = \langle x \mid x^n = 1 \rangle \cong \mathbb{Z}/n\mathbb{Z} \cong C_n$.

(iii) Every finite group G is finitely presented: Let $X := \{x_g : g \in G\}$, $R := \{x_g x_h x_{gh}^{-1} : g, h \in G\}$ and $N := \langle R \rangle^{F_X}$. Then there exists an epimorphism $\varphi\colon F_X \to G$ with $\varphi(x_g) = g$ and $R \subseteq \mathrm{Ker}(\varphi)$. Since $\mathrm{Ker}(\varphi) \trianglelefteq F_X$, we also have $N \subseteq \mathrm{Ker}(\varphi)$. Conversely, let $w := x_{g_1}^{\epsilon_1} \dots x_{g_n}^{\epsilon_n} \in \mathrm{Ker}(\varphi)$. Since $x_1 = x_1 x_1 x_1^{-1} \in R$ and $x_g x_{g^{-1}} x_1^{-1} \in R$, we have $x_g x_{g^{-1}} \in \langle R \rangle$ and $x_g^{-1} \equiv x_{g^{-1}} \pmod N$. It follows that

$$w \equiv x_{g_1^{\epsilon_1}} \dots x_{g_n^{\epsilon_n}} \equiv x_{g_1^{\epsilon_1} g_2^{\epsilon_2}} x_{g_3^{\epsilon_3}} \dots \equiv x_{g_1^{\epsilon_1} \dots g_n^{\epsilon_n}} \equiv x_1 \equiv 1 \pmod N.$$

Hence $\mathrm{Ker}(\varphi) \subseteq N$ and $G = \langle X \mid R \rangle$.

Theorem 1.15 (von Dyck)**.** *Let $G = \langle x_i : i \in I \rangle$ and $H = \langle y_i : i \in I \rangle$ be groups such that for every relation $x_{i_1}^{\epsilon_1} \dots x_{i_n}^{\epsilon_n} = 1$ in G the relation $y_{i_1}^{\epsilon_1} \dots y_{i_n}^{\epsilon_n} = 1$ also holds in H. Then there exists an epimorphism $G \to H$ with $f(x_i) = y_i$ for $i \in I$.*

Proof. By the universal property, there exist epimorphisms $f_G\colon F_I \to G$ and $f_H\colon F_I \to H$ with $f_G(i) = x_i$ and $f_H(i) = y_i$ for $i \in I$. By assumption, $\mathrm{Ker}(f_G) \leq \mathrm{Ker}(f_H)$. Hence

$$G \cong F_I/\mathrm{Ker}(f_G) \to (F_I/\mathrm{Ker}(f_G))/(\mathrm{Ker}(f_H)/\mathrm{Ker}(f_G)) \cong F_I/\mathrm{Ker}(f_H) \cong H$$

is the desired epimorphism. □

Remark 1.16. One can use von Dyck's theorem to approximate opaque group presentations by known groups.

Example 1.17.

(i) Let $G := \langle x_1, \dots, x_n \mid [x_i, x_j] = 1 \ \forall i, j \rangle$. Clearly, G is abelian and every element in G has the form $x_1^{a_1} \dots x_n^{a_n}$ with $a_1, \dots, a_n \in \mathbb{Z}$. Now let $H := \langle y_1 \rangle \oplus \dots \oplus \langle y_n \rangle \cong C_\infty^n$. By Theorem 1.15, there exists an epimorphism $f\colon G \to H$ with $f(x_i) = y_i$ for $i = 1, \dots, n$. Clearly, f is also injective and $G \cong H \cong C_\infty^n$. This also shows $F_n/F_n' \cong C_\infty^n$.

(ii) Let $G = \langle x, y \rangle$ with $x \neq y$ and $|\langle x \rangle| = |\langle y \rangle| = 2$. Then G consists of the elements of the form $xyxy \dots$ and $yxyx \dots$. If G is finite, then $n := |\langle xy \rangle| \in \mathbb{N}$ and every element has the form $x^i (xy)^j$ with $0 \leq i \leq 1$ and $0 \leq j \leq n-1$. Then we obtain

$$G \cong D_{2n} := \langle \sigma, \tau \mid \sigma^n = \tau^2 = 1, \ \tau\sigma\tau = \sigma^{-1} \rangle,$$

where $D_4 = C_2^2$.

Remark 1.18 (GAP)**.**

```
F:=FreeGroup("x","y");; #the double semicolon suppresses the output
AssignGeneratorVariables(F);;
G:=F/[x^2,y^3,(x*y)^5]; #finitely presented group
Size(G);
H:=Image(IsomorphismPermGroup(G)); #isomorphic permutation group (more efficient)
StructureDescription(H);

#The reverse direction:
G:=Image(IsomorphismFpGroup(H));; #isomorphic finitely presented group
RelatorsOfFpGroup(G); #new relations

#Computing in the free group:
x:=(1,2,3,4,5,6,7,8,9,10,11);; y:=(3,7,11,8)(4,10,5,6);; #permutations of S_11
G:=Group(x,y);;
epi:=GroupHomomorphismByImages(F,G,GeneratorsOfGroup(F),[x,y]);
z:=Random(G);
PreImagesRepresentative(epi,z); #representation of z as a word in x,y
```

Theorem 1.19 (Neumann). *Let G be finitely presented and X an arbitrary generating set of G. Then there exists a finite presentation $G = \langle X_0 \mid R\rangle$ with $X_0 \subseteq X$.*

Proof. Let $G = \langle y_1, \dots, y_n \mid s_1, \dots, s_m\rangle$ be a finite presentation. Every y_i can be expressed by finitely many $x \in X$, say $y_i = w_i(x)$. Hence there exists a finite subset $X_0 = \{x_1, \dots, x_k\} \subseteq X$ with $G = \langle X_0\rangle$. Conversely, the x_i can be expressed by the y_j, say $x_i = v_i(y)$. We obtain the following relations in X_0:

$$s_i(w_1(x), \dots, w_n(x)) = 1, \qquad x_i = v_i(w_1(x), \dots, w_n(x)).$$

Let R be the set of these relators and $H := \langle X_0 \mid R\rangle$. By von Dyck's theorem, there exists an epimorphism $\varphi\colon H \to G$ with $\varphi(x_i) = x_i$ for $i = 1, \dots, k$. Now we can define $y_i := w_i(x)$ in H. Since $x_i = v_i(y_1, \dots, y_n)$, we then have $H = \langle y_1, \dots, y_n\rangle$. Since the relations in s_i also hold in H, there exists an epimorphism $\psi\colon G \to H$ with $\psi(y_i) = y_i$ for $i = 1, \dots, n$. From

$$\psi(\varphi(x_i)) = \psi(x_i) = \psi(v_i(y)) = v_i(y) = x_i,$$
$$\varphi(\psi(y_i)) = \varphi(y_i) = \varphi(w_i(x)) = w_i(x) = y_i$$

it follows that φ and ψ are mutually inverse isomorphisms. □

Theorem 1.20 (Hall). *Let $N \trianglelefteq G$. If N and G/N are finitely presented, then so is G.*

Proof. Let $N = \langle x_1, \dots, x_n \mid r_1, \dots, r_m\rangle$ and $G/N = \langle y_1N, \dots, y_kN \mid s_1, \dots, s_l\rangle$. Clearly, $G = \langle x_1, \dots, x_n, y_1, \dots, y_k\rangle$. Since $s_i(y) \in N$, there are relations $s_i(y) = t_i(x)$. The normality can be expressed by relations $y_ix_jy_i^{-1} = u_{ij}(x)$ and $y_i^{-1}x_jy_i = v_{ij}(x)$. We define

$$H := \langle x_1, \dots, x_n, y_1, \dots, y_k \mid \forall i, j : r_i, s_i(y)t_i(x)^{-1}, u_{ij}(x)y_ix_j^{-1}y_i^{-1}, v_{ij}(x)y_i^{-1}x_j^{-1}y_i \ \rangle.$$

Then there exists an epimorphism $\varphi\colon H \to G$ with $\varphi(x_i) = x_i$ and $\varphi(y_j) = y_j$. Let $\tilde{N} := \langle x_1, \dots, x_n\rangle \leq H$. The relations u_{ij} and v_{ij} show $\tilde{N} \trianglelefteq H$. For $h \in \tilde{N} \cap \mathrm{Ker}(\varphi)$ we have $h \in \langle r_1, \dots, r_m\rangle^{F_X}$. This shows $h = 1$ and $\tilde{N} \cap \mathrm{Ker}(\varphi) = 1$. Clearly, φ induces an epimorphism $\overline{\varphi}\colon H/\tilde{N} \to G/N$ with $\overline{\varphi}(y_i\tilde{N}) = y_iN$. The relations r_i, t_i, u_{ij} and v_{ij} become trivial in $H/\tilde{N}$. Hence $H/\tilde{N}$ and G/N satisfy the same relations ($s_i(y) = 1$) and $\overline{\varphi}$ is an isomorphism. For $h \in \mathrm{Ker}(\varphi)$ we have $h\tilde{N} \in \mathrm{Ker}(\overline{\varphi}) = 1$, hence $h \in \tilde{N} \cap \mathrm{Ker}(\varphi) = 1$. Hence φ is an isomorphism. □

Theorem 1.21. *Let $G = \langle X \mid R\rangle$ with $X = \{x_1, \dots, x_n\}$ and $R = \{r_1, \dots, r_k\}$. For $i = 1, \dots, k$ let $r_i \equiv x_1^{a_{i1}} \dots x_n^{a_{in}} \pmod{F_X'}$. Let $d_1 \mid \dots \mid d_l$ be the elementary divisors of $A = (a_{ij}) \in \mathbb{Z}^{k\times n}$, where $l := \min\{n, k\}$. Then $G/G' \cong \mathbb{Z}/d_1\mathbb{Z} \times \dots \times \mathbb{Z}/d_l\mathbb{Z} \times \mathbb{Z}^{n-l}$. In particular, $|G| = |G/G'| = \infty$ if $k < n$.*

Proof. Let $F := F_X$. By Example 1.17, F/F' is a free abelian group of rank n. For $N := \langle rF' : r \in R\rangle \leq F/F'$ we now have $G/G' \cong (F/F')/N \cong \mathbb{Z}/d_1\mathbb{Z} \times \dots \times \mathbb{Z}/d_l\mathbb{Z} \times \mathbb{Z}^{n-l}$ (see the Algebra lecture notes). □

Remark 1.22.

(i) By Theorem 1.21, one can algorithmically decide whether a finitely presented group is perfect (nevertheless, one does not know whether the group is trivial).

(ii) Only few *finite* groups $G = \langle X \mid R\rangle$ with $|X| = |R|$ are known (for example Q_{2^n} by Exercise 2).[7] One even conjectures that every such group can be generated by three elements. This was proved for p-groups (cf. Theorem 5.25 and Exercise 22).

Example 1.23. Let $G = \langle x, y \mid x^2 = y^3 = (xy)^7 = 1\rangle$. Then we obtain the matrix

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \\ 7 & 7 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ 0 & 3 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 \\ 0 & -2 \\ 0 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Hence $G = G'$ is perfect. One can verify that the matrices

$$x = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

satisfy the relations in $\mathrm{GL}(3, 2)$. Hence $G \neq 1$. We will show in Theorem 11.50 that G is infinite.

Lemma 1.24. *Let N be the normal closure of $Y \subseteq X$ in F_X. Then F_X/N is free with respect to $X \setminus Y$.*

Proof. W. l. o. g. we may assume $\varnothing \neq Y \subset X$. Every element of N is a product of elements of the form gwg^{-1} with $w \in \langle Y\rangle$ and $g \in F_X$. The sum of the exponents of a letter $x \in X \setminus Y$ in gwg^{-1} is 0, since x can only occur in g and g^{-1}. This shows that the map $\sigma\colon X \setminus Y \to F/N$, $x \mapsto xN$ is injective. We show that F_X/N is free with respect to $\sigma(X \setminus Y)$. Let G be an arbitrary group and $\alpha\colon X \setminus Y \to G$ a function. Setting

$$\overline{\alpha}(x) := \begin{cases} 1 & \text{if } x \in Y \\ \alpha(x) & \text{if } x \notin Y \end{cases}$$

for $x \in X$, we obtain an extension $\overline{\alpha}\colon X \to G$ of α. Since F is free with respect to X, there exists a homomorphism $\overline{\beta}\colon F \to G$ with $\overline{\beta}(x) = \overline{\alpha}(x)$ for all $x \in X$. Clearly, $N \leq \mathrm{Ker}(\overline{\beta})$. Hence there exists a homomorphism $\beta\colon F/N \to (F/N)/(\mathrm{Ker}(\overline{\beta})/N) \to F/\mathrm{Ker}(\overline{\beta}) \to G$ with $\beta(\sigma(x)) = \alpha(x)$ for all $x \in X \setminus Y$. Now let $\beta'\colon F/N \to G$ be another homomorphism with $\beta'(\sigma(x)) = \alpha(x)$ for all $x \in X \setminus Y$. From

$$F/N = \langle xN : x \in X\rangle = \langle xN : x \in X \setminus Y\rangle = \langle \sigma(x) : x \in X \setminus Y\rangle$$

we immediately obtain $\beta = \beta'$. □

[7] One speaks of a *balanced* presentation.

Remark 1.25.

(i) The *Cayley graph* $\Omega(G, X)$ of a finitely generated group $G = \langle X \mid R \rangle$ is a graph with vertex set G such that $g, h \in G$ form an edge if and only if $g^{-1}h \in X \cup X^{-1}$ (by $g^{-1}h \in X$ one obtains a directed graph). Clearly, $\Omega(G, X)$ depends on X. To avoid loops, we assume $1 \notin X$. In any case, $\Omega(G, X)$ is connected and regular (i. e., all vertices have the same number of edges).

(ii) A cycle in $\Omega(G, X)$ corresponds to a reduced word $x_1^{\epsilon_1} \dots x_n^{\epsilon_n} = 1$. Hence $\Omega(G, X)$ is a tree (i. e., acyclic) if and only if G is free with respect to X.

(iii) The group G permutes the vertices and edges of $\Omega(G, X)$ by left multiplication. Hence G is a subgroup of $\mathrm{Aut}(\Omega(G, X))$. On this basis one can show that every finite group is the automorphism group of a graph (FRUCHT's theorem[8]).

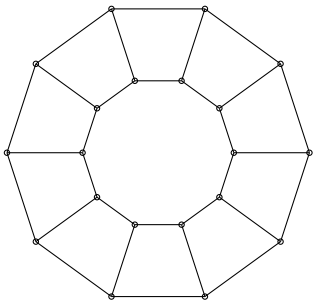
(iv) By the Euler-Hierholzer theorem, $\Omega(G, X)$ is Eulerian if and only if $|X \cup X^{-1}|$ is even. This means there is a closed path that visits every edge of $\Omega(G, X)$ exactly once. The open Lovász conjecture states that $\Omega(G, X)$ is Hamiltonian for $2 < |G| < \infty$. This means there is a closed path that visits every vertex of $\Omega(G, X)$ exactly once.

(v) If G is finite, one can study the eigenvalues of the adjacency matrix of $\Omega(G, X)$. One can show that they are closely connected to the values of the complex irreducible characters of G.

Example 1.26.

(i) The Cayley graph of $G = \langle x \mid x^n \rangle$ with respect to $X = \{x\}$ is an n-gon. Choosing $X = G$, we obtain the complete graph with n vertices.

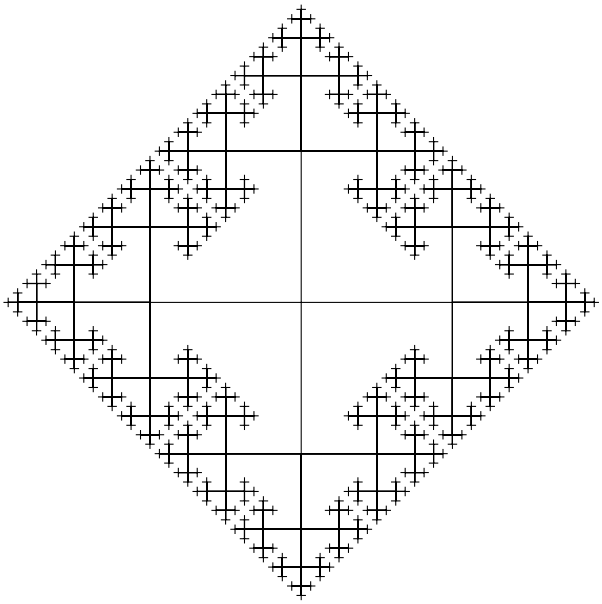
(ii) Let $G = D_{2n}$ be generated by two reflections $X = \{x, y\}$. Then $\Omega(G, X)$ is a $2n$-gon, just as for C_{2n}. Hence the Cayley graph does not determine whether G is abelian. If instead we choose $X = \{x, z\}$ with a rotation z by $360°/n$, then we obtain (with $n = 10$):



The two circles correspond to the cosets $\langle z \rangle$ and $x\langle z \rangle$.

[8] See the Algebra lecture notes

(iii) The Cayley graph of F_2 can be drawn (approximately) as a fractal:



Interactive examples can be found at `https://juliapoo.github.io/Cayley-Graph-Plotting/`.

(iv) Let $G = \langle X \rangle$ be the group of the $3\times3\times3$ Rubik's Cube and X the set of the 90° and 180° rotations of the six faces. Then *God's Number* (20) is the diameter of $\Omega(G, X)$, that is, the maximal length of a shortest path between two vertices (see the Christmas lecture of the Group Theory course).

Remark 1.27 (GAP)**.**

```
LoadPackage("grape",false);
G:=AlternatingGroup(5);;
Y:=[(1,2,3),(3,4,5)];; #X is write-protected
C:=CayleyGraph(G,Y); #replaces Y by Y∪Y^-1
Diameter(C);

LoadPackage("hap",false); #loads further packages
G:=DihedralGroup(22);;
Y:=GeneratorsOfGroup(G);
CayleyGraphOfGroupDisplay(G,Y,"chromium"); #output in the browser chromium, requires
    GraphViz
```

2 Subgroups of Free Groups

Remark 2.1. We already know that every group is a subgroup of a symmetric group (and a quotient group of a free group). The quotient groups of a finite symmetric group are themselves symmetric (since A_n is simple for $n \geq 5$). We show, dually, that the subgroups of a free group are themselves free. We start with a variation of the universal property.

Lemma 2.2 (Steinberg)**.** *A group F is free with respect to $X \subseteq F$ if and only if for every nonempty set Ω and every map $\sigma\colon X \to \mathrm{Sym}(\Omega)$ there exists exactly one action $F \to \mathrm{Sym}(\Omega)$ with ${}^x\omega = \sigma(x)(\omega)$ for all $x \in X$ and $\omega \in \Omega$.*

Proof. If F is free with respect to X, then σ can be extended to an action with the stated property by the universal property. For the converse, we first show $F = \langle X \rangle$. The map $\sigma\colon X \to \mathrm{Sym}(\langle X \rangle)$ given by left multiplication (i. e., $\sigma(x)(y) = xy$) can be extended to an action $\rho\colon F \to \mathrm{Sym}(\langle X \rangle)$. But we can also consider $\sigma\colon X \to \mathrm{Sym}(F)$, and by uniqueness we obtain exactly the same extension $\rho\colon F \to \mathrm{Sym}(F)$. Since the action by left multiplication is transitive, we have $\langle X \rangle = F$.

Now let a group G and a map $\sigma\colon X \to G$ be given. Let $\tau\colon G \to \mathrm{Sym}(G)$ be the monomorphism from Cayley's theorem. By hypothesis, there exists exactly one action $\rho\colon F \to \mathrm{Sym}(G)$ with $\rho(x) = \tau(\sigma(x))$ for all $x \in X$. Here we have $\rho(F) = \rho(\langle X \rangle) \leq \tau(G)$. Clearly, $\tau^{-1}\rho\colon F \to G$ is a homomorphism extending σ. If $\gamma\colon F \to G$ is also an extension of σ, then $\tau\gamma = \rho$ by hypothesis. This shows $\gamma = \tau^{-1}\rho$. □

Definition 2.3. Let F be free with respect to $X \subseteq F$ and $G \leq F$. A *Schreier transversal* for G is a set of representatives S for F/G with the following property: If $x^\epsilon s \in S$ is reduced with $x \in X$ and $\epsilon = \pm 1$, then also $s \in S$. In particular, $1 \in S$.

Lemma 2.4. *Every subgroup of a free group has a Schreier transversal.*

Proof. Let F be free with respect to $X \subseteq F$ and $G \leq F$. Let the *length* of a coset aG be the minimal length of a reduced word in aG. We construct a Schreier transversal S of G by induction on the length of the cosets. Clearly, $1G$ is the only coset of length 0. So let $1 \in S$. Now let aG be of length $l \geq 1$. Here, let a have length l. Let $a = x^\epsilon b$ be reduced with $x \in X$ and $\epsilon = \pm 1$. By induction, there exists $s \in S$ with $sG = bG$. We choose $x^\epsilon s \in S$ as the representative of aG. Clearly, this yields a Schreier transversal. □

Theorem 2.5 (Nielsen-Schreier)**.** *Subgroups of free groups are free.*

Proof. Let F be free with respect to $X \subseteq F$ and $G \leq F$. Let T be a Schreier transversal of G. For $a \in F$ let $\overline{a} \in T$ with $aG = \overline{a}G$. We use Lemma 2.2 to show that G is free with respect to the following set:

$$Y := \{(\overline{xt})^{-1}xt : (x,t) \in X \times T,\ xt \neq \overline{xt}\} \subseteq G$$

For this, let $\sigma\colon Y \to \mathrm{Sym}(\Omega)$ be an arbitrary map. We extend σ by setting $\sigma(1) := \mathrm{id}_\Omega$. For $x \in X$, the maps $T \to T$, $t \mapsto \overline{xt}$ (with inverse map $t \mapsto \overline{x^{-1}t}$) and $\sigma((\overline{xt})^{-1}xt)$ are bijections. The map $f\colon X \to \mathrm{Sym}(\Omega \times T)$ with

$$f(x)(\omega, t) = (\sigma((\overline{xt})^{-1}xt)(\omega), \overline{xt})$$

extends to an action $\widehat{f}\colon F \to \mathrm{Sym}(\Omega \times T)$. We show that $\widehat{f}(g)(\omega, 1) = (\widehat{\sigma}(g)(\omega), 1)$ defines the unique homomorphic extension $\widehat{\sigma}\colon G \to \mathrm{Sym}(\Omega)$ of σ. For well-definedness we must verify $\widehat{f}(g)(\Omega \times \{1\}) = \Omega \times \{1\}$ for all $g \in G$. By construction, $f(x)$ maps the set $\Omega \times \{t\}$ bijectively onto $\Omega \times \{\overline{xt}\}$. Since $\overline{x\overline{x^{-1}t}} = t$, it follows that $\widehat{f}(x^{-1}) = f(x)^{-1}$ maps the set $\Omega \times \{t\}$ bijectively onto $\Omega \times \{\overline{x^{-1}t}\}$. For $a = x^\epsilon b \in F$ reduced with $x \in X$ and $\epsilon = \pm 1$ we have $\overline{x^\epsilon \overline{b}} = \overline{x^\epsilon b} = \overline{a}$. By induction on the length of a it follows that

$$\widehat{f}(a)(\Omega \times \{1\}) = \widehat{f}(x^\epsilon)(\Omega \times \{\overline{b}\}) = \Omega \times \{\overline{a}\}.$$

For $g \in G$ we have $\overline{g} = 1$ and therefore $\widehat{f}(g)(\Omega \times \{1\}) = \Omega \times \{1\}$ as desired. Since $\widehat{f}$ is a homomorphism, $\widehat{\sigma}$ is a homomorphism as well.

We now show ${}^t(\omega,1) = (\omega,t)$ for all $t \in T$. This is clear for $t = 1$. So let $t = x^\epsilon s$ be reduced with $x \in X$, $\epsilon = \pm 1$. Since T is a Schreier transversal, we have $s \in T$. By induction on the length of t we may assume ${}^s(\omega,1) = (\omega,s)$. In the case $\epsilon = 1$ we obtain

$$ {}^t(\omega,1) = {}^x(\omega,s) = (\sigma((\overline{xs})^{-1}xs)(\omega),t) = (\sigma(1)(\omega),t) = (\omega,t). $$

Now let $\epsilon = -1$. Since $xt = s \in T$, we have ${}^x(\omega,t) = (\sigma((\overline{xt})^{-1}xt)(\omega),\overline{xt}) = (\omega,s)$ and

$$ {}^t(\omega,1) = {}^{x^{-1}s}(\omega,1) = {}^{x^{-1}}(\omega,s) = (\omega,t). $$

Now let $y := (\overline{xt})^{-1}xt \in Y$ be arbitrary. Then

$$ {}^{xt}(\omega,1) = {}^x(\omega,t) = (\sigma(y)(\omega),\overline{xt}) = {}^{\overline{xt}}(\sigma(y)(\omega),1) $$

and $\widehat{\sigma}(y) = \sigma(y)$. Hence $\widehat{\sigma}$ is an extension of σ. Let $\tau\colon G \to \mathrm{Sym}(\Omega)$ be an extension as well. Then $\rho\colon F \to \mathrm{Sym}(\Omega \times T)$ with $\rho(a)(\omega,t) := (\tau((\overline{at})^{-1}at)(\omega),\overline{at})$ is an action, because for $a,b \in F$ and $t \in T$ we have

$$ \begin{aligned} (\rho(a)\rho(b))(\omega,t) &= \rho(a)(\tau((\overline{bt})^{-1}bt)(\omega),\overline{bt}) = (\tau((\overline{a\overline{bt}})^{-1}a\overline{bt})(\tau((\overline{bt})^{-1}bt)(\omega)),\overline{a\overline{bt}}) \\ &= (\tau((\overline{abt})^{-1}abt)(\omega),\overline{abt}) = \rho(ab)(\omega,t). \end{aligned} $$

For $x \in X$ we have

$$ \rho(x)(\omega,t) = (\tau((\overline{xt})^{-1}xt)(\omega),\overline{xt}) = (\sigma((\overline{xt})^{-1}xt)(\omega),\overline{xt}) = {}^x(\omega,t). $$

Since $F = \langle X \rangle$, ρ therefore coincides with the action defined above. In particular, $\tau = \widehat{\sigma}$. □

Remark 2.6. For vector spaces (or free abelian groups) $U \leq V$ we have $\dim U \leq \dim V$, as is well known. For free groups this is completely false.

Theorem 2.7 (Schreier's formula)**.** *If F is free and $G \leq F$ with $|F:G| < \infty$, then*

$$ \boxed{\mathrm{rk}(G) = |F:G|(\mathrm{rk}(F)-1)+1.} $$

Proof. It suffices to determine the cardinality of the set Y in Theorem 2.5. Let $s,t \in T$ and $x,y \in X$ with

$$ (\overline{xt})^{-1}xt = (\overline{ys})^{-1}ys \neq 1. $$

Then $xt, ys \notin T$. Since T is a Schreier transversal, xt and ys must be reduced. Suppose that $\overline{xt} = xt'$ is reduced. Then $t' \in T$ and

$$ t'G = x^{-1}xt'G = x^{-1}\overline{xt}G = x^{-1}xtG = tG, $$

so $t = t'$. This contradicts the choice of x and t. Hence $\overline{xt}$ does not begin with x. Consequently, x cannot cancel in $(\overline{xt})^{-1}xt$. Similarly, y does not cancel in $(\overline{ys})^{-1}ys$. Since xt and ys are reduced, xt must occur at the end of ys (or ys at the end of xt). If xt were actually shorter than ys, then xt would already occur in s, and we would obtain the contradiction $xt \in T$. Hence $xt = ys$ and it follows that $(x,t) = (y,s)$. Thus the listed elements of Y are pairwise distinct.

Now let $t \in T \setminus \{1\}$. Let $t = x^\epsilon s$ in reduced form with $x \in X$ and $\epsilon = \pm 1$. Since T is a Schreier transversal, we have $s \in T$. In the case $\epsilon = 1$ we have $xs \in T$, and otherwise $xt \in T$. Hence every $t \neq 1$ determines exactly one pair $(x',t') \in X \times T$ with $x't' \in T$. Conversely, every such pair arises in this way. This yields

$$ |Y| = |X \times T| - |T \setminus \{1\}| = |T|(|X|-1)+1 = |F:G|(\mathrm{rk}(F)-1)+1. $$ □

Example 2.8. Let $F_2 = \langle x, y\rangle$ and let N be the normal closure of $\{x^2, y^2, (xy)^2\}$ in F_2. By Example 1.17, F_2/N is the Klein four-group. By Schreier's formula, $\operatorname{rk} N = 4(2-1)+1 = 5$. We choose the Schreier transversal $\{1, x, y, xy\}$ of N in F_2 and compute

$$(\overline{xx})^{-1}xx = x^2,\ (\overline{xxy})^{-1}xxy = y^{-1}x^2y,\ (\overline{yy})^{-1}yy = y^2,$$
$$(\overline{yx})^{-1}yx = y^{-1}x^{-1}yx,\ (\overline{yxy})^{-1}yxy = x^{-1}yxy$$

Hence N is free with respect to $\{x^2, y^2, y^{-1}x^2y, y^{-1}x^{-1}yx, x^{-1}yxy\}$.

Corollary 2.9. *The group F_2 has (free) subgroups of every finite or countable rank. For every countable group G there exist $N \trianglelefteq H \leq F_2$ with $G \cong H/N$ (cf. Theorem 12.25).*

Proof. Let $F_2 = \langle x, y\rangle$. Then $G := \langle x\rangle \leq F_2$ with $\operatorname{rk} G = 1$ (Schreier's formula does not apply here because $|F : G| = \infty$). For $n \in \mathbb{N}$ there exists $N \trianglelefteq F$ with $F/N \cong C_n$ by Theorem 1.11. Schreier's formula shows $\operatorname{rk} N = n+1$. Finally, let $N := F' = [F, F]$. By Example 1.17, $F/N \cong \mathbb{Z}^2$. Hence the elements $x^a y^b$ with $a, b \in \mathbb{Z}$ form a Schreier transversal of N in F. The set Y constructed in the proof of Theorem 2.5 comprises the pairwise distinct reduced words $(\overline{yxy^b})^{-1}yxy^b = y^{-b-1}x^{-1}yxy^b$ with $b \in \mathbb{Z}$. Hence $\operatorname{rk} N = \infty$. The second statement follows from Theorem 1.11. □

Remark 2.10.

(i) The proof of Corollary 2.9 shows that not every subgroup of a finitely generated group needs to be finitely generated ($\operatorname{rk}(F_2') = \infty$).

(ii)
```
F:=FreeGroup("x","y");;
AssignGeneratorVariables(F);;
H:=Subgroup(F,[x^2,y^3,(x*y)^5]);;
IsFreeGroup(H);
Rank(H); #= 3
FreeGeneratorsOfGroup(H);
Index(F,H); #= ∞
N:=NormalClosure(F,H);;
Index(F,N); #= 60
GeneratorsOfGroup(N);; #intermediate computation
Rank(N);
```

Theorem 2.11. *If G is finitely generated and $H \leq G$ with $|G : H| < \infty$, then H is also finitely generated.*

Proof. Let $X = X^{-1}$ be a finite generating set of G and R a set of representatives for G/H with $1 \in R$. For $x \in X$ and $r \in R$ there exist $\alpha(x, r) \in H$ and $\gamma(x, r) \in R$ with $xr = \gamma(x, r)\alpha(x, r)$. Every element in H has the form $h = x_1 \dots x_n$ with $x_1, \dots, x_n \in X$. Here we have

$$h = x_1 \dots x_n 1 = x_1 \dots x_{n-1}\gamma(x_n, 1)\alpha(x_n, 1) = x_1 \dots x_{n-2}\gamma(x_{n-1}, \gamma(x_n, 1))\alpha(x_{n-1}, \gamma(x_n, 1))\alpha(x_n, 1)$$
$$= \dots = \gamma(x_1, \dots)\alpha(x_1, \dots) \dots \alpha(x_n, 1).$$

Since $h \in H$, we have $\gamma(x_1, \dots) = 1$. It follows that $H = \langle \alpha(x, r) : x \in X, r \in R\rangle$. □

Remark 2.12. The proof of Theorem 2.11 shows that H can be generated by $|G : H||X|$ elements. The next theorem gives an optimal estimate.

Theorem 2.13 (Reidemeister-Schreier). *Let $G = \langle X \mid R\rangle$ be a group and $H \leq G$. Then a presentation $H = \langle Y \mid S\rangle$ can be derived from X and R. In the case $|G : H| < \infty$ we have $|Y| \leq |G : H|(|X| - 1) + 1$ and $|S| \leq |G : H||R|$.*

Proof. Let $F := F_X$ and $N := \langle R\rangle^F \trianglelefteq F$. Let $\varphi\colon F \to G$ be the epimorphism with $\mathrm{Ker}(\varphi) = N$. Let $\tilde{H} = \varphi^{-1}(H) \leq F$. Then $|F : \tilde{H}| = |F/N : \tilde{H}/N| = |G : H|$. Let T be a Schreier transversal of $\tilde{H}$ in F. As in the proof of Theorem 2.5, $\tilde{H}$ is free with respect to

$$Y := \{(\overline{xt})^{-1}xt : (x,t) \in X \times T,\ xt \neq \overline{xt}\} \subseteq \tilde{H}.$$

In particular, $\langle\varphi(Y)\rangle = \varphi(\tilde{H}) = H$. For $(x,t) \in X \times T$ we have

$$(\overline{x^{-1}t})^{-1}x^{-1}t = (t^{-1}x\overline{x^{-1}t})^{-1} = \left((\overline{xs})^{-1}xs\right)^{-1} \in Y^{-1} \cup \{1\}$$

with $s := \overline{x^{-1}t}$, since $\overline{xs} = t$. We define $\psi\colon F \to \tilde{H}$ recursively by $\psi(1) := 1$ and

$$\psi(w) := (\overline{x^\epsilon z})^{-1}x^\epsilon\overline{z} \cdot \psi(z)$$

for reduced $w = x^\epsilon z$ with $x \in X$ and $\epsilon = \pm 1$. By what we have just shown, $\psi(w)$ is a product of elements from $Y \cup Y^{-1}$. Let $S := \{\psi(t^{-1}rt) : t \in T,\ r \in R\}$. We show $H \cong \langle Y \mid S\rangle$. For this, let $M := \langle S\rangle^{\tilde{H}} \trianglelefteq \tilde{H}$. First we verify $\psi(w) = \overline{w}^{-1}w$ for $w \in F$ by induction on $|w|$. This is clear for $w = 1$. For reduced $w = x^\epsilon z$ it follows that

$$\psi(w) = (\overline{x^\epsilon z})^{-1}x^\epsilon\overline{z} \cdot \overline{z}^{-1}z = \overline{w}^{-1}w.$$

For $h \in \tilde{H}$ it follows that $\psi(h) = (\overline{h})^{-1}h = h$. In particular, $\psi(t^{-1}rt) = t^{-1}rt \in N$ for $r \in R$ and $t \in T$. Hence $S \subseteq N$ and $M \subseteq N$ because of $N \trianglelefteq \tilde{H}$. For the reverse inclusion it suffices to show $g^{-1}rg \in M$ for $g \in F$ and $r \in R$. Let $g = th$ with $t \in T$ and $h \in \tilde{H}$. Then

$$g^{-1}rg = h^{-1}t^{-1}rth = h^{-1}\psi(t^{-1}rt)h \in M.$$

Hence $H \cong \tilde{H}/N = \tilde{H}/M = \langle Y \mid S\rangle$.

The second claim follows from Schreier's formula and the construction of S. □

Remark 2.14 (GAP).

```
G:=SymmetricGroup(6);;
FG:=Image(IsomorphismFpGroup(G));;
FH:=DerivedSubgroup(FG);; #= A_6
P:=PresentationSubgroup(FG,FH,"y"); #presentation with generators y_i
TzPrintPresentation(P); #statistics on generators and relations
```

Theorem 2.15 (Todd-Coxeter algorithm). *Let $G = \langle X \mid R\rangle$ be finitely presented and $H \leq G$ with $|G : H| < \infty$. Then there exists an algorithm that determines the action of G on G/H. In particular, $|G : H|$ can be computed from a generating set of H.*

Proof sketch. By Reidemeister-Schreier there exists a finite generating set Y of H. Let $G/H = \{H = H_1, \ldots, H_n\}$. For $y = x_1^{\epsilon_1} \ldots x_k^{\epsilon_k} \in Y$ let $t_y := (t_1, \ldots, t_k)$ with $x_i^{\epsilon_i} \ldots x_k^{\epsilon_k}H = H_{t_i}$ for $i = 1, \ldots, k$. Since $y \in H$, we have $t_1 = 1$. For $r = x_1^{\epsilon_1} \ldots x_l^{\epsilon_l} \in R$ let $T_r = (t_{ij}) \in \mathbb{N}^{n\times l}$ with $x_m^{\epsilon_m} \ldots x_l^{\epsilon_l}H_i = H_{t_{im}}$ for $m = 1, \ldots, l$. Since $r = 1$ in G, we have $t_{i1} = i$ for $i = 1, \ldots, n$. We fill the vectors t_y $(y \in Y)$ and matrices T_r $(r \in R)$ from left to right and from top to bottom by assigning new cosets and taking all logical consequences into account. For every new coset, a new row is appended to T_r. Since $|G : H| < \infty$, eventually all entries in t_y and T_r are filled. Moreover, every $x \in X$ occurs in some $r \in R$ or in a generator of H. In this way, one can read off the action of G on G/H. Since G acts transitively, all cosets really do appear in the tables, i. e., we can read off $n = |G : H|$. □

Example 2.16. Let $G = \langle x, y \mid x^3 = y^5 = (xy)^2 = 1\rangle$ and $H := \langle y\rangle$. The vector $t_y = (1)$ contains no information here. The first row of T_{y^5} is $(1, 1, 1, 1, 1)$. The assignment $H_2 := x^{-1}H_1 = x^2H$ yields the following entries in T_{x^3}, T_{y^5} and $T_{(xy)^2}$:

x	x	x
1	2	
2		

y	y	y	y	y
1	1	1	1	1
2				

x	y	x	y
1	2		1
2			

Now let $H_3 := x^{-1}H_2 = xH$. Then $x^{-1}H_3 = H_1$ and $y^4H_2 = y^{-1}x^{-1}H = xyH = H_3$.

x	x	x
1	2	3
2	3	1
3	1	2

y	y	y	y	y
1	1	1	1	1
2	3			
3				

x	y	x	y
1	2	3	1
2	3		
3	1	1	2

We continue by defining $H_4 := y^{-1}H_3$, $H_5 := x^{-1}H_4$, $H_6 := y^{-1}H_4$, $H_7 := x^{-1}H_5$ etc.

x	x	x
1	2	3
4	5	7
6	9	8
10	11	12

y	y	y	y	y
1	1	1	1	1
2	3	4	6	5
7	8	10	11	9
12	12	12	12	12

x	y	x	y
1	2	3	1
2	3	4	5
5	7	8	6
6	9	7	4
9	8	10	11
11	12	12	10

(redundant rows have been deleted). We can now read off $|G| = |G : H||H| = 12|H| \leq 60$. Indeed, the permutations $x = (1, 2, 3)$ and $y := (1, 4, 3, 5, 2)$ in A_5 satisfy the relations. By von Dyck's theorem, $G \cong \langle (1, 2, 3), (1, 4, 3, 5, 2)\rangle = A_5$.

Remark 2.17 (GAP)**.**

```
F:=FreeGroup(2);;
G:=F/[F.1^3,F.2^5,(F.1*F.2)^2];; #F.n is the n-th generator
H:=Subgroup(G,[G.2]);;
CT:=CosetTable(G,H);;
Display(TransposedMat(CT)); #only columns F.1^±1 and F.2^±1
f:=FactorCosetAction(G,H); #action on G/H
StructureDescription(Image(f)); #corresponding permutation group

G:=F/[F.1^2,F.2^3,(F.1*F.2)^7];;
CT:=CosetTable(G,TrivialSubgroup(G));; #aborts after 4096000 cosets (|G| = ∞)
```

For more complicated examples and better graphical output one can use the packages `ACE` or `ICT` (requires `xgap`). The table for $A_6 \leq M_{11}$ is displayed on the cover.

Theorem 2.18 (Moore)**.** *For $n \geq 2$ we have*

$$S_n \cong \langle x_1, \ldots, x_{n-1} \mid 1 = x_i^2 = (x_jx_{j+1})^3 = (x_kx_l)^2 \text{ for } k < l - 1\rangle.$$

Proof. Let G be the group on the right-hand side. For $n = 2$ we have $G \cong C_2 \cong S_2$. So let $n > 2$ and $H := \langle x_1, \ldots, x_{n-2}\rangle \leq G$. By induction, H is a quotient group of S_{n-1}. In particular, $|H| \leq (n - 1)!$. We show that G permutes the following cosets:

$$H, x_{n-1}H, x_{n-2}x_{n-1}H, \ldots, x_1 \ldots x_{n-1}H.$$

Certainly we have

$$x_i(x_ix_{i+1}\dots x_{n-1}H) = x_{i+1}\dots x_{n-1}H,$$
$$x_{i-1}(x_ix_{i+1}\dots x_{n-1}H) = x_{i-1}x_i\dots x_{n-1}H.$$

For $j < i-1$ we have $x_ix_j = (x_ix_j)^{-1} = x_jx_i$ and

$$x_j(x_ix_{i+1}\dots x_{n-1}H) = x_ix_{i+1}\dots x_{n-1}x_jH = x_ix_{i+1}\dots x_{n-1}H.$$

Now let $j > i$. Since $(x_{j-1}x_j)^3 = 1$, we have $x_{j-1}x_jx_{j-1} = x_jx_{j-1}x_j$. It follows that

$$\begin{aligned} x_j(x_ix_{i+1}\dots x_{n-1}H) &= x_i\dots x_{j-2}(x_jx_{j-1}x_j)x_{j+1}\dots x_{n-1}H = x_i\dots x_{j-2}(x_{j-1}x_jx_{j-1})x_{j+1}\dots x_{n-1}H \\ &= x_i\dots x_{n-1}x_{j-1}H = x_i\dots x_{n-1}H. \end{aligned}$$

Since G acts transitively on G/H in general, we have $|G:H| \le n$ and $|G| \le n!$.

Conversely, the transpositions $x_i' := (i, i+1) \in S_n$ for $i = 1, \dots, n-1$ satisfy the same relations, and since $S_n = \langle x_1', \dots, x_{n-1}' \rangle$, the claim follows. □

Theorem 2.19 (Moore). *For $n \ge 2$ we have*

$$A_n \cong \langle x_1, \dots, x_{n-2} \mid 1 = x_1^3 = x_2^2 = \dots = x_{n-2}^2 = (x_ix_{i+1})^3 = (x_kx_l)^2 \text{ for } k < l-1 \rangle.$$

Proof. Again let G be the right-hand side. For $n \le 3$ the claim holds. Let $n \ge 4$ and $H := \langle x_1, \dots, x_{n-3} \rangle \le G$. By induction, $|H| \le \frac{1}{2}(n-1)!$. We show that G permutes the following n cosets

$$H, x_{n-2}H, x_{n-3}x_{n-2}H, \dots, x_1\dots x_{n-2}H, x_1^2x_2\dots x_{n-2}H.$$

As long as no x_1 is involved, this works as in Theorem 2.18. For $i \ge 3$ we have $x_1x_i = x_ix_1^{-1}$ and

$$x_1x_i\dots x_{n-2}H = x_i\dots x_{n-2}x_1^{\pm 1}H = x_i\dots x_{n-2}H,$$
$$x_ix_1^{\pm 1}x_2\dots x_{n-2}H = x_1^{\mp 1}x_ix_2\dots x_{n-2}H = x_1^{\mp 1}x_2\dots x_{n-2}H.$$

From $(x_1x_2)^3 = 1$ it follows that $x_2x_1x_2 = x_1^{-1}x_2x_1^{-1}$ and

$$x_2x_1^{\pm 1}\dots x_{n-2}H = x_1^{\mp 1}x_2x_1^{\mp 1}x_3\dots x_{n-2}H = x_1^{\mp 1}x_2\dots x_{n-2}H.$$

As in Theorem 2.18 we obtain $|G| \le n|H| \le |A_n|$. Conversely, the elements $x_1 = (1,2,3)$, $x_i = (1,2)(i+1, i+2)$ $(i = 2, \dots, n-2)$ of A_n satisfy the given relations. □

Remark 2.20. Guralnick-Kantor-Kassabov-Lubotzky showed that all symmetric and alternating groups can be presented with two generators and eight relations (or three generators and seven relations[9]).

Definition 2.21. A reduced word $x_1^{\epsilon_1}\dots x_n^{\epsilon_n}$ of a free group is called *cyclically reduced* if $x_1^{\epsilon_1} \ne x_n^{-\epsilon_n}$.

Remark 2.22. For a group $\langle X \mid r \rangle$ with only one relation, after conjugation we can always assume that r is cyclically reduced. The following generalization of Lemma 1.24 now holds.

Theorem 2.23 (Magnus' Freiheitssatz). *Let $G = \langle X \mid r \rangle$, where r is cyclically reduced. If $x \in X$ occurs in r, then $\langle X \setminus \{x\} \rangle \le G$ is free with respect to $X \setminus \{x\}$.*

Proof. See Theorem 7.1 in [Camps et al., *Einführung in die kombinatorische und geometrische Gruppentheorie*, Heldermann Verlag, Lemgo, 2008] □

[9] The presentation was erroneous and was corrected by Huxford.

3 Automorphisms of Free Groups

Definition 3.1. Let $X \subseteq X'$ be alphabets and $R \subseteq F_X$ with normal closure $N \trianglelefteq F$. The transformations

- $(X, R) \to (X, R \cup \{r\})$ with $r \in N \setminus R$,
- $(X, R) \to (X \cup \{x\}, R \cup \{x^{-1}w\})$ with $x \in X' \setminus X$ and $w \in F_X$

and their inverses are called *Tietze transformations*.

Theorem 3.2. *Two finitely presented groups $\langle X_1 \mid R_1\rangle$ and $\langle X_2 \mid R_2\rangle$ are isomorphic if and only if (X_1, R_1) can be transformed into (X_2, R_2) by finitely many Tietze transformations.*

Proof. W. l. o. g. we may assume $X_1 \cap X_2 = \varnothing$. Let $F_i := F_{X_i}$ and $N_i := \langle R_i\rangle^{F_i} \trianglelefteq F_i$. For $r \in N_i$ we clearly have $\langle R_i \cup \{r\}\rangle^{F_i} = N_i$ and $G_i := \langle X_i \mid R_i\rangle \cong \langle X_i \mid R_i \cup \{r\}\rangle$. Now let $x \in X' \setminus X_i$ and $w \in F_i$. Then there exists an epimorphism $\varphi\colon F_{X_i \cup \{x\}} \to G_i$ with $\varphi(x_i) = x_i N_i$ for $x_i \in X_i$ and $\varphi(x) = wN_i$. Clearly,

$$N_i' := \langle R_i \cup \{x^{-1}w\}\rangle^{F_{X_i \cup \{x\}}} \subseteq \operatorname{Ker}(\varphi).$$

Conversely, let $y := y_1^{\epsilon_1} \dots y_n^{\epsilon_n} \in \operatorname{Ker}(\varphi)$ with $y_1, \dots, y_n \in X_i \cup \{x\}$. Let $k := |\{1 \le j \le n : y_j = x\}|$. If $k = 0$, then $y \in \operatorname{Ker}(\varphi) \cap F_i = N_i \subseteq N_i'$. Now let $k > 0$. To show $y \in N_i'$, we may assume $y_n^{\epsilon_n} = x$ after conjugation. Then it suffices to show $yx^{-1}w \in N_i'$, but this follows by induction on k. Hence $\operatorname{Ker}(\varphi) = N_i'$ and $G_i \cong \langle X_i \cup \{x\} \mid R_i \cup \{x^{-1}w\}\rangle$. If $(X_1 \mid R_1)$ can be transformed into $(X_2 \mid R_2)$ by Tietze transformations, then G_1 and G_2 are therefore isomorphic (this does not require a finite presentation).

Conversely, let $G := G_1 \cong G_2$. Then there exist epimorphisms $\varphi_i\colon F_i \to G$ with kernel N_i for $i = 1, 2$. Let $X := X_1 \cup X_2$ and $F := F_X$. Then there exists a homomorphism $\varphi\colon F \to G$ extending φ_1 and φ_2. For $x \in X_1$ we choose $w_x \in F_2$ with $\varphi(x) = \varphi_2(w_x) = \varphi(w_x)$. Let $S_1 := \{s_x := x^{-1}w_x : x \in X_1\} \subseteq F$. We define S_2 analogously. By an obvious (finite) sequence of Tietze transformations we transform (X_1, R_1) into $(X, R_1 \cup S_2)$. Clearly, $N := \langle R_1 \cup S_2\rangle^F \subseteq \operatorname{Ker}(\varphi)$. As above, one shows $\operatorname{Ker}(\varphi) \subseteq N$. In particular, $R_2 \cup S_1 \subseteq N$. Hence we can pass from $(X, R_1 \cup S_2)$ to $(X, R_1 \cup R_2 \cup S_1 \cup S_2)$ by Tietze transformations. The situation is now symmetric. Therefore, we can also pass from (X_2, R_2) to $(X, R_1 \cup R_2 \cup S_1 \cup S_2)$. Finally, the inverse Tietze transformations transform (X_1, R_1) into (X_2, R_2). □

Remark 3.3. The commands

`SimplifyPresentation` (= `TzGo`), `SimplifiedFpGroup`, `IsomorphismSimplifiedFpGroup`

silently perform Tietze transformations in order to simplify a finite presentation $\langle X \mid R\rangle$, i. e., $|X|$, $|R|$, and $\sum_{r \in R} l(r)$ are minimized. However, one can also apply the transformations in a targeted way.

```
G:=PerfectGroup(768000,10); #perfect group from the database
H:=Image(IsomorphismFpGroup(G));
P:=PresentationFpGroup(H,2); #presentation, print level=2, i.e., more output
TzGoGo(P); #iterates simplifications using Tietze transformations
TzPrintPresentation(P); #5 generators, 33 relations, total length 224
gen:=SmallGeneratingSet(G); #also works with 2 generators
H:=Image(IsomorphismFpGroupByGenerators(G,gen)); #use these
P:=PresentationFpGroup(H,2);
TzGoGo(P);
TzPrintPresentation(P); #2 generators, 59 relations, total length 3220
```

```
gen:=GeneratorsOfPresentation(P);
TzSubstitute(P,gen[1]^2*gen[2]); #perform a second Tietze transformation with w = x^2y
```

Definition 3.4. Let $F = F_X$ and $x, y \in X$ with $x \neq y$. The automorphisms

$$\alpha_x \colon F \to F, \qquad z \mapsto \begin{cases} x^{-1} & \text{if } z = x \\ z & \text{if } z \neq x \end{cases},$$

$$\beta_{xy} \colon F \to F, \qquad z \mapsto \begin{cases} xy & \text{if } z = x \\ z & \text{if } z \neq x \end{cases}$$

(where $z \in X$) are called *Nielsen transformations* (note the analogy to Gaussian elimination). Let $\mathrm{Aut}_N(F) := \langle \alpha_x, \beta_{xy} : x, y \in X, x \neq y \rangle \leq \mathrm{Aut}(F)$.

Remark 3.5. For distinct $x, y \in X$ we have

$$\alpha_x \beta_{xy}^{-1} \alpha_x^{-1}(x) = \alpha_x \beta_{xy}^{-1}(x^{-1}) = \alpha_x(yx^{-1}) = yx.$$

Applied to the pair (x, y) we have

$$\begin{aligned} \alpha_y \beta_{xy} \alpha_x \beta_{yx} \alpha_y \beta_{xy}(x, y) &= \alpha_y \beta_{xy} \alpha_x \beta_{yx} \alpha_y(xy, y) = \alpha_y \beta_{xy} \alpha_x \beta_{yx}(xy^{-1}, y^{-1}) \\ &= \alpha_y \beta_{xy} \alpha_x(y^{-1}, x^{-1}y^{-1}) = \alpha_y \beta_{xy}(y^{-1}, xy^{-1}) = \alpha_y(y^{-1}, x) = (y, x). \end{aligned}$$

Hence every permutation of X can be realized by Nielsen transformations.

Theorem 3.6. *Let $w_1, \ldots, w_n \in F = F_X$ and $\gamma \in \mathrm{Aut}(F)$. Then there exists a $\delta \in \mathrm{Aut}_N(F)$ with $\delta(w_i) = \gamma(w_i)$ for $i = 1, \ldots, n$. In particular, $\mathrm{Aut}(F) = \mathrm{Aut}_N(F)$ if $|X| < \infty$.*

Proof. Let $Y \subseteq X$ with $|Y| < \infty$ and $w_1, \ldots, w_n \in F_Y$. It suffices to construct a $\delta \in \mathrm{Aut}_N(F)$ with $\gamma(y) = \delta(y)$ for all $y \in Y$ (the given $w_1, \ldots, w_n$ are irrelevant from now on). Since $\langle \gamma^{-1}(X) \rangle = F$, there exists a finite subset $Z \subseteq X$ with $Y \subseteq \langle \gamma^{-1}(Z) \rangle$. We may assume $Y \subseteq Z = \{x_1, \ldots, x_m\}$. Let $w_i := \gamma^{-1}(x_i)$ be reduced for $i = 1, \ldots, m$. Suppose that there exist $1 \leq i, j \leq m$ with $|w_i w_j| < |w_i|$. Clearly, $i \neq j$ then. By replacing γ with $\beta_{x_i x_j}^{-1} \gamma$, the word w_i is replaced by $w_i w_j$, while w_k does not change for $k \neq i$. In this way, $\sum_{i=1}^m |w_i|$ becomes smaller. In the case $|w_i w_j| < |w_j|$ we can analogously perform the transformation $w_j \mapsto w_i w_j$ by Remark 3.5. Hence we may assume

$$|w_i w_j| \geq \max\{|w_i|, |w_j|\}$$

for $1 \leq i, j \leq m$. This means that at most half of w_i and w_j can cancel in the product $w_i w_j$. Now suppose that i, j, k exist with $|w_i w_j w_k| \leq |w_i| - |w_j| + |w_k|$. Let $w_i = as^{-1}$, $w_j = sbt^{-1}$ and $w_k = tc$, such that $w_i w_j = abt^{-1}$ and $w_j w_k = sbc$ are reduced (note that only half of w_j can cancel on the left and on the right, respectively). We have

$$|a| + |b| + |c| = |w_i w_j w_k| \leq |w_i| - |w_j| + |w_k| = |a| - |b| + |c|,$$

i. e., $b = 1$ and $|s| = \frac{1}{2}|w_j| = |t|$. Since $|w_i| = |a| + |s| = |w_i w_j| \geq |w_j| = 2|s|$, we have $|s| \leq \frac{1}{2}|w_i|$ and analogously $|s| \leq \frac{1}{2}|w_k|$. Hence we can replace w_i by $w_i w_j$ or w_k by $w_j w_k$ without changing $\sum_{i=1}^m |w_i|$. For $w \in F_Z$ let $L(w)$ be the left subword of w of length $\lfloor (|w| + 1)/2 \rfloor$. Let $\leq$ be the lexicographic (well-)order on F_Z with

$$1 < x_m^{-1} < x_{m-1}^{-1} < \ldots < x_1^{-1} < x_1 < x_2 < \ldots < x_m < x_m^{-2} < x_m^{-1} x_{m-1}^{-1} < \ldots.$$

We write $w \prec v$ if $\min\{L(w), L(w^{-1})\} < \min\{L(v), L(v^{-1})\}$ or $(\min\{L(w), L(w^{-1})\} = \min\{L(v), L(v^{-1})\}$ and $\max\{L(w), L(w^{-1})\} < \max\{L(v), L(v^{-1})\})$, where the minimum/maximum is taken with respect to $\leq$. In the situation above, if $s < t$, then $w_j w_k = sc \prec tc = w_k$, because $L(c^{-1}s^{-1}) = L(c^{-1}) = L(c^{-1}t^{-1})$. If $t < s$, then $w_i w_j = at^{-1} \prec as^{-1} = w_i$. The case $s = t$ is excluded since $w_j = st^{-1} \neq 1$. By suitable Nielsen transformations we can therefore achieve that the w_i are as small as possible with respect to $\prec$ (since $\leq$ is a well-order, the w_i cannot become arbitrarily small). The property $|w_i w_j| \geq \max\{|w_i|, |w_j|\}$ is preserved in the process. In the end,

$$|w_i w_j w_k| > |w_i| - |w_j| + |w_k|$$

for $1 \leq i, j, k \leq m$, i. e., w_j does not cancel completely in $w_i w_j w_k$. The same procedure can be carried out more generally with the elements $w_i^{\pm 1}$.

Since $Y \subseteq \langle \gamma^{-1}(Z) \rangle$, every $x_i \in Y$ can be written as a product of the w_j, say $x_i = w_{j_1}^{\epsilon_1} \dots w_{j_k}^{\epsilon_k}$. By construction, however, $1 = |x_i| = |w_{j_1}^{\epsilon_1} \dots w_{j_k}^{\epsilon_k}| \geq k$, i. e., $x_i = w_{j_1}^{\epsilon_1}$. After further Nielsen transformations as in Remark 3.5, one finally achieves $w_i = x_i$ for all $x_i \in Y$. □

Remark 3.7.

(i) Nielsen gave a finite presentation of $\mathrm{Aut}(F_n)$ with respect to the Nielsen transformations. Newman showed that $\mathrm{Aut}(F_n)$ can be generated by only two automorphisms (of infinite order).

(ii) In GAP one can construct Nielsen transformations as follows:

```
F:=FreeGroup("x","y");;
AssignGeneratorVariables(F);;
FreeGroupAutomorphismsGeneratorO(F); #Nielsen transformation α_x
FreeGroupAutomorphismsGeneratorU(F); #β_xy
FreeGroupAutomorphismsGeneratorP(F); #(x,y) ↦ (y,x)
A:=AutomorphismGroup(F);; #Aut(F)
iso:=IsomorphismFpGroup(A);; #isomorphism from A to a finitely presented group
a:=GroupHomomorphismByImages(F,F,[x,y],[x^y,x*y]); #a ∈ A
a^iso; #a as a word in α_x, β_xy and (x,y) ↦ (y,x)
```

Theorem 3.8. *Let $\Phi\colon \mathrm{Aut}(F_n) \to \mathrm{GL}(n, \mathbb{Z})$, where $\Phi(\alpha)_{ij}$ is the exponent sum of x_j in $\alpha(x_i)$. Then Φ is an epimorphism.*

Proof. The map Φ arises from the restriction

$$\mathrm{Aut}(F_n) \to \mathrm{Aut}(F_n/F_n') \cong \mathrm{Aut}(\mathbb{Z}^n) \cong \mathrm{GL}(n, \mathbb{Z})$$

and is therefore a well-defined homomorphism. For surjectivity it suffices to show that every matrix $A \in \mathrm{GL}(n, \mathbb{Z})$ is a product of $\Phi(\alpha_x)$ and $\Phi(\beta_{xy})$. Left multiplication (or right multiplication, respectively) by $\Phi(\alpha_x)$ has the effect that a row (or column, respectively) of A is multiplied by -1. By $\Phi(\beta_{xy})$ one can add a row (or column) of A to another row (or column). As in Remark 3.5, one can also swap rows and columns of A. With these operations, A can be brought into Smith normal form. Since $\det A = \pm 1$, all elementary divisors are 1, i. e., A can be transformed into the identity matrix. □

4 Group Extensions

Remark 4.1. By the Jordan-Hölder theorem, all finite groups (or those possessing a composition series) are built up from simple groups. The finite simple groups are, as is well known, completely classified. It remains to investigate how a group is composed of normal subgroups and quotient groups.

Definition 4.2.

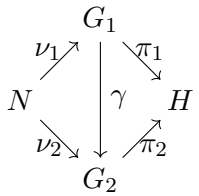
- A (*group*) *extension* of H by N is a short exact sequence of groups

$$1 \to N \xrightarrow{\nu} G \xrightarrow{\pi} H \to 1,$$

  i. e., $N \cong \nu(N) \trianglelefteq G$ and $G/\nu(N) \cong \pi(G) = H$. [10]

- We call G a *split* extension if there exists a homomorphism $\rho\colon H \to G$ with $\pi \circ \rho = \mathrm{id}_H$. In this case, $\rho(H)$ is a complement of $\nu(N)$ in G and $G \cong N \rtimes H$. Conversely, every semidirect product is a split extension.

- Two group extensions G_1, G_2 of H by N are called *equivalent* if there exists a homomorphism $\gamma\colon G_1 \to G_2$ with $\gamma \circ \nu_1 = \nu_2$ and $\pi_2 \circ \gamma = \pi_1$. That is, the following diagram commutes:



Remark 4.3.

(i) Every group G with normal subgroup N is an extension of $H = G/N$ by N, by choosing the inclusion for ν and the canonical epimorphism for π.

(ii) Let G_1 and G_2 be equivalent extensions via $\gamma\colon G_1 \to G_2$. Let $x \in \mathrm{Ker}(\gamma)$. Since $\pi_1(x) = \pi_2(\gamma(x)) = 1$ we have $x \in \nu_1(N)$, say $x = \nu_1(y)$. From $\nu_2(y) = \gamma(\nu_1(y)) = \gamma(x) = 1$ it follows that $y = 1 = x$. Hence γ is injective. For $g \in G_2$ there exists $x \in G_1$ with $\pi_2(\gamma(x)) = \pi_1(x) = \pi_2(g)$. Thus there exists $y \in N$ with $g^{-1}\gamma(x) = \nu_2(y) = \gamma(\nu_1(y))$. This shows $g \in \gamma(G_1)$ and γ is surjective. Altogether, $G_1 \cong G_2$. Hence equivalence of groups is an equivalence relation.

(iii) However, isomorphic extensions need not be equivalent: Let

$$G_1 = G_2 = \langle x, y \mid x^4 = y^2 = 1,\ yxy = x^{-1} \rangle \cong D_8,$$

$N = \langle x^2, y \rangle \cong C_2^2$ and $H = \langle xy \rangle$. Let ν_1 be the inclusion, $\nu_2(x^2) = y$ and $\nu_2(y) = x^2$. Let π_1 and π_2 be the canonical epimorphisms. Since $\mathrm{Z}(G_i) = \langle x^2 \rangle$, there cannot exist $\gamma \in \mathrm{Aut}(G_i)$ with $\gamma(x^2) = \gamma(\nu_1(x^2)) = \nu_2(x^2) = y$. Nevertheless, it is useful to be able to determine extensions up to equivalence.

(iv) Let G_1 and G_2 be equivalent extensions of H by N with respect to $\gamma\colon G_1 \to G_2$. Suppose that G_1 splits with $\rho_1\colon H \to G_1$ and $\pi_1\rho_1 = \mathrm{id}_H$. For $\rho_2 := \gamma\rho_1\colon H \to G_2$ we have $\pi_2\rho_2 = \pi_2\gamma\rho_1 = \pi_1\rho_1 = \mathrm{id}_H$. Thus G_2 splits as well.

[10] Caution: This usage is not uniform in the literature.

(v) Without the methods described below, one can construct extensions in GAP with the `grpconst` package:

```
LoadPackage("grpconst",false);
G:=AlternatingGroup(6);; #the package expects a permutation group
up:=UpwardsExtensions(G,2)[2]; #extensions of G by C_2
List(up,IdGroup);
P:=DihedralGroup(IsPermGroup,32);;
up:=CyclicExtensions(P,2);; #is faster, but does not reduce up to isomorphism
Size(up)=Size(Set(up,IdGroup)); #false, so the list is redundant
UpwardsExtensionsNoCentre(G,2); #works only for G = G' and Z(G) = 1
```

Definition 4.4. Let H and N be groups. For $x \in N$ let $c_x \in \mathrm{Inn}(N)$ be the inner automorphism (i. e., $c_x(y) = xyx^{-1}$). A pair of maps

$$\alpha\colon H \to \mathrm{Aut}(N),\ x \mapsto \alpha_x, \qquad \kappa\colon H \times H \to N$$

is called a *parameter system* of H by N if for all $x, y, z \in H$ we have:

- $\boxed{\alpha_x\alpha_y = c_{\kappa(x,y)}\alpha_{xy},}$
- $\boxed{\kappa(x,y)\kappa(xy,z) = \alpha_x(\kappa(y,z))\kappa(x,yz).}$

In this case, one calls α *automorphism system* and κ *factor set*. A parameter system (or factor set) is called *normalized* if $\kappa(x,1) = \kappa(1,x) = 1$ for all $x \in H$ (in this case, $\alpha_1 = \mathrm{id}_N$). Finally, κ is called *trivial* if $\kappa(x,y) = 1$ for all $x, y \in H$ (in this case, α is a homomorphism).

Lemma 4.5. *Every extension of H by N determines a (normalized) parameter system.*

Proof. Let $N \xrightarrow{\nu} G \xrightarrow{\pi} H$ be an extension. For $x \in H$ we choose $\tilde{x} \in G$ with $\pi(\tilde{x}) = x$ and $\tilde{1} = 1$. Then $\alpha_x\colon N \to N,\ y \mapsto \nu^{-1}(\tilde{x}\nu(y)\tilde{x}^{-1})$ is an automorphism (note: ν is injective) and $\alpha_1 = \mathrm{id}_N$. For $x, y \in H$ we have $\pi(\widetilde{xy}) = xy = \pi(\tilde{x})\pi(\tilde{y}) = \pi(\tilde{x}\tilde{y})$. Hence there exists $\kappa(x,y) \in N$ with $\tilde{x}\tilde{y} = \nu(\kappa(x,y))\widetilde{xy}$. Moreover, $\kappa(x,1) = \kappa(1,x) = 1$. For $g \in N$ we have

$$\begin{aligned}(\alpha_x\alpha_y)(g) &= \alpha_x(\nu^{-1}(\tilde{y}\nu(g)\tilde{y}^{-1})) = \nu^{-1}(\tilde{x}\tilde{y}\nu(g)\tilde{y}^{-1}\tilde{x}^{-1}) = \nu^{-1}(\nu(\kappa(x,y))\widetilde{xy}\nu(g)\widetilde{xy}^{-1}\nu(\kappa(x,y)^{-1}))\\ &= \kappa(x,y)\nu^{-1}(\widetilde{xy}\nu(g)\widetilde{xy}^{-1})\kappa(x,y)^{-1} = c_{\kappa(x,y)}\alpha_{xy}(g).\end{aligned}$$

For $x, y, z \in H$ we have

$$\begin{aligned}\nu(\kappa(x,y))\nu(\kappa(xy,z))\widetilde{xyz} &= \nu(\kappa(x,y))\widetilde{xy}\tilde{z} = (\tilde{x}\tilde{y})\tilde{z} = \tilde{x}(\tilde{y}\tilde{z}) = \tilde{x}\nu(\kappa(y,z))\widetilde{yz}\\ &= \nu(\alpha_x(\kappa(y,z)))\tilde{x}\widetilde{yz} = \nu(\alpha_x(\kappa(y,z)))\nu(\kappa(x,yz))\widetilde{xyz}.\end{aligned}$$

Since ν is injective, the claim follows. □

Example 4.6.

(i) If $G = N \rtimes H$, we can choose $\tilde{x} \in H$ in the above proof. We then obtain the trivial factor set and, for $\alpha\colon H \to \mathrm{Aut}(N)$, the conjugation action.

(ii) If $x \in H$ with $x^2 \neq 1$, we can choose $\widetilde{x^{-1}} = \tilde{x}^{-1}$ in the above proof. We then obtain $\kappa(x, x^{-1}) = 1$.

(iii) Let $G = \langle x, y\rangle = Q_8$, $N = \langle x\rangle \cong C_4$, ν the inclusion and $H = \langle z\rangle \cong C_2$. We can choose $\tilde{z} = y$. We obtain $\kappa(z,z) = \tilde{z}^2(\widetilde{z^2})^{-1} = y^2 \neq 1$.

(iv) Let $n \in \mathbb{N}$ and $H = \langle x \rangle \cong C_n \cong \langle a \rangle = N$. For $i, j \in \mathbb{Z}$ let $\alpha_{x^i} := \mathrm{id}_N$ and $\kappa(x^i, x^j) := a^{ij}$ (well defined!). Then we have

$$\kappa(x^i, x^j)\kappa(x^i x^j, x^k) = x^{ij+(i+j)k} = x^{jk+i(j+k)} = \kappa(x^j, x^k)\kappa(x^i, x^j x^k).$$

Hence (α, κ) is a normalized parameter system.

Lemma 4.7. *Every normalized parameter system of H by N determines an extension.*

Proof. Let (α, κ) be a normalized parameter system of H by N. We consider the set $G = N \times H$ with the multiplication

$$(a, x)(b, y) := (a\alpha_x(b)\kappa(x, y), xy).$$

Then we have

$$\begin{aligned}
\big((a, x)(b, y)\big)(c, z) &= (a\alpha_x(b)\kappa(x, y), xy)(c, z) = (a\alpha_x(b)\kappa(x, y)\alpha_{xy}(c)\kappa(xy, z), xyz) \\
&= (a\alpha_x(b)\kappa(x, y)\alpha_{xy}(c)\kappa(x, y)^{-1}\kappa(x, y)\kappa(xy, z), xyz) \\
&= (a\alpha_x(b)\alpha_x(\alpha_y(c))\alpha_x(\kappa(y, z))\kappa(x, yz), xyz) \\
&= (a\alpha_x(b\alpha_y(c)\kappa(y, z))\kappa(x, yz), xyz) \\
&= (a, x)(b\alpha_y(c)\kappa(y, z), yz) = (a, x)\big((b, y)(c, z)\big).
\end{aligned}$$

Hence the multiplication is associative. Since κ is normalized, $(1, 1)$ is an identity element. Inverse elements are obtained by

$$(\kappa(y^{-1}, y)^{-1}\alpha_{y^{-1}}(b)^{-1}, y^{-1})(b, y) = (\kappa(y^{-1}, y)^{-1}\alpha_{y^{-1}}(b)^{-1}\alpha_{y^{-1}}(b)\kappa(y^{-1}, y), y^{-1}y) = (1, 1).$$

Hence G is a group. Clearly, $\nu\colon N \to G$, $a \mapsto (a, 1)$ is a monomorphism and $\pi\colon G \to H$, $(a, x) \mapsto x$ is an epimorphism with $\nu(N) = N \times 1 = \mathrm{Ker}(\pi)$. □

Remark 4.8.

(i) For the trivial factor set, the above proof yields the semidirect product $N \rtimes H$.

(ii) Let (α, κ) be a normalized parameter system with $\alpha_x = \mathrm{id}_N$ for all $x \in H$ (we write $\alpha = 1$ for this). Since $c_{\kappa(x,y)} = c_{\kappa(x,y)}\alpha_{xy} = \alpha_x\alpha_y = \mathrm{id}_N$ we have $\kappa(H \times H) \leq \mathrm{Z}(N)$. Let G be the corresponding extension and $K := \langle (1, x) : x \in H \rangle$. Because of

$$\begin{aligned}
(a, x)(1, y)(a, x)^{-1} &= (a\kappa(x, y), xy)(\kappa(x^{-1}, x)^{-1}a^{-1}, x^{-1}) \\
&= (a\kappa(x, y)\kappa(x^{-1}, x)^{-1}a^{-1}\kappa(xy, x^{-1}), xyx^{-1}) = (1, x)(1, y)(1, x)^{-1} \in K
\end{aligned}$$

we have $K \trianglelefteq G$ and $N \cap K \leq \kappa(H \times H) \leq \mathrm{Z}(G)$. Hence $G = N * K$ is a central product.

Definition 4.9. Two parameter systems (α, κ), (α', κ') are called *equivalent* if there exists a map $\varphi\colon H \to N$ with

- $\boxed{\alpha'_x = c_{\varphi(x)}\alpha_x,}$
- $\boxed{\kappa'(x, y) = \varphi(x)\alpha_x(\varphi(y))\kappa(x, y)\varphi(xy)^{-1}}$

for all $x, y \in H$. In this case, we write $(\alpha, \kappa) \sim (\alpha', \kappa')$.

Remark 4.10. The maps $H \to N$ form a group $C^1(H,N)$ with respect to $(\varphi\psi)(x) := \varphi(x)\psi(x)$ $(x \in H)$, which is isomorphic to $\times_{h\in H} N$. We show that ${}^{\varphi}(\alpha,\kappa) := (\alpha',\kappa')$ as in Definition 4.9 defines an action of $C^1(H,N)$ on the set of parameter systems. First we must show that (α',κ') is a parameter system:

$$\begin{aligned}
\alpha'_x\alpha'_y &= c_{\varphi(x)}\alpha_x c_{\varphi(y)}\alpha_y = c_{\varphi(x)}\alpha_x c_{\varphi(y)}\alpha_x^{-1}\alpha_x\alpha_y = c_{\varphi(x)}c_{\alpha_x(\varphi(y))}c_{\kappa(x,y)}\alpha_{xy} = c_{\kappa'(x,y)}\alpha'_{xy}\\
\kappa'(x,y)\kappa'(xy,z) &= \varphi(x)\alpha_x(\varphi(y))\kappa(x,y)\alpha_{xy}(\varphi(z))\kappa(xy,z)\varphi(xyz)^{-1}\\
&= \varphi(x)\alpha_x(\varphi(y))(c_{\kappa(x,y)}\alpha_{xy})(\varphi(z))\kappa(x,y)\kappa(xy,z)\varphi(xyz)^{-1}\\
&= \varphi(x)\alpha_x(\varphi(y))(\alpha_x\alpha_y)(\varphi(z))\alpha_x(\kappa(y,z))\kappa(x,yz)\varphi(xyz)^{-1}\\
&= \varphi(x)\alpha_x(\varphi(y)\alpha_y(\varphi(z))\kappa(y,z))\kappa(x,yz)\varphi(xyz)^{-1}\\
&= (c_{\varphi(x)}\alpha_x)(\kappa'(y,z))\varphi(x)\alpha_x(\varphi(yz))\kappa(x,yz)\varphi(xyz)^{-1}\\
&= \alpha'_x(\kappa'(y,z))\kappa'(x,yz).
\end{aligned}$$

Clearly, ${}^{1}(\alpha,\kappa) = (\alpha,\kappa)$. Let $\varphi,\psi \in C^1(H,N)$ with ${}^{\psi}(\alpha,\kappa) = (\alpha',\kappa')$. Then we have

$$\begin{aligned}
c_{\varphi(x)}\alpha'_x &= c_{\varphi(x)}c_{\psi(x)}\alpha_x = c_{\varphi(x)\psi(x)}\alpha_x = c_{(\varphi\psi)(x)}\alpha_x,\\
\varphi(x)\alpha'_x(\varphi(y))\kappa'(x,y)\varphi(xy)^{-1} &= \varphi(x)\psi(x)\alpha_x(\varphi(y))\alpha_x(\psi(y))\kappa(x,y)\psi(xy)^{-1}\varphi(xy)^{-1}\\
&= (\varphi\psi)(x)\alpha_x((\varphi\psi)(y))\kappa(x,y)(\varphi\psi)(xy)^{-1}.
\end{aligned}$$

This shows ${}^{\varphi}({}^{\psi}(\alpha,\kappa)) = {}^{\varphi\psi}(\alpha,\kappa)$. Thus the equivalence classes of parameter systems are the orbits under $C^1(H,N)$. In particular, $\sim$ is an equivalence relation. The number of equivalence classes could be determined with Burnside's lemma.

Lemma 4.11. *Every parameter system is equivalent to a normalized parameter system.*

Proof. Let (α,κ) be an arbitrary parameter system and $\zeta := \kappa(1,1) \in N$. Since $\alpha_1\alpha_1 = c_\zeta\alpha_1$ we have $\alpha_1 = c_\zeta$. From

$$\zeta\kappa(1\cdot 1,x) = \alpha_1(\kappa(1,x))\kappa(1,1x) = \zeta\kappa(1,x)\zeta^{-1}\kappa(1,x)$$

we obtain $\kappa(1,x) = \zeta$ for all $x \in H$. With $\kappa(x,1)\kappa(x1,1) = \alpha_x(\zeta)\kappa(x,1\cdot 1)$ we have $\kappa(x,1) = \alpha_x(\zeta)$.

Now define $\varphi(1) := \zeta^{-1}$ and $\varphi(y) := 1$ for $y \in H \setminus \{1\}$. Then we have

$$\begin{aligned}
\varphi(x)\alpha_x(\varphi(1))\kappa(x,1)\varphi(x)^{-1} &= \varphi(x)\alpha_x(\zeta)^{-1}\alpha_x(\zeta)\varphi(x)^{-1} = 1,\\
\varphi(1)\alpha_1(\varphi(x))\kappa(1,x)\varphi(x)^{-1} &= \zeta^{-1}\zeta\varphi(x)\zeta^{-1}\zeta\varphi(x)^{-1} = 1
\end{aligned}$$

for all $x \in H$. In this way we obtain an equivalent normalized parameter system. □

Lemma 4.12. *Equivalent extensions define equivalent parameter systems.*

Proof. Let $N \xrightarrow{\nu_1} G_1 \xrightarrow{\pi_1} H$ and $N \xrightarrow{\nu_2} G_2 \xrightarrow{\pi_2} H$ be equivalent extensions via $\gamma\colon G_1 \to G_2$. For $x \in H$ let $\tilde{x} \in G_1$ and $\bar{x} \in G_2$ with $\pi_1(\tilde{x}) = x = \pi_2(\bar{x})$. As in the proof of Lemma 4.5 we obtain parameter systems (α,κ) and (α',κ') with

$$\begin{aligned}
\alpha_x(g) &= \nu_1^{-1}(\tilde{x}\nu_1(g)\tilde{x}^{-1}), & \kappa(x,y) &= \nu_1^{-1}(\tilde{x}\tilde{y}\widetilde{xy}^{-1}),\\
\alpha'_x(g) &= \nu_2^{-1}(\bar{x}\nu_2(g)\bar{x}^{-1}), & \kappa'(x,y) &= \nu_2^{-1}(\bar{x}\bar{y}\overline{xy}^{-1})
\end{aligned}$$

for $x,y \in H$ and $g \in N$. Since $\pi_2(\gamma(\tilde{x})) = \pi_1(\tilde{x}) = x = \pi_2(\bar{x})$ we can define

$$\varphi(x) := \nu_2^{-1}(\bar{x}\gamma(\tilde{x})^{-1}) \in N$$

for $x \in H$. Because of $\gamma\nu_1 = \nu_2$ we have $\varphi(x) = \nu_1^{-1}(\gamma^{-1}(\bar{x})\tilde{x}^{-1})$. It follows

$$\begin{aligned}(c_{\varphi(x)}\alpha_x)(g) &= \nu_1^{-1}\big(\gamma^{-1}(\bar{x})\tilde{x}^{-1}\big)\nu_1^{-1}\big(\tilde{x}\nu_1(g)\tilde{x}^{-1}\big)\nu_1^{-1}\big(\tilde{x}\gamma^{-1}(\bar{x})^{-1}\big) = \nu_1^{-1}\big(\gamma^{-1}(\bar{x})\nu_1(g)\gamma^{-1}(\bar{x})^{-1}\big)\\ &= \nu_1^{-1}\big(\gamma^{-1}(\bar{x}\nu_2(g)\bar{x}^{-1})\big) = \nu_2^{-1}\big(\bar{x}\nu_2(g)\bar{x}^{-1}\big) = \alpha'_x(g)\end{aligned}$$

and

$$\begin{aligned}\varphi(x)\alpha_x(\varphi(y))\kappa(x,y)\varphi(xy)^{-1} &= \nu_1^{-1}\big(\gamma^{-1}(\bar{x})\tilde{x}^{-1}\big)\nu_1^{-1}\big(\tilde{x}\nu_1(\varphi(y))\tilde{x}^{-1}\big)\nu_1^{-1}\big(\tilde{x}\tilde{y}\widetilde{xy}^{-1}\big)\nu_1^{-1}\big(\widetilde{xy}\gamma^{-1}(\overline{xy})^{-1}\big)\\ &= \nu_1^{-1}\big(\gamma^{-1}(\bar{x})\nu_1(\varphi(y))\tilde{y}\gamma^{-1}(\overline{xy})^{-1}\big)\\ &= \nu_1^{-1}\big(\gamma^{-1}(\bar{x})\gamma^{-1}(\bar{y})\gamma^{-1}(\overline{xy})^{-1}\big) = \nu_2^{-1}\big(\bar{x}\bar{y}\overline{xy}^{-1}\big) = \kappa'(x,y).\end{aligned}$$

Hence (α,κ) and (α',κ') are equivalent. □

Lemma 4.13. *Equivalent (normalized) parameter systems define equivalent extensions.*

Proof. Let (α,κ) and (α',κ') be equivalent parameter systems via $\varphi\colon H \to N$. By Lemma 4.11 we may assume that both parameter systems are normalized. Then we have

$$1 = \kappa'(1,x) = \varphi(1)\alpha_1(\varphi(x))\kappa(1,x)\varphi(x)^{-1} = \varphi(1).$$

We construct $(G_1,\cdot)$ and $(G_2,*)$ as in the proof of Lemma 4.7. Since G_1 and G_2 are equal as sets, we can define $\gamma\colon G_1 \to G_2$, $(a,x) \mapsto (a\varphi(x)^{-1},x)$. For $(a,x),(b,y) \in G_1$ we have

$$\begin{aligned}\gamma(a,x) * \gamma(b,y) &= (a\varphi(x)^{-1},x) * (b\varphi(y)^{-1},y) = (a\varphi(x)^{-1}\alpha'_x(b\varphi(y)^{-1})\kappa'(x,y),xy)\\ &= (a\alpha_x(b\varphi(y)^{-1})\alpha_x(\varphi(y))\kappa(x,y)\varphi(xy)^{-1},xy) = (a\alpha_x(b)\kappa(x,y)\varphi(xy)^{-1},xy)\\ &= \gamma(a\alpha_x(b)\kappa(x,y),xy) = \gamma((a,x)\cdot(b,y)),\end{aligned}$$

i. e., γ is a homomorphism. Moreover, we have

$$\begin{aligned}(\gamma\nu_1)(a) &= \gamma(a,1) = (a\varphi(1)^{-1},1) = (a,1) = \nu_2(a),\\ (\pi_2\gamma)(a,x) &= \pi_2(a\varphi(x)^{-1},x) = x = \pi_1(a,x).\end{aligned}$$

Hence G_1 and G_2 are equivalent. □

Theorem 4.14 (Schreier)**.** *There is a bijection between the set of equivalence classes of extensions of H by N and the set of equivalence classes of (normalized) parameter systems of H by N.*

Proof. By Lemma 4.12 and Lemma 4.13 there are well-defined maps between the sets of equivalence classes. We show that they are inverse to each other.

Let the extension $N \xrightarrow{\nu_1} G_1 \xrightarrow{\pi_1} H$ define the normalized parameter system (α,κ) with the elements $\tilde{x} \in G_1$ as in Lemma 4.5. From (α,κ) we construct the extension $N \xrightarrow{\nu_2} G_2 \xrightarrow{\pi_2} H$ as in Lemma 4.7. Every element in G_1 can be written uniquely in the form $\nu_1(a)\tilde{x}$ with $a \in N$ and $x \in H$. We define $\gamma\colon G_1 \to G_2$, $\nu_1(a)\tilde{x} \mapsto (a,x)$. For $\nu_1(a)\tilde{x},\nu_1(b)\tilde{y} \in G_1$ we have

$$\begin{aligned}\gamma\big(\nu_1(a)\tilde{x}\cdot\nu_1(b)\tilde{y}\big) &= \gamma\big(\nu_1(a)\tilde{x}\nu_1(b)\tilde{x}^{-1}\cdot\tilde{x}\tilde{y}\big) = \gamma\big(\nu_1(a)\nu_1(\alpha_x(b))\cdot\nu_1(\kappa(x,y))\widetilde{xy}\big)\\ &= (a\alpha_x(b)\kappa(x,y),xy) = (a,x) * (b,y) = \gamma(\nu_1(a)\tilde{x}) * \gamma(\nu_1(b)\tilde{y}),\end{aligned}$$

i. e., γ is a homomorphism with

$$(\gamma\nu_1)(a) = \gamma(\nu_1(a)\tilde{1}) = (a,1) = \nu_2(a),$$
$$(\pi_2\gamma)(\nu_1(a)\tilde{x}) = \pi_2(a,x) = x = \pi_1(\tilde{x}) = \pi_1(\nu_1(a)\tilde{x}).$$

Hence G_1 and G_2 are equivalent.

Conversely, let first (α,κ) be given normalized. We construct the extension G and from it the parameter system (α',κ') by means of the elements $\tilde{x} = (1,x) \in G$ for $x \in H$. First we have

$$\kappa(x,x^{-1}) = \kappa(x,x^{-1})\kappa(xx^{-1},x) = \alpha_x(\kappa(x^{-1},x))\kappa(x,x^{-1}x) = \alpha_x(\kappa(x^{-1},x)) \tag{4.1}$$

for $x \in H$. From this we obtain

$$\begin{aligned}
\alpha'_x(g) &= \nu^{-1}(\tilde{x}\nu(g)\tilde{x}^{-1}) = \nu^{-1}\big((1,x)(g,1)(1,x)^{-1}\big) = \nu^{-1}\big((\alpha_x(g),x)(\kappa(x^{-1},x)^{-1},x^{-1})\big)\\
&= \nu^{-1}\big(\alpha_x(g)\alpha_x(\kappa(x^{-1},x)^{-1})\kappa(x,x^{-1}),1\big) = \alpha_x(g),\\
\kappa'(x,y) &= \nu^{-1}\big(\tilde{x}\tilde{y}\widetilde{xy}^{-1}\big) = \nu^{-1}\big((1,x)(1,y)(\kappa((xy)^{-1},xy)^{-1},(xy)^{-1})\big)\\
&= \nu^{-1}\big((\kappa(x,y),xy)(\kappa((xy)^{-1},xy)^{-1},(xy)^{-1})\big)\\
&= \nu^{-1}\big(\kappa(x,y)\alpha_{xy}(\kappa((xy)^{-1},xy)^{-1})\kappa(xy,(xy)^{-1}),1\big)\\
&= \nu^{-1}(\kappa(x,y),1) = \kappa(x,y)
\end{aligned}$$

for $x,y \in H$ and $g \in N$. This shows $(\alpha,\kappa) = (\alpha',\kappa')$. □

Remark 4.15. By Remark 4.3, the split extensions correspond, up to equivalence, exactly to the parameter systems with trivial factor set.

Theorem 4.16. *Let $H = \langle x\rangle \cong C_n$ and $\beta \in \mathrm{Aut}(N)$. There exists a parameter system (α,κ) with $\alpha_x = \beta$ if and only if there exists $a \in N$ with $\beta(a) = a$ and $\beta^n = c_a$.*

Proof. Let (α,κ) be a parameter system with $\alpha_x = \beta$. In the case $n = 1$ the claim holds with $a := \kappa(1,1)$, since

$$\alpha_1\alpha_1 = c_{\kappa(1,1)}\alpha_1,$$
$$\kappa(1,1)\kappa(1,1) = \alpha_1(\kappa(1,1))\kappa(1,1).$$

For $n \geq 2$ we may pass to a normalized parameter system without changing α_x (see the proof of Lemma 4.11). We construct the corresponding extension $G = N \times H$. For $\tilde{x} := (1,x) \in G$ and $b \in N$ we have

$$\tilde{x}\nu(b)\tilde{x}^{-1} = (1,x)(b,1)(1,x)^{-1} = (\alpha_x(b),x)(\kappa(x^{-1},x)^{-1},x^{-1}) \overset{(4.1)}{=} (\alpha_x(b),1) = \nu(\beta(b)).$$

Hence $\beta = \nu^{-1}c_{\tilde{x}}\nu$. Let $(a,1) := \tilde{x}^n$. Then $\beta^n = \nu^{-1}c_{(a,1)}\nu = c_a$ and

$$\beta(a) = \nu^{-1}(\tilde{x}\nu(a)\tilde{x}^{-1}) = \nu^{-1}(\tilde{x}\tilde{x}^n\tilde{x}^{-1}) = \nu^{-1}(a,1) = a.$$

Conversely, let $\beta^n = c_a$ and $\beta(a) = a$ for some $a \in N$. We define $\alpha_{x^i} := \beta^i$ and

$$\kappa(x^i,x^j) := \begin{cases} 1 & \text{if } i+j < n,\\ a & \text{if } i+j \geq n \end{cases}$$

for $0 \leq i, j \leq n-1$. Then

$$\alpha_{x^i}\alpha_{x^j} = \beta^{i+j} = \begin{cases} \alpha_{x^{i+j}} = c_{\kappa(x^i,x^j)}\alpha_{x^ix^j} & \text{if } i+j < n, \\ c_a\beta^{i+j-n} = c_{\kappa(x^i,x^j)}\alpha_{x^ix^j} & \text{if } i+j \geq n \end{cases}$$

and

$$\kappa(x^i,x^j)\kappa(x^{i+j},x^k) = \begin{cases} 1 & \text{if } i+j+k < n, \\ a & \text{if } n \leq i+j+k < 2n, \\ a^2 & \text{if } i+j+k \geq 2n. \end{cases}$$

Since $\alpha_{x^i}(a) = \beta^i(a) = a$, we have $\kappa(x^i,x^j)\kappa(x^{i+j},x^k) = \alpha_{x^i}(\kappa(x^j,x^k))\kappa(x^i,x^{j+k})$ in all cases. Thus (α,κ) is a parameter system. □

Example 4.17.

(i) If β is an inner automorphism in the situation of Theorem 4.16, say $\beta = c_b$ with $b \in N$, then the assumptions hold with $a = b^n$. Hence there always exists an extension such that β is induced by x.

(ii) If N is abelian, the conditions reduce to $\beta^n = 1$. For nonabelian groups this is not sufficient (Exercise 17).

Remark 4.18. Let N be abelian. Then the automorphism systems are precisely the homomorphisms $\alpha\colon H \to \mathrm{Aut}(N)$. The corresponding factor sets form a subgroup $F_\alpha \leq C^2(H,N) := C^1(H \times H, N)$ (for fixed α). Moreover, $(\alpha,\kappa) \sim (\alpha',\kappa')$ implies $\alpha = \alpha'$. Hence one can study the equivalence classes inside F_α.

Theorem 4.19. *Let N be abelian and $\alpha\colon H \to \mathrm{Aut}(N)$ a homomorphism. Then*

$$P_\alpha := \{\kappa \in F_\alpha : (\alpha,\kappa) \sim (\alpha,1)\} \leq F_\alpha.$$

The equivalence classes of parameter systems with automorphism system α correspond to the elements of $\overline{F_\alpha} := F_\alpha/P_\alpha$.

Proof. For $\kappa,\kappa' \in P_\alpha$ there exist $\varphi,\psi \in C^1(H,N)$ with $\kappa(x,y) = \varphi(x)\alpha_x(\varphi(y))\varphi(xy)^{-1}$ and $\kappa'(x,y) = \psi(x)\alpha_x(\psi(y))\psi(xy)^{-1}$ for $x,y \in H$. Since N is abelian, we obtain

$$(\kappa\kappa')(x,y) = (\varphi\psi)(x)\alpha_x((\varphi\psi)(y))(\varphi\psi)(xy)^{-1}$$

and $\kappa\kappa' \in P_\alpha$. One easily obtains $P_\alpha \leq F_\alpha$. Since $C^2(H,N) \cong \times_{x\in H\times H} N$ is abelian, we have $P_\alpha \trianglelefteq F_\alpha$ and $\overline{F_\alpha}$ is well defined. Moreover,

$$\begin{aligned} (\alpha,\kappa) \sim (\alpha,\kappa') &\iff \exists\varphi \in C^1(H,N)\,\forall x,y \in H : \kappa'(x,y) = \varphi(x)\alpha_x(\varphi(y))\kappa(x,y)\varphi(xy)^{-1} \\ &\iff \exists\varphi \in C^1(H,N) : \kappa^{-1}\kappa' \in P_\alpha \iff \kappa P_\alpha = \kappa' P_\alpha. \end{aligned}$$

□

Remark 4.20. Extensions can be computed particularly efficiently on a computer for *polycyclic* groups H. This means that H has a subnormal series with cyclic factors. For finite groups, polycyclic is equivalent to solvable. We further assume that N is an elementary abelian p-group. Via the homomorphism $\alpha\colon H \to \mathrm{Aut}(N)$, N then becomes an $\mathbb{F}_pH$-module.

```
H:=SymmetricGroup(4);; #permutation group
H:=Image(IsomorphismPcGroup(H));; #isomorphic polycyclic group
IrrM:=IrreducibleModules(H,GF(2),2); #simple F_2H-modules of dimension ≤ 2
N:=IrrM[2][2];; #a 2-dimensional simple F_2H-module, so N ≅ C_2^2
ext:=Extensions(H,N); #two extensions of order 96
List(ext,IdGroup); #number in the small groups library
IdGroup(SplitExtension(H,N)); #split extension
Fa:=VectorSpace(GF(2),TwoCocycles(H,N));; #= F_α
Pa:=Subspace(Fa,TwoCoboundaries(H,N));; #= P_α
t:=Difference(Fa,Pa)[1];; #nontrivial factor set
G:=Extension(H,N,t);; #this should be the nonsplit extension
IdGroup(G);
TwoCohomology(H,N); #more information about F̄_α
```

For nonpolycyclic groups one can load the `cohomolo` package (so far only under Linux). As input it expects a permutation group and a finitely presented group isomorphic to it.

```
LoadPackage("cohomolo",false); #false suppresses the banner
H:=GL(5,2);;
gen:=GeneratorsOfGroup(H);;
Hp:=Image(IsomorphismPermGroup(H));;
Hf:=Image(IsomorphismFpGroupByGenerators(H,gen));;
chr:=CHR(Hp,2,Hf,gen); #here N = F_2^5 and α: H → Aut(N) is the natural isomorphism
SecondCohomologyDimension(chr); #=1 means F̄_α ≅ F_2, i.e., there are two extensions
G:=NonsplitExtension(chr); #this is the DEMPWOLFF group
Size(G); #takes forever
```

Theorem 4.21. *Let N be abelian, $K \le H$ finite and $\alpha\colon H \to \mathrm{Aut}(N)$ a homomorphism. Let $\alpha_K\colon K \to \mathrm{Aut}(N)$ be the restriction of α. Then the map*

$$\overline{F_\alpha} \to \overline{F_{\alpha_K}}, \qquad \kappa P_\alpha \mapsto \kappa_{K\times K} P_{\alpha_K}$$

is a homomorphism. If $\kappa_{K\times K} \in P_{\alpha_K}$, then $\kappa^{|H:K|} \in P_\alpha$. In particular, $\exp(\overline{F_\alpha})$ divides $|H|$.

Proof. For $\kappa \in F_\alpha$ clearly $\kappa_{K\times K} \in F_{\alpha_K}$, and the map $F_\alpha \to F_{\alpha_K}$, $\kappa \mapsto \kappa_{K\times K}$ is a homomorphism. From $\kappa \equiv \kappa' \pmod{P_\alpha}$ we obtain $\kappa_{K\times K} \equiv \kappa'_{K\times K} \pmod{P_{\alpha_K}}$. Hence $\overline{F_\alpha} \to \overline{F_{\alpha_K}}$ is a well-defined homomorphism.

Now let $\kappa_{K\times K} \in P_{\alpha_K}$. Let G be an extension of H by N with respect to (α,κ). Then there exists a homomorphism $\rho\colon K \to G$ with $\pi\rho = \mathrm{id}_K$. Let R be a set of coset representatives for H/K with $1 \in R$. For $r \in R$ let $\tilde{r} \in G$ with $\pi(\tilde{r}) = r$ and $\tilde{1} = 1$. For $x \in H$ let $r_x \in R$ with $r_x^{-1}x \in K$. We define $\tilde{x} := \widetilde{r_x}\rho(r_x^{-1}x)$. Then $\pi(\tilde{x}) = r_x r_x^{-1} x = x$ and

$$\kappa'(x,y) := \nu^{-1}(\tilde{x}\tilde{y}\widetilde{xy}^{-1})$$

defines a normalized parameter system equivalent to (α,κ). For $x,y \in H$ we have $r_{xy}K = xyK = xr_yK = r_{xr_y}K$ and $r_{xy} = r_{xr_y}$. It follows that

$$\begin{aligned}\kappa'(x,y) &= \nu^{-1}\big(\widetilde{x}\widetilde{r_y}\rho(r_y^{-1}y)\widetilde{xy}^{-1}\big) = \nu^{-1}\big(\widetilde{x}\widetilde{r_y}\widetilde{xr_y}^{-1}\widetilde{xr_y}\rho(r_y^{-1}y)\widetilde{xy}^{-1}\big)\\ &= \kappa'(x,r_y)\nu^{-1}\big(\widetilde{xr_y}\rho(r_y^{-1}y)\widetilde{xy}^{-1}\big) = \kappa'(x,r_y)\nu^{-1}\big(\widetilde{r_{xy}}\rho(r_{xy}^{-1}xr_y)\rho(r_y^{-1}y)\widetilde{xy}^{-1}\big) = \kappa'(x,r_y).\end{aligned}$$

For $r \in R$ it follows that

$$\kappa'(x,y) = \alpha_x(\kappa'(y,r))\kappa'(x,yr)\kappa'(xy,r)^{-1} = \alpha_x(\kappa'(y,r))\kappa'(x,r_{yr})\kappa'(xy,r)^{-1}.$$

As r runs through R, so does r_{yr}. Let $\varphi(x) := \prod_{r\in R} \kappa'(x,r)$ for $x \in H$. Then

$$\kappa'(x,y)^{|H:K|} = \prod_{r\in R} \alpha_x(\kappa'(y,r))\kappa'(x,r_{yr})\kappa'(xy,r)^{-1} = \alpha_x(\varphi(y))\varphi(x)\varphi(xy)^{-1}.$$

This shows $(\kappa')^{|H:K|} \in P_\alpha$. Since $(\alpha,\kappa) \sim (\alpha,\kappa')$, we also have $\kappa^{|H:K|} \in P_\alpha$ by Theorem 4.19.

The second claim follows by choosing $K = 1$. □

Remark 4.22. The following two theorems were added later in the Group Theory lecture notes with alternative proofs.

Theorem 4.23 (Gaschütz)**.** *Let G be a finite group with abelian normal subgroup N. Let $N \le H \le G$ with $\gcd(|N|, |G:H|) = 1$ such that N has a complement in H. Then N has a complement in G.*

Proof. As is well known, G is an extension of G/N by N. Let (α,κ) be the corresponding parameter system. By assumption, the extension of $K := H/N$ by N splits. Hence $\kappa_{K\times K} \in P_{\alpha_K}$. By Theorem 4.21 we obtain $\kappa^{|G:H|} \in P_\alpha$. Since $F_\alpha \le C^2(G/N, N)$, we also have $\kappa^{|N|} \in P_\alpha$. From $\gcd(|G:H|, |N|) = 1$ we obtain $\kappa \in P_\alpha$, i. e., G splits. □

Theorem 4.24 (Gaschütz)**.** *Let G be a finite group with abelian normal subgroup N. Then N has a complement in G if and only if N has a complement in P for every Sylow subgroup P/N of G/N.*

Proof. If K is a complement of N in G, then $K \cap P$ is a complement of N in P for $N \le P \le G$, since $N(K \cap P) = NK \cap P = P$ and $N \cap (K \cap P) \le N \cap K = 1$.

Conversely, suppose now that N has a complement in P for all Sylow subgroups P/N of G/N. Let (α,κ) be the parameter system of the extension G. By Theorem 4.21 we have $\kappa^{|G:P|} \in P_\alpha$. This shows $\kappa^{|G/N|_{p'}} \in P_\alpha$ for all prime divisors p of $|G/N|$. Since the numbers $|G/N|_{p'}$ are coprime (where p runs over all prime divisors of $|G|$), we obtain $\kappa \in P_\alpha$, i. e., G splits. □

Definition 4.25. Now let N be arbitrary and (α,κ) a parameter system.

- The map $\omega\colon H \to \mathrm{Out}(N)$, $x \mapsto \alpha_x\mathrm{Inn}(N)$ is a homomorphism, called the *coupling* of (α,κ). Let $\mathrm{Par}(\omega)$ be the set of parameter systems with coupling ω. Equivalent parameter systems define the same coupling. Let $\overline{\mathrm{Par}(\omega)}$ be the corresponding set of equivalence classes.
- Let $Z := \mathrm{Z}(N)$. Every homomorphism $\omega\colon H \to \mathrm{Out}(N)$, $x \mapsto \omega_x\mathrm{Inn}(N)$ defines a well-defined homomorphism $\omega_Z\colon H \to \mathrm{Aut}(Z)$, $x \mapsto (\omega_x)_{|Z}$.

Remark 4.26. For a homomorphism $\omega\colon H \to \mathrm{Out}(N)$, a corresponding parameter system need not exist (Exercise 17).

Theorem 4.27. *Let $\omega\colon H \to \mathrm{Out}(N)$ be a homomorphism with $\mathrm{Par}(\omega) \neq \varnothing$. Let $Z := \mathrm{Z}(N)$ and $\beta := \omega_Z$. Then $\overline{F_\beta}$ acts regularly on $\overline{\mathrm{Par}(\omega)}$. In particular, $\overline{\mathrm{Par}(\omega)}$ and $\overline{F_\beta}$ have the same cardinality.*

Proof. Let $\lambda \in F_\beta$ and $(\alpha,\kappa) \in \mathrm{Par}(\omega)$. Since $\lambda(H\times H) \le Z$, we have

$$\begin{aligned}\alpha_x\alpha_y &= c_{\kappa(x,y)}\alpha_{xy} = c_{(\lambda\kappa)(x,y)}\alpha_{xy},\\ (\lambda\kappa)(x,y)(\lambda\kappa)(xy,z) &= \alpha_x(\kappa(y,z))\beta_x(\lambda(y,z))(\lambda\kappa)(x,yz) = \alpha_x((\lambda\kappa)(y,z))(\lambda\kappa)(x,yz)\end{aligned}$$

for $x,y,z \in H$. Hence $(\alpha,\lambda\kappa) \in \mathrm{Par}(\omega)$. One easily sees that ${}^\lambda(\alpha,\kappa) := (\alpha,\lambda\kappa)$ defines an action of F_β on $\mathrm{Par}(\omega)$. Let $(\alpha',\kappa') \sim (\alpha,\kappa)$. Then there exists $\varphi \in C^1(H,N)$ with

$$\begin{aligned}\alpha'_x &= c_{\varphi(x)}\alpha_x,\\ (\lambda\kappa')(x,y) &= \varphi(x)\alpha_x(\varphi(y))(\lambda\kappa)(x,y)\varphi(xy)^{-1}.\end{aligned}$$

Hence ${}^\lambda(\alpha,\kappa) \sim {}^\lambda(\alpha',\kappa')$ and F_β acts on $\overline{\mathrm{Par}(\omega)}$. Let $\lambda \in P_\beta$, i. e., $\lambda(x,y) = \delta(x)\beta_x(\delta(y))\delta(xy)^{-1}$ for some $\delta \in C^1(H,Z)$. Then

$$(\lambda\kappa)(x,y) = \delta(x)\alpha_x(\delta(y))\kappa(x,y)\delta(xy)^{-1},$$

i. e., ${}^\lambda(\alpha,\kappa) \sim (\alpha,\kappa)$. Thus P_β acts trivially on $\overline{\mathrm{Par}(\omega)}$, and we obtain a well-defined action of $\overline{F_\beta}$ on $\overline{\mathrm{Par}(\omega)}$. Let $\lambda \in F_\beta$ with ${}^\lambda(\alpha,\kappa) \sim (\alpha,\kappa)$. Then there exists $\varphi \in C^1(H,N)$ with

$$\begin{aligned}\alpha_x &= c_{\varphi(x)}\alpha_x,\\ (\lambda\kappa)(x,y) &= \varphi(x)\alpha_x(\varphi(y))\kappa(x,y)\varphi(xy)^{-1}\end{aligned}$$

for $x,y \in H$. It follows that $\varphi(H) \le Z$ and $\lambda(x,y) = \varphi(x)\beta_x(\varphi(y))\varphi(xy)^{-1}$. This shows $\lambda \in P_\beta$. Hence $\overline{F_\beta}$ acts fixed-point-freely on $\overline{\mathrm{Par}(\omega)}$. For transitivity, let $(\alpha,\kappa),(\alpha',\kappa') \in \mathrm{Par}(\omega)$. Since $\alpha_x\mathrm{Inn}(N) = \omega(x) = \alpha'_x\mathrm{Inn}(N)$ for $x \in H$, there exists $\varphi \in C^1(H,N)$ with $\alpha_x = c_{\varphi(x)}\alpha'_x$ for $x \in H$. Hence we can replace (α',κ') by an equivalent parameter system of the form (α,κ'). Now

$$c_{\kappa(x,y)}\alpha_{xy} = \alpha_x\alpha_y = c_{\kappa'(x,y)}\alpha_{xy}$$

and $\lambda := \kappa^{-1}\kappa' \colon H\times H \to Z$. Because of

$$\begin{aligned}\lambda(x,y)\lambda(xy,z) &= \kappa(xy,z)^{-1}\lambda(x,y)\kappa'(xy,z) = \kappa(xy,z)^{-1}\kappa(x,y)^{-1}\kappa'(x,y)\kappa'(xy,z)\\ &= \big(\alpha_x(\kappa(y,z))\kappa(x,yz)\big)^{-1}\alpha_x(\kappa'(y,z))\kappa'(x,yz)\\ &= \kappa(x,yz)^{-1}\alpha_x\big(\kappa(y,z)^{-1}\kappa'(y,z)\big)\kappa'(x,yz) = \kappa(x,yz)^{-1}\beta_x(\lambda(y,z))\kappa'(x,yz)\\ &= \beta_x(\lambda(y,z))\kappa(x,yz)^{-1}\kappa'(x,yz) = \beta_x(\lambda(y,z))\lambda(x,yz)\end{aligned}$$

for $x,y \in H$ we have $\lambda \in F_\beta$ and ${}^\lambda(\alpha,\kappa) = (\alpha,\kappa\lambda) = (\alpha,\kappa')$. Hence $\overline{F_\beta}$ acts regularly on $\overline{\mathrm{Par}(\omega)}$. □

Corollary 4.28. *Let* $\mathrm{Z}(N) = 1$. *Then:*

(i) For every homomorphism $\omega \colon H \to \mathrm{Out}(N)$ *we have* $|\overline{\mathrm{Par}(\omega)}| = 1$.

(ii) If $\mathrm{Inn}(N) \cong N$ *has a complement in* $\mathrm{Aut}(N)$, *then every extension by* N *splits.*

Proof.

(i) Let $\omega(x) = \alpha_x\mathrm{Inn}(N)$ for $x \in H$. Since $\alpha_{xy}\mathrm{Inn}(N) = \omega(xy) = \omega(x)\omega(y) = \alpha_x\alpha_y\mathrm{Inn}(N)$, there exist $\kappa(x,y) \in N$ with $\alpha_x\alpha_y = c_{\kappa(x,y)}\alpha_{xy}$ for all $x,y \in H$. Here

$$\begin{aligned}c_{\kappa(x,y)\kappa(xy,z)}\alpha_{xyz} &= c_{\kappa(x,y)}\alpha_{xy}\alpha_z = \alpha_x\alpha_y\alpha_z = \alpha_x c_{\kappa(y,z)}\alpha_{yz}\\ &= c_{\alpha_x(\kappa(y,z))}\alpha_x\alpha_{yz} = c_{\alpha_x(\kappa(y,z))\kappa(x,yz)}\alpha_{xyz}.\end{aligned}$$

From $\mathrm{Inn}(N) \cong N/\mathrm{Z}(N) \cong N$ we obtain $\kappa(x,y)\kappa(xy,z) = \alpha_x(\kappa(y,z))\kappa(x,yz)$. This shows $(\alpha,\kappa) \in \mathrm{Par}(\omega) \ne \varnothing$ and $|\overline{\mathrm{Par}(\omega)}| = 1$ by Theorem 4.27.

(ii) By assumption, there exists a homomorphism $\tau\colon \mathrm{Out}(N) \to \mathrm{Aut}(N)$ with $\tau(\gamma)\mathrm{Inn}(N) = \gamma$ for all $\gamma \in \mathrm{Out}(N)$. Hence every homomorphism $\omega\colon H \to \mathrm{Out}(N)$ can be lifted to a homomorphism $\alpha := \tau\omega\colon H \to \mathrm{Aut}(N)$ (i. e., $\omega(x) = \alpha_x\mathrm{Inn}(N)$ for all $x \in H$). By (i), the split extension with respect to $(\alpha, 1)$ is the only extension with coupling ω. □

Example 4.29. If N is *complete*, i. e., $\mathrm{Z}(N) = 1 = \mathrm{Out}(N)$, then the only extensions are $N \times H$. This was already shown in GT, Exercise 26.

Remark 4.30. The number of extensions with a nonabelian group N for a given coupling $\omega\colon K \to \mathrm{Out}(N)$ can be determined with the `hap` package:

```
LoadPackage("hap",false); #loads further packages
N:=QuaternionGroup(8);;
H:=AutomorphismGroup(N);;
NH:=SemidirectProduct(H,N);;
omega:=GOuterGroup(NH,Image(Embedding(NH,2))); #defines coupling H → Out(N)
beta:=Center(omega); #associated factor set H → Aut(Z(N))
A:=ActingGroup(beta);; #preimage of β
R:=ResolutionFiniteGroup(A,3);; #three terms of a resolution
C:=HomToGModule(R,beta);; #associated chain complex
Cohomology(C,2); #orders of the cyclic factors of F_β
```

Theorem 4.31 (Johnson-Zassenhaus). *Two extensions $N \xrightarrow{\nu_i} G_i \xrightarrow{\pi_i} H$ $(i = 1, 2)$ are equivalent if and only if for every Sylow subgroup P of H the extensions $\pi_1^{-1}(P)$ and $\pi_2^{-1}(P)$ of P by N are equivalent.*

Proof. Let G_1 and G_2 be equivalent via $\gamma\colon G_1 \to G_2$. For a prime p and $P \in \mathrm{Syl}_p(H)$ let $P_i := \pi_i^{-1}(P) \leq G_i$ and $\delta := \gamma_{P_1}$. Then

$$\delta(P_1) = \gamma(\pi_1^{-1}(P)) = \gamma\big(\gamma^{-1}(\pi_2^{-1}(P))\big) = P_2,$$

$\delta\nu_1 = \nu_2$ and $(\pi_2)_{|P_2}\delta = (\pi_1)_{|P_1}$. Hence P_1 and P_2 are equivalent.

Now suppose that P_1 and P_2 are equivalent for all Sylow subgroups P of H. Let (α, κ) and (α', κ') be the parameter systems of G_1 and G_2, respectively. Then $(\alpha_P, \kappa_{P\times P})$ and $(\alpha'_P, \kappa'_{P\times P})$ are equivalent, and we obtain $\alpha_x\mathrm{Inn}(N) = \alpha'_x\mathrm{Inn}(N)$ for all $x \in P$. Since H is generated by its Sylow subgroups, (α, κ) and (α', κ') have the same coupling ω. Let $Z := \mathrm{Z}(N)$ and $\beta := \omega_Z$. By Theorem 4.27 there exists $\lambda \in F_\beta$ with $(\alpha', \kappa') \sim (\alpha, \lambda\kappa)$. By assumption, $\lambda_P \in P_{\beta_P}$ for every Sylow subgroup P of H, i. e., the extension of P by Z splits. With Theorem 4.24, the extension of H by Z now splits as well, i. e., $\lambda \in P_\beta$. This shows $(\alpha', \kappa') \sim (\alpha, \lambda\kappa) \sim (\alpha, \kappa)$. □

Remark 4.32. Theorem 4.31 is not suited to prove Theorem 4.24 of Gaschütz for nonabelian N, since for a homomorphism $\omega\colon H \to \mathrm{Out}(N)$ there need not exist a split extension to compare with (Exercise 17). Even if such a split extension of H by N exists, the restricted extension of H by $\mathrm{Z}(N)$ need not split. For example, $G = D_8 * C_4$ is a split extension of $H = C_2$ by $N = D_8$, but the restricted extension $\mathrm{C}_G(N) = C_4$ of H by $\mathrm{Z}(N) \cong C_2$ does not split. A similar example shows that Theorem 4.24 is false for nonabelian groups N in general (Exercise 15).

5 Central Extensions

Definition 5.1. If N is abelian and $\alpha\colon H \to \mathrm{Aut}(N)$ is trivial, then

$$\begin{aligned} Z^2(H,N) &:= F_\alpha = \big\{\kappa \in C^2(H,N) : \kappa(x,y)\kappa(xy,z) = \kappa(y,z)\kappa(x,yz)\big\},\\ B^2(H,N) &:= P_\alpha = \big\{\kappa \in Z^2(H,N) : \exists \varphi \in C^1(H,N) : \kappa(x,y) = \varphi(x)\varphi(y)\varphi(xy)^{-1}\big\},\\ H^2(H,N) &:= \overline{F_\alpha} = F_\alpha/P_\alpha. \end{aligned}$$

The elements of $Z^2(H,N)$ are called (2-)*cocycles*. For $\varphi \in C^1(H,N)$ let $\partial\varphi \in B^2(H,N)$ be given by $\partial\varphi(x,y) := \varphi(x)\varphi(y)\varphi(xy)^{-1}$ for $x,y \in H$. $H^2(H,N)$ is called the second *cohomology group* of H with values in N. If G is a corresponding extension, then $\nu(N) \leq \mathrm{Z}(G)$ (cf. Remark 4.8). Hence one speaks of *central* extensions.

Remark 5.2.

(i) In the following, A always denotes an abelian group.

(ii) Let H be abelian and G the central extension on $A \times H$ by means of $\kappa \in Z^2(H,A)$, i. e.,

$$(a,x)(b,y) = (ab\kappa(x,y), xy)$$

for all $(a,x),(b,y) \in G$. Then G is abelian if and only if κ is symmetric, i. e., $\kappa(x,y) = \kappa(y,x)$ for all $x,y \in H$. Since H is abelian, the cocycles in $B^2(H,A)$ are symmetric.

Definition 5.3. For abelian groups H let

$$Z_s^2(H,A) := \big\{\kappa \in Z^2(H,A) : \forall x,y \in H : \kappa(x,y) = \kappa(y,x)\big\}$$

and $H_s^2(H,A) := Z_s^2(H,A)/B^2(H,A)$.

Theorem 5.4. *For $n \in \mathbb{N}$ we have $H^2(C_n,A) = H_s^2(C_n,A) \cong A/\langle a^n : a \in A\rangle$.*

Proof. Let G be a central extension of $H = \langle x\rangle \cong C_n$ by A. Since $A \leq \mathrm{Z}(G)$, the group $G/\mathrm{Z}(G)$ is cyclic and G is abelian. This shows $H^2(H,A) = H_s^2(H,A)$. For $a \in A$ we have constructed in the proof of Theorem 4.16 a cocycle $\kappa_a \in Z^2(H,A)$ with

$$\kappa_a(x^i,x^j) = \begin{cases} 1 & \text{if } i+j < n,\\ a & \text{if } i+j \geq n \end{cases}$$

The map $F\colon A \to H^2(H,A)$, $a \mapsto \kappa_a$ is clearly a homomorphism. Let $\varphi\colon H \to A$, $x^i \mapsto a^i$ for $i = 0,\dots,n-1$. Then we have

$$\partial\varphi(x^i,x^j) = \begin{cases} 1 & \text{if } i+j < n,\\ a^n & \text{if } i+j \geq n. \end{cases}$$

This shows $a^n \in \mathrm{Ker}(F)$. For $a \in \mathrm{Ker}(F)$ there exists a $\varphi \in C^1(H,A)$ with $\kappa_a = \partial\varphi$. We have

$$\begin{aligned} \varphi(1) &= \partial\varphi(1,1) = \kappa_a(1,1) = 1,\\ \varphi(x)^2 &= \kappa_a(x,x)\varphi(x^2) = \varphi(x^2),\\ \varphi(x)^3 &= \kappa_a(x^2,x)\varphi(x^3) = \varphi(x^3),\\ &\vdots\\ \varphi(x)^n &= \kappa_a(x^{n-1},x)\varphi(1) = a. \end{aligned}$$

Hence $a \in \langle b^n : b \in A\rangle$ and $\mathrm{Ker}(F) = \langle a^n : a \in A\rangle$.

For the surjectivity of F, let $A \to G \xrightarrow{\pi} H$ be an extension of H by A. Let $\tilde{x} \in G$ with $\pi(\tilde{x}) = x$. For $i = 0, \ldots, n-1$ we can choose $\widetilde{x^i} := \tilde{x}^i$ as a preimage of x^i. For the cocycle κ we then have

$$\kappa(x^i, x^j) = \begin{cases} \tilde{x}^i \tilde{x}^j \widetilde{x^{i+j}}^{-1} = 1 & \text{if } i+j < n, \\ \tilde{x}^i \tilde{x}^j \widetilde{x^{i+j-n}}^{-1} = \tilde{x}^n & \text{if } i+j \geq n. \end{cases}$$

With $a := \tilde{x}^n \in A$ we thus have $\kappa = \kappa_a$. Hence F is surjective. □

Theorem 5.5. *For all abelian groups G and H we have $H^2_s(G \times H, A) \cong H^2_s(G, A) \times H^2_s(H, A)$.*

Proof. We follow the proof of the Künneth formula from group theory. We regard G and H as subgroups of $G \times H$. For $\kappa \in Z^2_s(G \times H, A)$ let $\kappa_G \in Z^2_s(G, A)$ be the restriction of κ to $G \times G$ and analogously $\kappa_H \in Z^2_s(H, A)$. Then

$$F\colon Z^2(G \times H, A) \to Z^2(G, A) \times Z^2(H, A), \qquad \kappa \mapsto (\kappa_G, \kappa_H)$$

is a homomorphism. For $\varphi \in C^1(G \times H, A)$ we clearly have $(\partial\varphi)_G = \partial\varphi_G \in B^2(G, A)$ and $(\partial\varphi)_H \in B^2(H, A)$. Thus F induces a homomorphism $\overline{F}\colon H^2_s(G \times H, A) \to H^2_s(G, A) \times H^2_s(H, A)$.

For the surjectivity of $\overline{F}$, let $\kappa_1 \in Z^2_s(G, A)$ and $\kappa_2 \in Z^2_s(H, A)$ be normalized. For $x_i \in G$ and $y_i \in H$ let $\kappa(x_1y_1, x_2y_2) := \kappa_1(x_1, x_2)\kappa_2(y_1, y_2)$. Then

$$\begin{aligned} \kappa(x_1y_1, x_2y_2)\kappa(x_1x_2y_1y_2, x_3y_3) &= \kappa_1(x_1, x_2)\kappa_2(y_1, y_2)\kappa_1(x_1x_2, x_3)\kappa_2(y_1y_2, y_3) \\ &= \kappa_1(x_2, x_3)\kappa_1(x_1, x_2x_3)\kappa_2(y_2, y_3)\kappa_2(y_1, y_2y_3) \\ &= \kappa(x_2y_2, x_3y_3)\kappa(x_1y_1, x_2y_2x_3y_3). \end{aligned}$$

This shows $\kappa \in Z^2_s(G \times H, A)$ with $\kappa_G = \kappa_1$ and $\kappa_H = \kappa_2$. Hence $\overline{F}$ is surjective.

For the injectivity, let $F(\kappa) = (\partial\varphi_1, \partial\varphi_2)$ with $\varphi_1 \in C^1(G, A)$ and $\varphi_2 \in C^1(H, A)$. Let $\varphi \in C^1(G \times H, A)$ with $\varphi(xy) := \varphi_1(x)\varphi_2(y)\kappa(x, y)^{-1}$ for $x \in G$ and $y \in H$. Then we have

$$\begin{aligned} \partial\varphi(x_1y_1, x_2y_2) &= \varphi(x_1y_1)\varphi(x_2y_2)\varphi(x_1x_2y_1y_2)^{-1} \\ &= \varphi_1(x_1)\varphi_2(y_1)\kappa(x_1, y_1)^{-1}\varphi_1(x_2)\varphi_2(y_2)\kappa(x_2, y_2)^{-1}\varphi_1(x_1x_2)^{-1}\varphi_2(y_1y_2)^{-1}\kappa(x_1x_2, y_1y_2) \\ &= \kappa(x_1, x_2)\kappa(y_1, y_2)\kappa(x_1, y_1)^{-1}\kappa(x_2, y_2)^{-1}\kappa(x_1x_2, y_1y_2) \\ &= \kappa(x_2, y_1)\kappa(x_1, x_2y_1)\kappa(x_1x_2, y_1)^{-1}\kappa(y_1, y_2)\kappa(x_1, y_1)^{-1}\kappa(x_2, y_2)^{-1}\kappa(x_1x_2, y_1y_2) \\ &= \kappa(x_1y_1, x_2)\kappa(x_1x_2y_1, y_2)\kappa(x_2, y_2)^{-1} = \kappa(x_1y_1, x_2y_2) \end{aligned}$$

for $x_i \in G$ and $y_i \in H$. Hence $\kappa = \partial\varphi \in B^2(G \times H, A)$ and $\overline{F}$ is an isomorphism. □

Corollary 5.6. *For every finite abelian group A we have $H^2_s(A, \mathbb{C}^\times) = 1$.*

Proof. Since every complex number has an n-th root, we have $\langle z^n : z \in \mathbb{C}^\times\rangle = \mathbb{C}^\times$. The claim follows from Theorem 5.5 and Theorem 5.4. □

Corollary 5.7. *For finite abelian groups $G \cong C_{d_1} \times \ldots \times C_{d_k}$ and $A \cong C_{e_1} \times \ldots \times C_{e_l}$ we have*

$$H^2_s(G, A) \cong \mathrm{Hom}(G, A) \cong \bigtimes_{\substack{1 \leq i \leq k \\ 1 \leq j \leq l}} C_{\gcd(d_i, e_j)}.$$

Proof. For $d \in \mathbb{N}$ we have $H^2(C_d, A) \cong A/\langle a^d : a \in A\rangle \cong \bigtimes_{i=1}^{l} C_{\gcd(d,e_i)}$ by Theorem 5.4. Theorem 5.5 yields $H_s^2(G, A) \cong \bigtimes_{i,j} C_{\gcd(d_i,e_j)}$. Let $G = \langle x_1\rangle \times \ldots \times \langle x_k\rangle \cong C_{d_1} \times \ldots \times C_{d_k}$. Then $f \in \mathrm{Hom}(G, A)$ is uniquely determined by

$$f(x_i) \in \langle a \in A : a^{d_i} = 1\rangle \cong \bigtimes_{j=1}^{l} C_{\gcd(d_i,e_j)} \qquad (i = 1, \ldots, k)$$

and every choice defines a homomorphism $G \to A$. Clearly, the map $\mathrm{Hom}(G, A) \to \bigtimes_{i,j} C_{\gcd(d_i,e_j)}$, $f \mapsto (f(x_1), \ldots, f(x_k))$ is an isomorphism. The claim follows. □

Example 5.8. By Corollary 5.7 there are four equivalence classes of abelian extensions of C_4 by C_4. However, there are only two isomorphism types of such groups: C_4^2 and C_{16}. In Theorem 5.19 we determine the structure of $H^2(G, A)$ for every finite group G.

Remark 5.9. For coprime $d_1, d_2 \in \mathbb{N}$ and $e \in \mathbb{N}$ we have $\gcd(d_1d_2, e) = \gcd(d_1, e)\gcd(d_2, e)$. Hence the right-hand side in Corollary 5.7 does not depend on the decomposition of G and A. Since the expression is symmetric, we have $H_s^2(A, B) \cong H_s^2(B, A)$ and $\mathrm{Hom}(A, B) \cong \mathrm{Hom}(B, A)$ for all finite abelian groups A and B.

Definition 5.10. An extension G of H by Z is called a *Schur extension*[11] if $Z \leq \mathrm{Z}(G) \cap G'$ holds. The *Schur multiplier* of H is $M(H) := H^2(H, \mathbb{C}^\times)$.

Example 5.11.

(i) Let G be a nonabelian finite nilpotent group. By GT, Theorem 3.14 we have $G' \cap \mathrm{Z}(G) \neq 1$. Hence G is a Schur extension of a smaller group. In particular, D_8 and Q_8 are Schur extensions of C_2^2. Moreover, $G/G^{[3]}$ is a Schur extension of G/G'.

(ii) Every quasisimple group is a Schur extension of a simple group.

Remark 5.12.

(i) In the following we study the possible Schur extensions of a fixed group H, which we rename G for this purpose.

(ii) Let $\widehat{G}$ be a Schur extension of G with $1 \neq Z \leq \widehat{G}' \cap \mathrm{Z}(\widehat{G})$ and $\widehat{G}/Z \cong G$. If Z has a complement K in $\widehat{G}$, then $\widehat{G} = Z \times K$ and $Z \nleq 1 \times K' = \widehat{G}'$. Hence proper Schur extensions are nonsplit.

(iii) If $n := |G| < \infty$, then $\exp(M(G)) \mid n$ by Theorem 4.21. Hence the values of $\kappa \in M(G)$ are n-th roots of unity. This shows $|M(G)| \leq |Z^2(G, \mathbb{C}^\times)| \leq n^{n^2} < \infty$. In the following, G is always finite and A abelian.

(iv) By Theorem 5.4 we have $M(C_n) = 1$ for all $n \in \mathbb{N}$.

(v) By GT, Lemma 11.14 we have $A^* := \mathrm{Hom}(A, \mathbb{C}^\times) \cong A$, if $|A| < \infty$.

Lemma 5.13. *The map*

$$\Phi\colon H_s^2(G/G', A) \to H^2(G, A), \qquad \kappa B^2(G/G', A) \mapsto \Phi_\kappa B^2(G, A)$$

with $\Phi_\kappa(x, y) = \kappa(xG', yG')$ *for* $x, y \in G$ *is a monomorphism.*

[11] In the literature, such extensions are often called *stem extensions*.

Proof. Clearly, $Z^2(G/G',A) \to Z^2(G,A)$, $\kappa \mapsto \Phi_\kappa$ is a homomorphism. For $\varphi \in C^1(G/G',A)$ we have $\Phi_{\partial\varphi} \in B^2(G,A)$. Hence Φ is well defined on $H^2(G/G',A)$. Let $\kappa \in Z^2_s(G/G',A)$ and $\varphi \in C^1(G,A)$ with $\Phi_\kappa = \partial\varphi$. Let $R \subseteq G$ be a set of representatives for G/G'. For $x \in G$ let $r_x \in R$ with $xG' = r_xG'$. Let $\overline{\varphi} \in C^1(G/G',A)$ with $\overline{\varphi}(xG') := \varphi(r_x)$ for $x \in G$. Then we have

$$\kappa(xG',yG') = \kappa(r_xG',r_yG') = \Phi_\kappa(r_x,r_y) = \partial\varphi(r_x,r_y) = \partial\overline{\varphi}(xG',yG')$$

for $x,y \in G$. Hence Φ is injective. □

Lemma 5.14. *The map*

$$\Psi\colon H^2(G,A) \to \mathrm{Hom}(A^*,M(G)), \qquad \kappa B^2(G,A) \mapsto \Psi_\kappa$$

with $\Psi_\kappa(\lambda) = (\lambda\circ\kappa)B^2(G,\mathbb{C}^\times)$ *is a homomorphism.*

Proof. For $\lambda,\mu \in A^*$ we clearly have $\lambda\circ\kappa \in Z^2(G,\mathbb{C}^\times)$ and $(\lambda\mu)\circ\kappa = (\lambda\circ\kappa)(\mu\circ\kappa)$. Hence $\Psi_\kappa \in \mathrm{Hom}(A^*,M(G))$. For $\varphi \in C^1(G,A)$ and $\lambda \in A^*$ we have

$$\Psi_{\partial\varphi}(\lambda) = (\lambda\circ\partial\varphi)B^2(G,\mathbb{C}^\times) = \partial(\lambda\circ\varphi)B^2(G,\mathbb{C}^\times) = 1.$$

Hence Ψ is well defined. Since $\lambda\circ(\kappa_1\kappa_2) = (\lambda\circ\kappa_1)(\lambda\circ\kappa_2)$, Ψ is a homomorphism. □

Theorem 5.15 (Schur)**.** *Let $\widehat{G}$ be a Schur extension of G with $\widehat{G}/Z \cong G$. Then Z is isomorphic to a subgroup of $M(G)$. In particular, $|\widehat{G}| \le |G||M(G)|$ and G has only finitely many Schur extensions up to isomorphism.*

Proof. For $x \in G$ let $\widehat{x} \in \widehat{G}$ with $\widehat{x}Z = x$ and $\widehat{1} = 1$. Let $\kappa \in Z^2(G,Z)$ be the corresponding normalized factor set of $\widehat{G}$ (Lemma 4.5). We have $\kappa(x,y) = \widehat{xy}\widehat{x}\widehat{y}^{-1}$ for $x,y \in G$. Since $Z^* \cong Z$, it suffices to show that the map Ψ_κ from Lemma 5.14 is injective. So let $\lambda \in Z^*$ with $\lambda\circ\kappa = \partial\varphi$ for some $\varphi \in C^1(G,\mathbb{C}^\times)$. Then we have

$$\varphi(1) = \varphi(1)\varphi(1)\varphi(1)^{-1} = \partial\varphi(1) = \lambda(\kappa(1,1)) = \lambda(1) = 1.$$

Let $\widehat{\lambda}\colon \widehat{G} \to \mathbb{C}^\times$ with $\widehat{\lambda}(\widehat{x}a) := \varphi(x)\lambda(a)$ for $x \in G$ and $a \in Z$. Since $\widehat{\lambda}(a) = \widehat{\lambda}(\widehat{1}a) = \lambda(a)$, $\widehat{\lambda}$ is an extension of λ. For $x,y \in G$ and $a,b \in Z$ we have

$$\begin{aligned}\widehat{\lambda}(\widehat{x}a\cdot\widehat{y}b) &= \widehat{\lambda}(\widehat{xy}ab) = \widehat{\lambda}(\kappa(x,y)\widehat{xy}ab) = \lambda(\kappa(x,y))\varphi(xy)\lambda(a)\lambda(b)\\ &= \varphi(x)\varphi(y)\varphi(xy)^{-1}\varphi(xy)\lambda(a)\lambda(b) = \varphi(x)\lambda(a)\varphi(y)\lambda(b) = \widehat{\lambda}(\widehat{x}a)\widehat{\lambda}(\widehat{y}b).\end{aligned}$$

Hence $\widehat{\lambda}$ is a homomorphism with $\widehat{G}/\mathrm{Ker}(\widehat{\lambda}) \le \mathbb{C}^\times$. It follows that $Z \le \widehat{G}' \le \mathrm{Ker}(\widehat{\lambda})$. This shows $\lambda = 1$. □

Definition 5.16. A Schur extension $\widehat{G}$ of G is called *maximal* if $|\widehat{G}| = |G||M(G)|$.

Theorem 5.17 (Schur)**.** *Every finite group G has a maximal Schur extension.*

Proof. By Remark 5.12, $M(G) = \langle\overline{\kappa_1}\rangle \oplus \ldots \oplus \langle\overline{\kappa_n}\rangle$ is finite. Let $d_i := |\langle\overline{\kappa_i}\rangle|$ and $A_i \leq \mathbb{C}^\times$ with $|A_i| = d_i$ for $i = 1, \ldots, n$. Let $\kappa_i \in Z^2(G, \mathbb{C}^\times)$ with $\kappa_i B^2(G, \mathbb{C}^\times) = \overline{\kappa_i}$. Then $\kappa_i^{d_i} = \partial\gamma_i$ for some $\gamma_i \in C^1(G, \mathbb{C}^\times)$. Let $\delta_i(x) \in \mathbb{C}^\times$ with $\delta_i(x)^{d_i} = \gamma_i(x)^{-1}$ for $x \in G$. After replacing κ_i by $\kappa_i\partial\delta_i$, we have $\kappa_i^{d_i} = 1$ for $i = 1, \ldots, n$. In particular, $\kappa_i \in Z^2(G, A_i)$ for $i = 1, \ldots, n$. By Lemma 4.11 we may also assume $\kappa_i(x, 1) = \kappa_i(1, x) = 1$ for $x \in G$. Let $A := A_1 \times \ldots \times A_n \cong M(G)$ and $\kappa \in C^2(G, A)$ with $\kappa(x, y) = (\kappa_1(x, y), \ldots, \kappa_n(x, y))$ for $x, y \in G$. Clearly, we then have $\kappa \in Z^2(G, A)$ with $\kappa(1, x) = \kappa(x, 1) = 1$ for $x \in G$.

Let $\widehat{G} = A \times G$ be the central extension of G by A with respect to κ. We regard A as a subgroup of $\widehat{G}$ via $a \mapsto (a, 1)$. We choose preimages $\widehat{x} \in \widehat{G}$ of $x \in G$ such that $\kappa(x, y) = \widehat{x}\widehat{y}\widehat{xy}^{-1}$ holds for all $x, y \in G$. Let $\pi_i \colon A \to A_i \leq \mathbb{C}^\times$ be the i-th projection. With the map Ψ_κ from Lemma 5.14 we then have

$$\Psi_\kappa(\pi_i) = (\pi_i \circ \kappa)B^2(G, \mathbb{C}^\times) = \overline{\kappa_i}$$

for $i = 1, \ldots, n$. Since $M(G) = \langle\overline{\kappa_1}, \ldots, \overline{\kappa_n}\rangle$, Ψ_κ is surjective. By Remark 5.12 we have $A^* \cong A \cong M(G)$. Hence Ψ_κ is also injective. By the fundamental theorem of finite abelian groups (applied to $\widehat{G}/\widehat{G}'$) there exist normal subgroups $N_1, \ldots, N_s \trianglelefteq \widehat{G}$ with $\widehat{G}' = N_1 \cap \ldots \cap N_s$ such that $\widehat{G}/N_i$ is cyclic for $i = 1, \ldots, s$.

Suppose that $A \not\leq \widehat{G}'$. Then there exists an i with $A \not\leq N_i$. Embedding $\widehat{G}/N_i$ into $\mathbb{C}^\times$, we obtain a homomorphism $\lambda \colon \widehat{G} \to \widehat{G}/N_i \to \mathbb{C}^\times$ with $\lambda(A) \neq 1$. Hence the restriction λ_A is a nontrivial element of A^*. For $x \in G$ we set $\varphi(x) := \lambda(\widehat{x})$. Then we have

$$\Psi_\kappa(\lambda_A)(x, y) = \lambda(\kappa(x, y)) = \lambda(\widehat{x}\widehat{y}\widehat{xy}^{-1}) = \varphi(x)\varphi(y)\varphi(xy)^{-1} = \partial\varphi(x, y)$$

for $x, y \in G$. This yields $\Psi_\kappa(\lambda_A) = 1$, contradicting the injectivity of Ψ_κ. Hence $A \leq \widehat{G}'$ and $\widehat{G}$ is a Schur extension of G. □

Remark 5.18. In GAP there are several ways to determine the Schur multiplier and a maximal Schur extension:

```
G:=AlternatingGroup(7);;
AbelianInvariantsMultiplier(G); #orders of the cyclic factors of M(G)
S:=SchurCover(G); #maximal Schur extension
S:=Image(IsomorphismPermGroup(S));; #more efficient representation as a permutation group
NrMovedPoints(S); #degree of the permutation group
S:=Image(SmallerDegreePermutationRepresentation(S));; #even more efficient representation
NrMovedPoints(S);
StructureDescription(S); #= C6.A7

epi:=EpimorphismSchurCover(G,[3]); #epimorphism Ĝ → G with kernel O3(M(G))
S:=Source(epi); #nonmaximal Schur extension Ĝ
M:=Kernel(epi); #O3(M(G))

P:=SmallGroup(256,111);;
SchurCovers(P); #all maximal Schur extensions, only for p-groups

LoadPackage("cohomolo",false);
G:=PSL(3,4);;
for p in PrimeDivisors(Size(G)) do
  chr:=CHR(G,p);;
  Print(SchurMultiplier(chr)); #= Op(M(G)) faster than AbelianInvariantsMultiplier
od; #altogether we obtain M(G) = O2(M(G)) ⊕ O3(M(G)) ≅ C12 × C4

LoadPackage("hap",false); #loads further packages
GroupHomology(G,2); #M(G) ≅ H2(G,Z)
```

The command `SchurExtension`, on the other hand, computes the infinite central extension $F/[F,N]$ from Theorem 5.24.

Theorem 5.19 (Universal coefficient theorem[12])**.** *For every finite abelian group A, the sequence*

$$1 \longrightarrow H_s^2(G/G', A) \xrightarrow{\Phi} H^2(G, A) \xrightarrow{\Psi} \mathrm{Hom}(A^*, M(G)) \longrightarrow 1.$$

is exact and split. In particular, we have

$$H^2(G, A) \cong \mathrm{Hom}(G/G' \times M(G), A).$$

Proof. The maps Φ and Ψ were defined in Lemma 5.13 and Lemma 5.14. Since Φ is injective, the sequence is exact at the first term. For $\kappa \in Z_s^2(G/G', A)$ and $\lambda \in A^*$ we have $\lambda \circ \kappa \in Z_s^2(G/G', \mathbb{C}^\times) = B_s^2(G/G', \mathbb{C}^\times)$ by Corollary 5.6. This shows $\Psi_{\Phi_\kappa} = 1$ and $\Phi(H_s^2(G/G', A)) \leq \mathrm{Ker}(\Psi)$. Conversely, let $\kappa \in Z^2(G, A)$ with $\Psi_\kappa = 1$. Let $\widehat{G} := A \times G$ be the extension with respect to κ. Via $a \mapsto (a, 1)$ we regard A as a subgroup of $\widehat{G}$. For $\lambda \in A^*$ there exists a $\varphi \in C^1(G, \mathbb{C}^\times)$ with $\lambda \circ \kappa = \partial\varphi$. We define $\widehat{\lambda}\colon \widehat{G} \to \mathbb{C}^\times$, $(a, x) \mapsto \lambda(a)\varphi(x)$. For $(a, x), (b, y) \in \widehat{G}$ we then have

$$\widehat{\lambda}\big((a,x)(b,y)\big) = \widehat{\lambda}\big(ab\kappa(x,y), xy\big) = \lambda(ab)\partial\varphi(x,y)\varphi(xy) = \lambda(a)\varphi(x)\lambda(b)\varphi(y) = \widehat{\lambda}(a,x)\widehat{\lambda}(b,y).$$

This shows $\widehat{\lambda} \in \widehat{G}^*$ and $\widehat{G}' \cap A \leq \mathrm{Ker}(\lambda)$. As is well known, for every $a \in A \setminus \{1\}$ there exists a $\lambda \in A^*$ with $\lambda(a) \neq 1$. This implies $\widehat{G}' \cap A = 1$. Hence $\widehat{G}/\widehat{G}'$ is an extension of

$$\widehat{G}/\widehat{G}'A \cong (\widehat{G}/A)/(\widehat{G}'A/A) \cong G/G'$$

by $A\widehat{G}'/\widehat{G}' \cong A/(A \cap \widehat{G}') \cong A$. Let $\pi\colon \widehat{G} \to G$, $(a, x) \mapsto x$ be the projection and $\overline{\pi}\colon \widehat{G}/\widehat{G}' \to G/G'$, $x\widehat{G}' \mapsto \pi(x)G'$. Because of $\widehat{G}' \cong \widehat{G}'A/A \cong G'$, the restriction of π to $\widehat{G}'$ is injective. For $s \in G'$ let $\widehat{s} := \pi^{-1}(s) \in \widehat{G}'$. Let $R \subseteq G$ be a set of coset representatives for G/G'. For $x \in G$ let $r_x \in R$ with $xG' = r_xG'$. For $r \in R$ let $\widehat{r} \in \widehat{G}$ be an arbitrary preimage under π. For every $x \in G$ there exists exactly one $s_x \in G'$ with $x = r_xs_x$. Now $\widehat{x} := \widehat{r_x}\widehat{s_x}$ is a preimage of x under π. Hence κ is equivalent to $\kappa' \in Z^2(G, A)$ with $\kappa'(x, y) = \widehat{x}\widehat{y}\widehat{xy}^{-1}$. Clearly, $\widehat{x}\widehat{G}'$ is a preimage of xG' under $\overline{\pi}$ that does not depend on the choice of the representative of the coset xG'. Thus there exists a cocycle $\overline{\kappa} \in Z^2(G/G', A)$ with

$$\overline{\kappa}(xG', yG') = \widehat{x}\widehat{y}\widehat{xy}^{-1}\widehat{G}' = \kappa'(x, y)\widehat{G}'.$$

This shows $\kappa B^2(G, A) = \kappa' B^2(G, A) = \Phi_{\overline{\kappa}} B^2(G, A)$, by identifying A with $A\widehat{G}'/\widehat{G}'$. Since $\widehat{G}/\widehat{G}'$ is abelian, we have $\overline{\kappa} \in Z_s^2(G/G', A)$ and $\Phi(H_s^2(G/G', A)) = \mathrm{Ker}(\Psi)$. Hence the sequence is also exact at the second term.

We now construct a homomorphism $\Gamma\colon \mathrm{Hom}(A^*, M(G)) \to H^2(G, A)$ with $\Psi \circ \Gamma = \mathrm{id}$. From this we obtain both the surjectivity of Ψ (exactness at the third term) and the splitting of the sequence. Let $Z := M(G)$ and let $\kappa \in Z^2(G, Z)$ be the cocycle of a maximal Schur extension. In the proof of Theorem 5.17 we have seen that the map $Z^* \to Z$, $\lambda \mapsto (\lambda \circ \kappa)B^2(G, \mathbb{C}^\times)$ is an isomorphism. Let $f \in \mathrm{Hom}(A^*, Z)$. For $\mu \in A^*$ there exists exactly one $\lambda_\mu \in Z^*$ with $f(\mu) = (\lambda_\mu \circ \kappa)B^2(G, \mathbb{C}^\times)$. For $x, y \in G$ we define $\alpha_f(x, y)\colon A^* \to \mathbb{C}^\times$, $\mu \mapsto \lambda_\mu(\kappa(x, y))$. From

$$(\lambda_{\mu_1\mu_2} \circ \kappa)B^2(G, \mathbb{C}^\times) = f(\mu_1\mu_2) = f(\mu_1)f(\mu_2) = (\lambda_{\mu_1} \circ \kappa)(\lambda_{\mu_2} \circ \kappa)B^2(G, \mathbb{C}^\times) = (\lambda_{\mu_1}\lambda_{\mu_2} \circ \kappa)B^2(G, \mathbb{C}^\times)$$

[12]The name describes the fact that the range A of the cocycles can be replaced by the *universal* range $\mathbb{C}^\times$.

we obtain $\lambda_{\mu_1\mu_2} = \lambda_{\mu_1}\lambda_{\mu_2}$. This shows $\alpha_f(x,y) \in (A^*)^*$. Since $|A| < \infty$, the map $A \to (A^*)^*$, $a \mapsto (\mu \mapsto \mu(a))$ is an isomorphism (GT, Exercise 79). Hence there exists exactly one $\Gamma_f(x,y) \in A$ with $\alpha_f(x,y)(\mu) = \mu(\Gamma_f(x,y))$ for all $\mu \in A^*$. For $x,y,z \in G$ and $\mu \in A^*$ we have

$$\begin{aligned}\mu\big(\Gamma_f(x,y)\Gamma_f(xy,z)\big) &= \alpha_f(x,y)(\mu)\alpha_f(xy,z)(\mu) = \lambda_\mu\big(\kappa(x,y)\kappa(xy,z)\big)\\ &= \lambda_\mu\big(\kappa(y,z)\kappa(x,yz)\big) = \mu\big(\Gamma_f(y,z)\Gamma_f(x,yz)\big).\end{aligned}$$

This shows $\Gamma_f \in Z^2(G,A)$. For $f' \in \mathrm{Hom}(A^*,Z)$ and $\mu \in A^*$ let $\lambda'_\mu \in Z^*$ with $f'(\mu) = (\lambda'_\mu \circ \kappa)B^2(G,\mathbb{C}^\times)$. Then $(ff')(\mu) = (\lambda_\mu\lambda'_\mu \circ \kappa)B^2(G,\mathbb{C}^\times)$ and

$$\mu\big(\Gamma_{ff'}(x,y)\big) = \alpha_{ff'}(x,y)(\mu) = (\lambda_\mu\lambda'_\mu)(\kappa(x,y)) = \alpha_f(x,y)(\mu)\alpha_{f'}(x,y)(\mu) = \mu\big(\Gamma_f(x,y)\Gamma_{f'}(x,y)\big).$$

Thus $\Gamma\colon \mathrm{Hom}(A^*,M(G)) \to H^2(G,A)$, $f \mapsto \Gamma_f B^2(G,A)$ is a homomorphism. From

$$\Psi_{\Gamma_f}(\mu)(x,y) = (\mu \circ \Gamma_f)(x,y) = \alpha_f(x,y)(\mu) = \lambda_\mu(\kappa(x,y)) = f(\mu)(x,y)$$

we obtain $\Psi \circ \Gamma = \mathrm{id}$.

For the second claim we use Remark 5.9 and $A^* \cong A$:

$$H^2(G,A) \cong H^2_s(G/G',A) \times \mathrm{Hom}(A^*,Z) \cong \mathrm{Hom}(G/G',A) \times \mathrm{Hom}(Z,A) \cong \mathrm{Hom}(G/G' \times Z, A). \quad \square$$

Theorem 5.20. *Let $G/G' \cong C_{d_1} \times \cdots \times C_{d_k}$ and $M(G) \cong C_{e_1} \times \cdots \times C_{e_l}$. Then G has at most*

$$\prod_{\substack{1\le i\le k\\ 1\le j\le l}} \gcd(d_i,e_j)$$

maximal Schur extensions up to isomorphism.

Proof. Let $\widehat{G} := Z \times G$ be a maximal Schur extension of G with respect to $\kappa \in Z^2(G,Z)$ with $Z \cong M(G)$. As in the proof of Theorem 5.17, $\Psi_\kappa\colon Z^* \to M(G)$ is an isomorphism. We show that the isomorphism type of $\widehat{G}$ does not depend on Ψ_κ. For this, let $f \in \mathrm{Aut}(Z)$ and $\kappa' \in Z^2(G,Z)$ with $\kappa'(x,y) = f(\kappa(x,y))$ for all $x,y \in G$. Let $\widehat{H} := Z \times G$ be the extension of G by Z with respect to κ'. Then

$$F\colon \widehat{G} \to \widehat{H}, \qquad (a,x) \mapsto (f(a),x)$$

is an isomorphism, since

$$\begin{aligned}F\big((a,x)(b,y)\big) &= F\big(ab\kappa(x,y),xy\big) = \big(f(ab\kappa(x,y)),xy\big)\\ &= \big(f(a)f(b)\kappa'(x,y),xy\big) = (f(a),x)(f(b),y) = F(a,x)F(b,y)\end{aligned}$$

for $(a,x),(b,y) \in \widehat{G}$. Let $f^* \in \mathrm{Aut}(Z^*)$, $\lambda \mapsto \lambda \circ f$ be the map dual to f. Then $\Psi_{\kappa'} = \Psi_\kappa \circ f^*$. In this way every isomorphism $Z^* \to M(G)$ can be realized. For every other maximal Schur extension of G with respect to some $\alpha \in Z^2(G,Z)$ we may therefore assume $\Psi_\alpha = \Psi_\kappa$. Hence the number of maximal Schur extensions is bounded by $|\Psi^{-1}(\Psi_\kappa)| = |\mathrm{Ker}(\Psi)| = |H^2_s(G/G',Z)|$ by Theorem 5.19. Corollary 5.7 yields the claim. $\square$

Example 5.21.

(i) Let p be a prime and $G \cong C_p^n$ elementary abelian. By GT, Example 11.38, we have $M(G) \cong C_p^{\binom{n}{2}}$. Hence G has at most $p^{n\binom{n}{2}}$ maximal Schur extensions up to isomorphism. This can also be proved

directly: Let $\widehat{G}$ be a maximal Schur extension and $Z \le \widehat{G}' \cap \mathrm{Z}(\widehat{G})$ with $\widehat{G}/Z \cong G$. Since $Z \le \Phi(\widehat{G})$, G can be generated by n elements $x_1, \ldots, x_n$. Then $\{[x_i, x_j] : 1 \le i < j \le n\}$ is a basis of $Z \cong M(G)$. For $i = 1, \ldots, n$ there exist $0 \le e_{ijk} < p$ with

$$x_i^p = \prod_{j<k} [x_j, x_k]^{e_{ijk}}.$$

Since $Z \le \mathrm{Z}(\widehat{G})$, $\widehat{G}$ is uniquely determined by the $n\binom{n}{2}$ parameters e_{ijk}. Clearly, many parameter values lead to isomorphic groups (for example by permuting the x_i). In Remark 13.38 we construct the maximal Schur extension with $p > 2$ and $e_{ijk} = 0$ for all i, j, k. For $n = 2$, the two nonabelian groups of order p^3 are the only maximal Schur extensions. For $p^n \in \{2^3, 2^4, 3^3, 5^3\}$ there are 10, 989, 16, and 20 maximal Schur extensions of G, respectively. In general, the number is unknown.

(ii) The minimal nonabelian 2-group $Q(2,1)$ from Exercise 6 has exactly seven maximal Schur extensions:

```
G:=SmallGroup(16,3);; #=Q(2,1)
ab:=AbelianInvariants(G); #=[2,4]
mG:=AbelianInvariantsMultiplier(G); #=[2,2]
ProductX(ab,mG,{d,e}->Gcd(d,e)); #=16
Size(SchurCovers(G)); #=7
```

Remark 5.22.

(i) If G is perfect (or more generally $\gcd(|G/G'|, |M(G)|) = 1$), then G has only one maximal Schur extension $\widehat{G}$ up to isomorphism. We call $\widehat{G}$ the *universal* Schur extension of G, since by Theorem 5.24 every Schur extension of G is a quotient of $\widehat{G}$.

(ii) For complete groups G, equality holds in Theorem 5.20 (without proof).

Theorem 5.23. *Let G be an arbitrary group with $|G : \mathrm{Z}(G)| < \infty$. Then $|G'| < \infty$.*

Proof (*Rosenlicht*). Let $Z := \mathrm{Z}(G)$ and $n := |G : Z| < \infty$. Let R be a set of coset representatives for G/Z. For $\Gamma := \{[r, s] : r, s \in R\}$ we have $|\Gamma| \le |R|^2 = |G/Z|^2 = n^2$. For $r, s \in R$ and $z \in Z$ we have $[rz, s] = [r, s] = [r, sz]$. Hence every element $g \in G'$ has the form $g = c_1 \ldots c_m$ with $c_1, \ldots, c_m \in \Gamma$. It suffices to show that we can choose $m \le n^3$ here (then $|G'| \le n^{2n^3} < \infty$ follows). Suppose $m > n^3$. Then there exists a $\gamma \in \Gamma$ with $|\{i \in \{1, \ldots, m\} : c_i = \gamma\}| > n$. Since $c_i c_{i+1} = c_{i+1}(c_{i+1}^{-1} c_i c_{i+1}) = c_{i+1}\delta$ with $\delta \in \Gamma$, we may assume $c_1 = \ldots = c_{n+1} = \gamma$. In contradiction to the minimality of m we will show that γ^{n+1} is a product of n commutators. For this, let $\gamma = [r, s]$ with $r, s \in R$. Since $\gamma^n = \gamma^{|G:Z|} \in Z$, we have

$$\gamma^{n+1} = \gamma\gamma^n = \gamma s \gamma^n s^{-1} = \gamma s \gamma s^{-1} (s\gamma s^{-1})^{n-1} = [r, s] s [r, s] s^{-1} [srs^{-1}, s]^{n-1} = [r, s^2][srs^{-1}, s]^{n-1}. \quad \square$$

Theorem 5.24. *Let $G = F/N$ be a finite group with $F = F_n$. Then:*

(i) $N/[F, N]$ is a finitely generated abelian group with free part of rank n and torsion part $(F' \cap N)/[F, N]$.

(ii) For $N/[F, N] = (F' \cap N)/[F, N] \oplus K/[F, N]$, F/K is a maximal Schur extension of G.

(iii) For every Schur extension $\widehat{G}$ of G there exists an $L \trianglelefteq F$ with $N = (F' \cap N)L$ and $\widehat{G} \cong F/L$. In particular, $\widehat{G}$ is a quotient group of a maximal Schur extension.

(iv) $\boxed{M(G) \cong (F' \cap N)/[F,N]}$ (Hopf formula).

Proof.

(i) By Theorem 2.11, N is finitely generated. Since $N \trianglelefteq F$, we have $[F,N] \trianglelefteq F$ and $[F,N] \leq F' \cap N$. Because of $N/[F,N] \leq \mathrm{Z}(F/[F,N])$, $\mathrm{Z}(F/[F,N])$ has finite index in $F/[F,N]$. By Theorem 5.23, $F'/[F,N]$ is finite. Hence $(F' \cap N)/[F,N]$ is finite as well. Since $N' \leq [F,N]$, $N/[F,N]$ is abelian. Furthermore,

$$(N/[F,N])/((F' \cap N)/[F,N]) \cong N/(F' \cap N) \cong NF'/F' \leq F/F'.$$

By Example 1.17, F/F' is a free abelian group of rank n. Since $|F/N| = |G| < \infty$, NF'/F' must also be a free abelian group of rank n. Hence $(F' \cap N)/[F,N]$ is the torsion part of $N/[F,N]$.

(ii) Since $K/[F,N] \leq N/[F,N] \leq \mathrm{Z}(F/[F,N])$, we have $K \trianglelefteq F$. Let $\widehat{G} := F/K$ and $Z := N/K$. Then $\widehat{G}/Z \cong F/N \cong G$ and $Z \leq \mathrm{Z}(\widehat{G})$ because of $[F,N] \leq K$. From $N/[F,N] \leq F'K/[F,N]$ it follows that

$$Z = N/K \leq F'K/K = (F/K)' = \widehat{G}'.$$

Hence $\widehat{G}$ is a Schur extension with $Z \cong (F' \cap N)/[F,N]$. From Theorem 5.15 we obtain $|M(G)| \geq |(F' \cap N)/[F,N]|$. For the reverse inequality we first show (iii).

(iii) Let $\alpha\colon F \to G$ and $\beta\colon \widehat{G} \to G$ be the canonical epimorphisms with $N = \mathrm{Ker}(\alpha)$ and $Z := \mathrm{Ker}(\beta)$. Since F is free, there exists a homomorphism $\rho\colon F \to \widehat{G}$ with $\beta\rho = \alpha$. Then $\widehat{G} = \rho(F)Z$ and $Z \leq \widehat{G}' \leq \rho(F)' \leq \rho(F)$, hence $\rho(F) = \widehat{G}$. Clearly, $L := \mathrm{Ker}(\rho) \leq \mathrm{Ker}(\alpha) = N$. Because of $\beta\rho(N) = \alpha(N) = 1$ we have $\rho(N) \leq \mathrm{Ker}(\beta) = Z$. This shows $\rho([F,N]) \leq [\widehat{G}, Z] = 1$ and $[F,N] \leq L$. From $\rho(N) = Z \leq \widehat{G}' = \rho(F')$ it also follows that $N \leq F'L$ and $(F' \cap N)L = F'L \cap N = N$ by Dedekind's modular law. Now we have

$$|Z| = \frac{|\widehat{G}|}{|G|} = |N : L| = |(F' \cap N)L : L| = |F' \cap N : F' \cap L| \leq |(F' \cap N)/[F,N]|.$$

Hence the Schur extension constructed in (ii) is indeed maximal.

By (i) and the fundamental theorem of finitely generated abelian groups we have

$$L/[F,N] = (L \cap F')/[F,N] \oplus M/[F,N],$$

where $(L \cap F')/[F,N]$ is the torsion part. Since $|N : L| = |F' \cap N : F' \cap L|$, $M/[F,N]$ is the torsion-free part of $N/[F,N]$. By (ii), F/M is a maximal Schur extension and $\widehat{G} \cong F/L \cong (F/M)/(L/M)$.

(iv) Follows from the proof of (ii). □

Theorem 5.25. *Let $G = \langle x_1, \ldots, x_n \mid r_1, \ldots, r_k \rangle$ be finite. Then $M(G)$ can be generated by $k - n$ elements. In particular, $M(G) = 1$ if $n = k$.*

Proof. By Theorem 1.21, $n \leq k$. Let F be the free group on $x_1, \ldots, x_n$ and $N := \langle r_1, \ldots, r_k \rangle^F$. Because of $N/[F,N] \leq \mathrm{Z}(F/[F,N])$ we have $N/[F,N] = \langle \overline{r_1}, \ldots, \overline{r_k} \rangle$, where $\overline{r_i} := r_i[F,N]$. By Theorem 5.24, the free part of $N/[F,N]$ has rank n and the torsion part is isomorphic to $M(G)$. Hence $M(G)$ can be generated by $k - n$ elements. □

Example 5.26. By Theorem 5.25 and Exercise 2, $M(Q_{2^n}) = 1$ for $n \geq 3$ (cf. GT, Example 11.38). For infinite groups, Theorem 5.25 is false: For example, $F_2/F_2^{[3]}$ is a Schur extension of $F_2/F_2' \cong \langle x, y \mid [x,y] = 1 \rangle \cong C_\infty^2$.

Theorem 5.27. *Let $\widehat{G}_1$, $\widehat{G}_2$ be maximal Schur extensions of G with $\widehat{G}_1/Z_1 \cong G \cong \widehat{G}_2/Z_2$. Then*

(i) (Schur) $\widehat{G}_1' \cong \widehat{G}_2'$ *and* $\widehat{G}_1/\widehat{G}_1' \cong G/G' \cong \widehat{G}_2/\widehat{G}_2'$.

(ii) (Gaschütz) $\widehat{G}_1/\mathrm{Z}(\widehat{G}_1) \cong \widehat{G}_2/\mathrm{Z}(\widehat{G}_2)$.

(iii) (Read) $\mathrm{Z}(\widehat{G}_1)/Z_1 \cong \mathrm{Z}(\widehat{G}_2)/Z_2$.

Proof. Let $G = F/N$, $K_i \trianglelefteq F$ and $\widehat{G}_i \cong F/K_i$ as in Theorem 5.24. We show that the stated groups do not depend on i.

(i) We have

$$\begin{aligned}\widehat{G}_i' &\cong F'K_i/K_i \cong F'/(F' \cap K_i) = F'/(F' \cap N \cap K_i) = F'/[F,N],\\ \widehat{G}_i/\widehat{G}_i' &\cong (\widehat{G}_i/Z_i)/(\widehat{G}_i'/Z_i) \cong G/G'.\end{aligned}$$

(ii) For $L/[F,N] := \mathrm{Z}(F/[F,N])$ we have $[F,L] \leq [F,N] \leq K_i$ and $L/K_i \leq \mathrm{Z}(F/K_i)$. Conversely, let $xK_i \in \mathrm{Z}(F/K_i)$. Then $[x,F] \leq K_i \cap F' = K_i \cap F' \cap N = [F,N]$ and it follows $x[F,N] \in \mathrm{Z}(F/[F,N]) = L/[F,N]$. This shows

$$\widehat{G}_i/\mathrm{Z}(\widehat{G}_i) \cong (F/K_i)/\mathrm{Z}(F/K_i) = (F/K_i)/(L/K_i) \cong F/L.$$

(iii) With the notation from (ii) we have

$$\mathrm{Z}(\widehat{G}_i)/Z_i \cong (L/K_i)/(N/K_i) \cong L/N. \qquad \square$$

Remark 5.28.

(i) In general, $\mathrm{Z}(\widehat{G}_1) \ncong \mathrm{Z}(\widehat{G}_2)$ in the situation of Theorem 5.27. For example, the minimal nonabelian groups

$$\begin{aligned}\widehat{G}_1 &= \langle x,y \mid x^8 = y^2 = 1,\ yxy^{-1} = x^5\rangle,\\ \widehat{G}_2 &= \langle x,y \mid x^4 = y^2 = [x,y]^2 = [x,x,y] = [y,x,y] = 1\rangle\end{aligned}$$

of order 16 from Exercise 4 and Exercise 6 are maximal Schur extensions of $G = C_4 \times C_2$ with $\mathrm{Z}(\widehat{G}_1) = \langle x^2\rangle \cong C_4$ and $\mathrm{Z}(\widehat{G}_2) = \langle x^2, [x,y]\rangle \cong C_2^2$ (GT, Theorem 11.37 implies $M(G) \cong C_2$).

(ii) Schur multipliers also occur in the representation theory of finite groups: Let $N \trianglelefteq G$ and $\Delta\colon N \to \mathrm{GL}(d,\mathbb{C})$ be an irreducible representation of N. For $g \in G$, ${}^g\Delta\colon N \to \mathrm{GL}(d,\mathbb{C})$, $x \mapsto \Delta(g^{-1}xg)$ is also a representation. Suppose that Δ and ${}^g\Delta$ are equivalent for all $g \in G$, i. e., for $g \in G$ there exists $T_g \in \mathrm{GL}(d,\mathbb{C})$ with ${}^g\Delta(x) = T_g\Delta(x)T_g^{-1}$ for all $x \in N$. Let $g_1,\dots,g_k \in G$ be a set of coset representatives for G/N and $T_i := T_{g_i}$. For $x \in N$ we define

$$\Gamma\colon G \to \mathrm{GL}(d,\mathbb{C}), \quad g_ix \mapsto \Delta(x^{-1})T_i.$$

For $y \in N$ we have

$$\Gamma(g_ix)\Delta(y)\Gamma(g_ix)^{-1} = \Delta(x^{-1})T_i\Delta(y)T_i^{-1}\Delta(x) = \Delta(x^{-1}g_i^{-1}yg_ix) = {}^{g_ix}\Delta(y).$$

This shows

$$\Gamma(gh)\Delta(y)\Gamma(gh)^{-1} = {}^{gh}\Delta(y) = {}^g({}^h\Delta)(y) = \Gamma(g)\Gamma(h)\Delta(y)\Gamma(h)^{-1}\Gamma(g)^{-1}$$

for $g, h \in G$ and all $y \in N$. From Schur's lemma[13] we obtain $\Gamma(gh) = \kappa(g,h)\Gamma(g)\Gamma(h)$ for some $\kappa(g,h) \in \mathbb{C}^\times$. One calls Γ a *projective representation* of G, since $G \to \mathrm{PGL}(d,\mathbb{C})$, $g \mapsto \Gamma(g)\mathbb{C}^\times$ is a homomorphism. Because of

$$\begin{aligned}\kappa(x,yz)\kappa(y,z)\Gamma(x)\Gamma(y)\Gamma(z) &= \kappa(x,yz)\Gamma(x)\Gamma(yz) = \Gamma(x(yz)) = \Gamma((xy)z)\\ &= \kappa(xy,z)\Gamma(xy)\Gamma(z) = \kappa(xy,z)\kappa(x,y)\Gamma(x)\Gamma(y)\Gamma(z)\end{aligned}$$

we have $\kappa \in Z^2(G,\mathbb{C}^\times)$. In fact, κ depends only on G/N, i. e., $\kappa \in Z^2(G/N,\mathbb{C}^\times)$. Other choices of the T_i yield equivalent κ. In the case $M(G/N) = 1$ one can achieve $\kappa = 1$. Then Γ is an ordinary representation extending Δ.

(iii) We now deal with the question which groups have the form $G/\mathrm{Z}(G)$. Here we again allow infinite groups G.

Definition 5.29. We call G *capable* if there exists a group H with $H/\mathrm{Z}(H) \cong G$. Let $\mathrm{Z}^*(G)$ be the intersection of all normal subgroups N such that G/N is capable. We call $\mathrm{Z}^*(G)$ the *epicenter* of G.

Remark 5.30.

(i) Capable groups are inner automorphism groups because of $H/\mathrm{Z}(H) \cong \mathrm{Inn}(H)$.

(ii) Since $G/\mathrm{Z}(G)$ is capable, we have $\mathrm{Z}^*(G) \leq \mathrm{Z}(G)$. Clearly, $\mathrm{Z}^*(G)$ is characteristic in G.

Example 5.31. Groups with trivial center are capable. Because of $D_{4n}/\mathrm{Z}(D_{4n}) \cong D_{2n}$ all dihedral groups are capable. Nontrivial cyclic groups are not capable.

Lemma 5.32. *For every group G, $G/\mathrm{Z}^*(G)$ is capable.*

Proof. Let $\{N_i : i \in I\}$ be the set of normal subgroups of G such that G/N_i is capable. Then there exist groups $\{H_i : i \in I\}$ and epimorphisms $\varphi_i \colon H_i \to G/N_i$ with $\mathrm{Ker}(\varphi_i) = \mathrm{Z}(H_i)$. Let $H := \times_{i\in I} H_i$ and

$$L := \left\{(h_i)_i \in H : \exists g \in G : \forall i \in I : \varphi_i(h_i) = gN_i\right\} \leq H.$$

Clearly, $\mathrm{Z}(H) = \times \mathrm{Z}(H_i) \leq \mathrm{Z}(L)$. Since the projection of L to H_i is surjective for every i, we have $\mathrm{Z}(H) = \mathrm{Z}(L)$. For every $g \in G$ there exists exactly one tuple $(h_i)_i\mathrm{Z}(L) \in L/\mathrm{Z}(L)$ with $\varphi_i(h_i) = gN_i$ for all $i \in I$. One easily sees that the map $G \to L/\mathrm{Z}(L)$, $g \mapsto (h_i)_i\mathrm{Z}(L)$ is an epimorphism with kernel $\mathrm{Z}^*(G) = \bigcap_{i\in I} N_i$. Hence $G/\mathrm{Z}^*(G) \cong L/\mathrm{Z}(L)$ is capable. □

Corollary 5.33. *G is capable if and only if $\mathrm{Z}^*(G) = 1$.*

Lemma 5.34. *For every group G, $\mathrm{Z}^*(G)$ is the intersection of all subgroups of the form $\pi(\mathrm{Z}(H))$, where $Z \to H \xrightarrow{\pi} G$ is a central extension of G.*

Proof. Let $D \leq G$ be the intersection over $\pi(\mathrm{Z}(H))$ as stated. For every central extension $\pi \colon H \to G$ we have $\pi(\mathrm{Z}(H)) \trianglelefteq \pi(H) = G$ and

$$G/\pi(\mathrm{Z}(H)) \cong (H/\mathrm{Ker}(\pi))/(\mathrm{Z}(H)/\mathrm{Ker}(\pi)) \cong H/\mathrm{Z}(H).$$

Hence $G/\pi(\mathrm{Z}(H))$ is capable and $\mathrm{Z}^*(G) \leq D$.

[13] see Lemma 1.9 in the Character Theory lecture notes

Conversely, let G/N be capable for some $N \trianglelefteq G$. Then there exists an epimorphism $\alpha\colon H \to G/N$ with $\mathrm{Ker}(\alpha) = \mathrm{Z}(H)$. Let

$$L := \big\{(x,y) \in G \times H : \alpha(y) = xN\big\} \leq G \times H.$$

The projection $\rho\colon L \to G$, $(x,y) \mapsto x$ is surjective with $\mathrm{Ker}(\rho) = 1 \times \mathrm{Ker}(\alpha) \leq \mathrm{Z}(L)$. Hence ρ is a central extension of G. It follows that $D \leq \rho(\mathrm{Z}(L))$. Since the projection of L to H is surjective, we have $\mathrm{Z}(L) \leq G \times \mathrm{Z}(H)$. Because of $\mathrm{Z}(H) = \mathrm{Ker}(\alpha)$ we have $\mathrm{Z}(L) \leq N \times \mathrm{Z}(H)$ and $D \leq \rho(\mathrm{Z}(L)) \leq N$. This shows $D \leq \mathrm{Z}^*(G)$. □

Remark 5.35. We show that in Lemma 5.34 (for finite groups) one only has to consider a single central extension.

Theorem 5.36 (Beyl-Felgner-Schmid)**.** *Let $\widehat{G}$ be a maximal Schur extension of a finite group G and $\gamma\colon \widehat{G} \to G$ the corresponding epimorphism. Then $\mathrm{Z}^*(G) = \gamma(\mathrm{Z}(\widehat{G}))$.*

Proof. Since every Schur extension is central, we have $\mathrm{Z}^*(G) \leq \gamma(\mathrm{Z}(\widehat{G}))$ by Lemma 5.34. Conversely, let $\alpha\colon H \to G$ be a central extension. We have to show $\gamma(\mathrm{Z}(\widehat{G})) \leq \alpha(\mathrm{Z}(H))$. Let F be a free group and $\pi\colon F \to G$ an epimorphism with kernel N. Then there exists a homomorphism $\varphi\colon F \to H$ with $\alpha\varphi = \pi$. Since $\alpha(\varphi(N)) = \pi(N) = 1$ we have $\varphi(N) \leq \mathrm{Ker}(\alpha) \leq \mathrm{Z}(H)$. It follows that $\varphi([F,N]) \leq [H, \mathrm{Z}(H)] = 1$. Since π is surjective, we have $H = \mathrm{Ker}(\alpha)\varphi(F) = \mathrm{Z}(H)\varphi(F)$.

By Theorem 5.24 we may assume $\widehat{G} = F/K$, where $N/[F,N] = (F' \cap N)/[F,N] \oplus K/[F,N]$. Moreover, let $\gamma(xK) = \pi(x)$ for all $x \in F$. Let $W/K := \mathrm{Z}(\widehat{G})$. Then

$$[\varphi(W), \varphi(F)] = \varphi([W,F]) \leq \varphi(F' \cap K) = \varphi(F' \cap N \cap K) = \varphi([F,N]) = 1.$$

It follows that $\varphi(W) \leq \mathrm{Z}(H)$ and $\gamma(\mathrm{Z}(\widehat{G})) = \pi(W) \leq \alpha(\mathrm{Z}(H))$. □

Corollary 5.37. *Let G be finite with $M(G) = 1$. Then G is capable if and only if $\mathrm{Z}(G) = 1$.*

Example 5.38. For $n \geq 3$, Q_{2^n} is not capable, since $M(Q_{2^n}) = 1$ (Example 5.26) and $\mathrm{Z}(Q_{2^n}) \neq 1$.

Theorem 5.39 (Baer)**.** *Let $A = C_{d_1} \times \ldots \times C_{d_k}$ be a finite abelian group with $1 \neq d_1 \mid d_2 \mid \ldots \mid d_k$. Then A is capable if and only if $k \geq 2$ and $d_{k-1} = d_k$.*

Proof. For $k = 1$, $A \neq 1$ is cyclic and therefore not capable. Now let $k \geq 2$. Let $A = \langle a_1 \rangle \times \ldots \times \langle a_k \rangle$ with $|\langle a_i \rangle| = d_i$ for $i = 1, \ldots, k$. Suppose first that A is capable. Let $\gamma\colon H \to A$ be an epimorphism with $\mathrm{Ker}(\gamma) = \mathrm{Z}(H)$. Let $H_i := \gamma^{-1}(\langle a_i a_k \rangle)$ for $i = 1, \ldots, k-1$ and $H_k := \gamma^{-1}(\langle a_k \rangle)$. Since $H_i/\mathrm{Z}(H)$ is cyclic, H_i is abelian for $i = 1, \ldots, k$. Because of $A = \langle a_1 a_k, \ldots, a_{k-1} a_k, a_k \rangle$ we have $H = \langle H_1, \ldots, H_k \rangle$ and $\bigcap_{i=1}^k H_i = \mathrm{Z}(H)$. Because of $(a_i a_k)^{d_{k-1}} = a_k^{d_{k-1}}$ for $i = 1, \ldots, k-1$ we have $\gamma^{-1}(a_k^{d_{k-1}}) \subseteq \mathrm{Z}(H) = \mathrm{Ker}(\gamma)$, i. e., $d_k = |\langle a_k \rangle| = d_{k-1}$.

Now let $Z := \mathrm{Z}^*(A) \neq 1$. Let F be a free group and $N, L \trianglelefteq F$ with $F/N \cong A$ and $F/L \cong (F/N)/(L/N) \cong A/Z$. Let $\widehat{A} := F/K$ be a maximal Schur extension of A with $N/[F,N] = (F' \cap N)/[F,N] \oplus K/[F,N]$. By Theorem 5.36 we have $\mathrm{Z}(\widehat{A}) = L/K$ and $[F,L] \leq F' \cap K \leq [F,N]$. Hence the map

$$M(A) = (F' \cap N)/[F,N] \to (F' \cap L)/[F,L] = M(A/Z), \qquad x[F,N] \mapsto x[F,L]$$

is injective. In particular, $|M(A)| \leq |M(A/Z)|$. By GT, Example 11.38 we have $|M(A)| = d_1^{k-1} d_2^{k-2} \ldots d_{k-1}$. On the other hand, there exist $e_i \mid d_i$ with $|M(A/Z)| = e_1^{k-1} \ldots e_{k-1}$. From $|M(A)| \leq |M(A/Z)|$ it now easily follows that $d_{k-1} < d_k$. □

Definition 5.40. Groups G and H are called *isoclinic*[14], if there exist isomorphisms

$$\varphi\colon G' \to H', \qquad \psi\colon G/\mathrm{Z}(G) \to H/\mathrm{Z}(H), \quad x\mathrm{Z}(G) \mapsto \tilde{x}\mathrm{Z}(H)$$

such that $\varphi([x,y]) = [\tilde{x},\tilde{y}]$ for all $x, y \in G$. In this case, ψ is called an *isoclinism* and we write $G \approx H$.

Remark 5.41.

(i) The property $\varphi([x,y]) = [\tilde{x},\tilde{y}]$ does not depend on the choice of the representatives $\tilde{x}$. Moreover, $\varphi(g)\mathrm{Z}(H) = \psi(g\mathrm{Z}(G))$ for all $g \in G'$.

(ii) Clearly, isoclinism is an equivalence relation and isomorphic groups are isoclinic.

(iii) The map $\varphi\colon G' \to H'$ is uniquely determined by an isoclinism.

(iv) For $G \approx H$ we have $G' \cong H'$ and

$$G/\mathrm{Z}(G)G' \cong (G/\mathrm{Z}(G))/(G/\mathrm{Z}(G))' \cong (H/\mathrm{Z}(H))/(H/\mathrm{Z}(H))' \cong H/\mathrm{Z}(H)H'.$$

Hence $|G|$ and $|H|$ differ only by the so-called *branching factor*

$$\mathrm{Z}(G)G'/G' \cong \mathrm{Z}(G)/(\mathrm{Z}(G) \cap G').$$

For $x \in G' \cap \mathrm{Z}(G)$ we have $\varphi(x)\mathrm{Z}(H) = \psi(x\mathrm{Z}(G)) = 1$, i. e., $\varphi(x) \in \mathrm{Z}(H)$. The restriction of φ therefore yields an isomorphism $G' \cap \mathrm{Z}(G) \cong H' \cap \mathrm{Z}(H)$.

(v) For $G \approx H$ we have $\varphi(G') = H'$ by assumption. Suppose that $\varphi(G^{[k]}) = H^{[k]}$ for some $k \geq 2$. For $y \in G^{[k]}$ we have $\tilde{y} \in \varphi(y)\mathrm{Z}(H) \subseteq H^{[k]}\mathrm{Z}(H)$. For $x \in G$ it follows that $\varphi([x,y]) = [\tilde{x},\tilde{y}] \in H^{[k+1]}$. This shows $\varphi(G^{[k+1]}) \subseteq H^{[k+1]}$. By symmetry, equality must hold. In particular, $G^{[k]} \cong H^{[k]}$ for all $k \geq 2$. Moreover, $G/\mathrm{Z}_k(G) \cong H/\mathrm{Z}_k(H)$ for $k \geq 1$. In particular, G and H have the same nilpotency class (if nilpotent) and derived length (if solvable).

(vi) $G/\mathrm{Z}(G) \cong H/\mathrm{Z}(H)$ and $G' \cong H'$ do not imply $G \approx H$ (see Example 5.48).

Example 5.42.

(i) All abelian groups are isoclinic to each other. More generally, $G \approx G \times A$ for every abelian group A.

(ii) Let $\widehat{G}_1$ and $\widehat{G}_2$ be maximal Schur extensions of G. As in the proof of Theorem 5.27 let $G = F/N$ and $\widehat{G}_i = F/K_i$ for a free group F. For $L/[F,N] = \mathrm{Z}(F/[F,N])$ the map

$$\psi\colon \widehat{G}_1/\mathrm{Z}(\widehat{G}_1) \cong F/L \cong \widehat{G}_2/\mathrm{Z}(\widehat{G}_2), \qquad xK_1\mathrm{Z}(\widehat{G}_1) \mapsto xK_2\mathrm{Z}(\widehat{G}_2)$$

is a canonical isomorphism. Since the isomorphism $\widehat{G}_1' \cong F'/[F,N] \cong \widehat{G}_2'$ is also canonical, ψ is an isoclinism. Conversely, not every group isoclinic to $\widehat{G}_1$ (of the same order) is a (maximal) Schur extension of G. For example, Q_8 is a maximal Schur extension of itself by Example 5.26, but $Q_8 \approx D_8$.

(iii) Let $H \leq G$ with $G = H\mathrm{Z}(G)$. Then $\mathrm{Z}(H) = H \cap \mathrm{Z}(G)$, $H' = G'$ and the map $H/\mathrm{Z}(H) \to G/\mathrm{Z}(G)$, $h\mathrm{Z}(H) \mapsto h\mathrm{Z}(G)$ is an isoclinism. Hence $G \approx H$.

[14] Hall introduced the term as *isoclinic* in the English original and suggested *gleich schief* ("equally skewed") as a German translation.

(iv) Let G be finite. If $\mathrm{Z}(G) \not\subseteq \Phi(G)$, then there exists a maximal subgroup $M < G$ with $G = M\mathrm{Z}(G)$. By (iii) we have $G \approx M$. Repeating the argument with M, one finally obtains a subgroup $H \leq G$ with $G \approx H$ and $\mathrm{Z}(H) \leq \Phi(H)$.

(v) Let $N \trianglelefteq G$ with $N \cap G' = 1$. Because of $[G, N] \leq N \cap G' = 1$ we have $N \leq \mathrm{Z}(G)$. Clearly, $G' \to G'N/N \cong (G/N)'$, $x \mapsto xN$ is an isomorphism. For $xN \in \mathrm{Z}(G/N)$ we have $[x, G] \leq G' \cap N = 1$ and $x \in \mathrm{Z}(G)$. This shows $\mathrm{Z}(G/N) = \mathrm{Z}(G)/N$. Hence $G/\mathrm{Z}(G) \cong (G/N)/\mathrm{Z}(G/N)$, $x\mathrm{Z}(G) \mapsto xN\mathrm{Z}(G/N)$ is an isoclinism. It follows that $G \approx G/N$.

(vi) If H is a maximal Schur extension of $G/\mathrm{Z}(G)$, then $H/\mathrm{Z}(H) \cong G/\mathrm{Z}(G)$ by Theorem 5.36. In general, however, G and H are not isoclinic (consider $G = A_4$ and $H = \mathrm{SL}(2, 3)$). The following theorem shows that G is nevertheless isoclinic to a Schur extension of $G/\mathrm{Z}(G)$.

Definition 5.43. We call G a *stem group* if $\mathrm{Z}(G) \leq G'$.

Theorem 5.44 (Hall)**.** *Every group is isoclinic to a stem group.*

Proof. Let $\{g_i : i \in I\}$ be a generating set of G and $A = \bigoplus_{i\in I}\langle a_i\rangle \cong \bigoplus_{i\in I} C_\infty$. We consider

$$H := \langle (g_i, a_i) : i \in I\rangle \leq G \times A.$$

Since $1 \times A \leq \mathrm{Z}(G \times A)$ we have $HA = G \times A$ and $G \approx G \times A \approx H$ by Example 5.42. Because of

$$[(g_i, a_i), (g_j, a_j)] = ([g_i, g_j], 1)$$

we have $H' = G' \times 1$. In H/H' the cosets of (g_i, a_i) generate a free abelian group, i. e., $H/H' \cong A$. Hence $\overline{Z} := \mathrm{Z}(H)/(\mathrm{Z}(H) \cap H') \cong \mathrm{Z}(H)H'/H'$ is also a free abelian group.[15] By the universal property of free (abelian) groups there exists a homomorphism $\alpha\colon \overline{Z} \to \mathrm{Z}(H)$ with $\mathrm{Z}(H) = \alpha(\overline{Z}) \oplus (\mathrm{Z}(H) \cap H')$. For $K := \alpha(\overline{Z})$ we have $H \approx H/K$, since $K \cap H' = 1$ (Example 5.42). Because of

$$\mathrm{Z}(H/K) = \mathrm{Z}(H)/K = (\mathrm{Z}(H) \cap H')K/K \leq H'K/K = (H/K)'$$

we see that H/K is a stem group isoclinic to G. □

Remark 5.45. The stem groups in an equivalence class of isoclinic groups are exactly the representatives of minimal order, since the branching factor is trivial.

Definition 5.46. For $n \in \mathbb{N}$ let $k_n(G)$ be the number of conjugacy classes of G of length n.

Theorem 5.47. *For finite groups $G \approx H$ and $n \in \mathbb{N}$ we have $k_n(G)|\mathrm{Z}(H)| = k_n(H)|\mathrm{Z}(G)|$.*

Proof. Let $\psi\colon G/\mathrm{Z}(G) \to H/\mathrm{Z}(H)$, $x\mathrm{Z}(G) \mapsto \tilde{x}\mathrm{Z}(H)$ be an isoclinism. For $g, x \in G$ we have

$$g \in \mathrm{C}_G(x) \iff [g, x] = 1 \iff \varphi([g, x]) = 1 \iff [\tilde{g}, \tilde{x}] = 1 \iff \tilde{g} \in \mathrm{C}_H(\tilde{x}).$$

This shows

$$|^G x| = |G : \mathrm{C}_G(x)| = \frac{|G/\mathrm{Z}(G)|}{|\mathrm{C}_G(x)/\mathrm{Z}(G)|} = \frac{|H/\mathrm{Z}(H)|}{|\mathrm{C}_H(\tilde{x})/\mathrm{Z}(H)|} = |H : \mathrm{C}_H(\tilde{x})| = |^H\tilde{x}|.$$

All elements in $x\mathrm{Z}(G)$ lie in conjugacy classes of the same length. For $n \in \mathbb{N}$ it follows that

$$|\mathrm{Z}(H)|k_n(G)n = |\mathrm{Z}(H)|\left|\bigcup_{\substack{x\in G\\ |^G x|=n}} x\mathrm{Z}(G)\right| = |\mathrm{Z}(H)|\sum_{\substack{x\mathrm{Z}(G)\in G/\mathrm{Z}(G)\\ |^G x|=n}} |\mathrm{Z}(G)| = |\mathrm{Z}(G)|\sum_{\substack{\tilde{x}\mathrm{Z}(H)\in H/\mathrm{Z}(H)\\ |^H\tilde{x}|=n}} |\mathrm{Z}(H)| = |\mathrm{Z}(G)|k_n(H)n. \quad \square$$

[15] See the appendix in the Algebra lecture notes

Example 5.48. The following code yields nonisoclinic groups G and H of order 64 with $G' \cong H'$ and $G/\mathrm{Z}(G) \cong H/\mathrm{Z}(H)$:

```
G:=SmallGroup(64,128);;
H:=SmallGroup(64,149);;
IdGroup(DerivedSubgroup(G))=IdGroup(DerivedSubgroup(H));
IdGroup(G/Center(G))=IdGroup(H/Center(H));
Number(ConjugacyClasses(G),K->Size(K)=2)=Number(ConjugacyClasses(H),K->Size(K)=2);
```

6 Extensions of the Alternating Groups

Remark 6.1. We know from group theory that the alternating groups A_n are simple for $n \geq 5$. In this chapter we determine extensions of A_n by other groups and of other groups by A_n.

Lemma 6.2. *For $n \geq 4$ we have $\mathrm{C}_{S_n}(A_n) = 1$. In particular, $\mathrm{Z}(S_n) = \mathrm{Z}(A_n) = 1$.*

Proof. Let $\sigma \in \mathrm{C}_{S_n}(A_n)$ and let $\tau := (a, b, c) \in A_n$ be a 3-cycle. From $\sigma\tau = \tau\sigma$ we obtain ${}^{\sigma}\{a, b, c\} = \{a, b, c\}$. Now $\{a\}$ is the intersection of all 3-element sets $\Delta \subseteq \{1, \ldots, n\}$ containing a (note that $n \geq 4$). This shows $\sigma(a) = a$ for $a = 1, \ldots, n$. □

Remark 6.3. For every homomorphism $\omega\colon H \to \mathrm{Out}(A_n)$ ($n \geq 4$), Corollary 4.28 therefore yields exactly one extension of A_n by H with coupling ω. Hence we have to determine $\mathrm{Out}(A_n)$.

Lemma 6.4 (GT, Exercise 39)**.** *For $n \geq 3$ we have $A_n = \langle (1, 2, 3), \ldots, (1, 2, n) \rangle$.*

Proof. We argue by induction on n. The case $n = 3$ is clear. Let now $n \geq 4$. It suffices to show that $A_n = \langle A_{n-1}, (1, 2, n) \rangle =: H$. Suppose for contradiction that $\sigma \in A_n \setminus H$. Then there exists $k \neq n$ with ${}^{\sigma}k = n$. Choose $\tau \in A_{n-1}$ with ${}^{\tau}1 = k$. Then $\sigma\tau(1, 2, n) \in A_{n-1}$ and we obtain the contradiction $\sigma \in H$. □

Lemma 6.5. *Let Ω be a set and $G \leq \mathrm{Sym}(\Omega)$ with exactly one minimal normal subgroup. Then there exists an orbit $\Delta \subseteq \Omega$ such that G acts faithfully on Δ.*

Proof. Let $\Delta_1, \ldots, \Delta_s$ be the orbits of G. Let $N \trianglelefteq G$ be the unique minimal normal subgroup and $x \in N \setminus \{1\}$. Then there exists $1 \leq i \leq s$ such that x acts nontrivially on Δ_i. Hence the kernel K of the restricted action $G \to \mathrm{Sym}(\Delta_i)$ satisfies $K \cap N = 1$. This shows $K = 1$, and G acts faithfully on Δ_i. □

Remark 6.6.

(i) By GT, Theorem 6.21 and GT, Theorem 6.39, the hypothesis of Lemma 6.5 holds for all symmetric and alternating groups.

(ii) For a k-cycle $(a_1, \dots, a_k) \in S_n$ and an arbitrary permutation $\sigma \in S_n$ we have

$$\sigma(a_1, \dots, a_k)\sigma^{-1} = (\sigma(a_1), \dots, \sigma(a_k)).$$

This shows that all k-cycles in S_n are conjugate. By decomposing into disjoint cycles one obtains that two arbitrary permutations $\sigma, \tau \in S_n$ are conjugate if and only if they have the same cycle type. For $\sigma \in A_n$, the set ${}^{S_n}\sigma \subseteq A_n$ is a union of conjugacy classes of A_n. Because of

$$|A_n : \mathrm{C}_{A_n}(\sigma)| = |A_n : A_n \cap \mathrm{C}_{S_n}(\sigma)| = |A_n\mathrm{C}_{S_n}(\sigma) : \mathrm{C}_{S_n}(\sigma)|$$

we have ${}^{S_n}\sigma = {}^{A_n}\sigma$ if and only if $\mathrm{C}_{S_n}(\sigma) \not\subseteq A_n$. Let $\sigma = \sigma_1 \dots \sigma_l$ with disjoint cycles σ_i of length λ_i (including trivial cycles). If λ_i is even, then $\sigma_i \in \mathrm{C}_{S_n}(\sigma) \setminus A_n$. Let us now assume that $\lambda_1, \dots, \lambda_l$ are odd. Suppose that there exist $i \neq j$ with $\lambda_i = \lambda_j$. For $\sigma_i = (a_1, \dots, a_{\lambda_i})$ and $\sigma_j = (b_1, \dots, b_{\lambda_i})$ we have $(a_1, b_1) \dots (a_{\lambda_i}, b_{\lambda_i}) \in \mathrm{C}_{S_n}(\sigma) \setminus A_n$.

In the case ${}^{S_n}\sigma \neq {}^{A_n}\sigma$, the numbers $\lambda_1, \dots, \lambda_l$ must therefore be odd and pairwise distinct. One can show that the converse holds as well.

Lemma 6.7. *Let $n \geq 4$ and $A_{n-1} \cong H \leq A_n$. Then $H = \mathrm{Alt}(\{1, \dots, n\} \setminus \{i\})$ for some $i \in \{1, \dots, n\}$ or $n = 6$.*

Proof. W. l. o. g. we may assume $n \geq 5$. By Lemma 6.5, H acts faithfully on an orbit $\Delta \subseteq \{1, \dots, n\}$. Since $|H| = (n-1)!/2$, we have $|\Delta| \geq n - 1$. Hence we may assume that H acts transitively on $\{1, \dots, n\}$. In particular, $n \mid |H| \mid (n-1)!$. Therefore we may assume $n \geq 8$.

Let $f\colon A_{n-1} \to H$ be an isomorphism and let $\sigma \in A_{n-1}$ be a 3-cycle. Clearly, A_{n-4} is isomorphic to a subgroup of $\mathrm{C}_{A_{n-1}}(\sigma)$. In particular, $\mathrm{C}_H(f(\sigma))$ also has a subgroup $C \cong A_{n-4}$. Let $f(\sigma)$ be the disjoint product of k many 3-cycles. Clearly, C permutes the orbits of $f(\sigma)$ (including trivial orbits). The kernel of this action is a 3-group and hence trivial. Thus there exists a monomorphism $F\colon C \to S_{n-2k}$. Since $|C| = (n-4)!/2$, it follows that $k \leq 2$. In the case $k = 2$, F is transitive and we obtain the contradiction $n = 3k = 6$. Hence $f(\sigma)$ is also a 3-cycle.

Let $f((1,2,3)) = (\alpha, \beta, \gamma)$ and $f((1,2,4)) = (\delta, \epsilon, \varphi)$. In the case $\{\alpha, \beta, \gamma\} \cap \{\delta, \epsilon, \varphi\} = \varnothing$ the group

$$A_4 \cong \langle (1,2,3), (1,2,4) \rangle \cong \langle f((1,2,3)), f((1,2,4)) \rangle$$

would be abelian. In the case $|\{\alpha, \beta, \gamma\} \cap \{\delta, \epsilon, \varphi\}| = 1$, the group $\langle f((1,2,3)), f((1,2,4)) \rangle$ would act transitively on the 5-element set $\{\alpha, \beta, \gamma\} \cup \{\delta, \epsilon, \varphi\}$. Hence $|\{\alpha, \beta, \gamma\} \cap \{\delta, \epsilon, \varphi\}| = 2$ and $\langle f((1,2,3)), f((1,2,4)) \rangle$ can move only four points. By induction on k one sees that $\langle f((1,2,3)), \dots, f((1,2,k)) \rangle$ can move at most k points. Hence $H = f(A_{n-1}) = \langle f((1,2,3)), \dots, f((1,2,n-1)) \rangle$ (Lemma 6.4) is intransitive. But this was already excluded. □

Theorem 6.8 (Hölder)**.** *We have $\mathrm{Aut}(A_6) \cong S_6 \rtimes C_2$ and $\mathrm{Aut}(A_n) \cong S_n$ for $4 \leq n \neq 6$.*

Proof. Clearly, S_n acts on A_n by conjugation with kernel $\mathrm{C}_{S_n}(A_n) = 1$ (Lemma 6.2). This yields a monomorphism $\Psi\colon S_n \to \mathrm{Aut}(A_n)$. Let $H_i := \mathrm{Alt}(\{1, \dots, n\} \setminus \{i\})$ for $i = 1, \dots, n$. Let first $n \neq 6$. By Lemma 6.7, $\mathrm{Aut}(A_n)$ then acts on $\{H_1, \dots, H_n\}$. This yields a homomorphism $\Gamma\colon \mathrm{Aut}(A_n) \to S_n$. Let $f \in \mathrm{Ker}(\Gamma)$ and $\sigma \in A_n$. Then

$$H_{f(\sigma)i} = f(\sigma)H_i f(\sigma)^{-1} = f(\sigma)f(H_i)f(\sigma)^{-1} = f(\sigma H_i \sigma^{-1}) = f(H_{\sigma i}) = H_{\sigma i}$$

and $f(\sigma) = \sigma$. Hence Γ is injective. Thus Ψ and Γ are in fact isomorphisms. Let now $n = 6$.

Step 1: $\Psi(S_6) < \mathrm{Aut}(A_6)$.
Clearly, A_5 acts faithfully and transitively on $\mathrm{Syl}_5(A_5)$ by conjugation. This yields a monomorphism $f\colon A_5 \to \mathrm{Sym}(\mathrm{Syl}_5(A_5)) \cong S_6$. In the case $f(A_5) \not\subseteq A_6$ we would have $1 \neq f(A_5) \cap A_6 \lhd f(A_5)$, contradicting the simplicity of $f(A_5) \cong A_5$. Hence $f(A_5) \leq A_6$. Since f is transitive, we have $f(A_5) \neq H_i$ for $i = 1, \ldots, 6$. The action of A_6 on the cosets $A_6/f(A_5)$ yields a monomorphism $\varphi\colon A_6 \to S_6$. As before, $\varphi(A_6) = A_6$, i. e., $\varphi \in \mathrm{Aut}(A_6)$. Since $f(A_5)$ is the stabilizer of the trivial coset, we must have $\varphi(f(A_5)) = H_i$ for some $i \in \{1, \ldots, 6\}$. In particular, $\varphi \notin \Psi(S_6)$.

Step 2: $|\mathrm{Out}(A_6)| = 4$.
Let $\varphi, \psi \in \mathrm{Aut}(A_6) \setminus \Psi(S_6)$. Clearly, every automorphism maps conjugacy classes to conjugacy classes. By Remark 6.6, the set of 3-cycles is a conjugacy class C of A_6. If $\varphi(C) = C$, then one sees as in the proof of Lemma 6.7 that φ permutes the subgroups H_i. But then we would have $\varphi \in \Psi(S_6)$. Hence $\varphi(C) = \psi(C)$ must be the conjugacy class of the elements of cycle type $(3,3)$, since these are the only other elements of order 3. It follows that $(\varphi\psi)(C) = C$ and $\varphi\psi \in \Psi(S_6)$. This implies the claim.

Step 3: $\mathrm{Aut}(A_6) \cong S_6 \rtimes C_2$.
As a subgroup of index 2, $\Psi(S_6) \trianglelefteq \mathrm{Aut}(A_6)$. Hence it suffices to find an element $\varphi \in \mathrm{Aut}(A_6) \setminus \Psi(S_6)$ of order 2. Let first $\varphi \in \mathrm{Aut}(A_6) \setminus \Psi(S_6)$ be arbitrary and let $x := (1,2,3,4,5) \in A_6$. Then $\varphi(x)$ is also a 5-cycle and hence conjugate to x in S_6 (Remark 6.6). If we replace φ by a suitable element of the coset $\varphi\Psi(S_6)$, we may assume $\varphi(x) = x$. Hence there exists $y \in \mathrm{C}_{S_6}(x)$ with $\varphi^2 = \Psi(y)$. Since S_6 has exactly $6 \cdot 4!$ cycles of length 5, we have $|\mathrm{C}_{S_6}(x)| \leq 5$ and thus $y \in \mathrm{C}_{S_6}(x) = \langle x \rangle$. This shows $|\langle \varphi \rangle| \in \{2, 10\}$. Hence $\varphi^5 \in \mathrm{Aut}(A_6) \setminus \Psi(S_6)$ has order 2. □

Theorem 6.9. *For $4 \leq n \neq 6$, every extension by A_n splits.*

Proof. By Theorem 6.8, $\mathrm{Inn}(A_n) \cong A_n$ has a complement in $\mathrm{Aut}(A_n) \cong S_n$. The claim follows from Lemma 6.2 and Corollary 4.28. □

Remark 6.10. By Theorem 6.8, $\mathrm{Out}(A_6) \cong C_2^2$. The next theorem shows that Theorem 6.9 is false for $n = 6$.

Theorem 6.11. *The extension* $\mathrm{Aut}(A_6)$ *of* $\mathrm{Inn}(A_6) \cong A_6$ *does not split.*

Proof. Suppose that $A \leq \mathrm{Aut}(A_6)$ is a complement of $\mathrm{Inn}(A_6)$. Let $\Psi(S_6) \leq \mathrm{Aut}(A_6)$ be as in the proof of Theorem 6.8. Then $|\Psi(S_6) \cap A| = 2$. Let $\sigma \in S_6$ be an involution with $\alpha := \Psi(\sigma) \in A$. Since $\mathrm{Inn}(A_6) \cap A = 1$, σ has cycle type (2) or $(2,2,2)$. W. l. o. g. we may assume $\sigma \in \{(1,2), (1,2)(3,4)(5,6)\}$. Let $\beta \in A \setminus \Psi(S_6)$. As in the proof of Theorem 6.8, β interchanges the 3-cycles with the permutations of cycle type $(3,3)$. Since $|A| = 4$, A is abelian.

Case 1: $\sigma = (1,2)$.
We have

$$\beta((3,4,5)) = (\beta\alpha)((3,4,5)) = (\alpha\beta)((3,4,5)) = (1,2)\beta((3,4,5))(1,2),$$

but $(1,2)$ cannot centralize a permutation of cycle type $(3,3)$. Contradiction.

Case 2: $\sigma = (1,2)(3,4)(5,6)$.
Clearly, $(1,3,5)(2,4,6)$ is centralized by σ. The same calculation as in Case 1 with $(1,3,5)(2,4,6)$ instead of $(3,4,5)$ shows that σ centralizes a 3-cycle. This is also impossible. □

Remark 6.12. The exceptional behavior of $\mathrm{Aut}(A_6)$ can be explained by the (also exceptional) isomorphism $A_6 \cong \mathrm{PSL}(2,9)$ (Exercise 18).

```
G:=AlternatingGroup(6);;
A:=AutomorphismGroup(G);;
sub:=SubgroupsOfIndexTwo(A);;
List(sub,IdGroup);
IdGroup(SymmetricGroup(6));
IdGroup(PGL(2,9));
IdGroup(MathieuGroup(10));
```

Lemma 6.13. *For $n \in \mathbb{N}$ we have $|M(S_n)| \leq 2$.*

Proof. We use the presentation

$$S_n = \langle x_1, \ldots, x_{n-1} \mid x_i^2 = (x_i x_{i+1})^3 = [x_i, x_j] = 1 \text{ for } i < j-1 \rangle$$

from Theorem 2.18. Let F and N be as in Theorem 5.24. Let $\overline{F} := F/[F,N]$ and $\overline{x_i} := x_i[F,N] \in \overline{F}$ for $i = 1, \ldots, n-1$. Let $a_i := \overline{x_i}^2$, $b_i := (\overline{x_i x_{i+1}})^3$ and $c_{ij} := \overline{[x_i, x_j]}$ be elements in $\overline{N}$. Since $\overline{N} \leq \mathrm{Z}(\overline{F})$, these are generators of $\overline{N}$ (one does not need the normal closure). Since we can realize x_i by the transposition $(i, i+1)$, there exist $g \in F$, $y, z \in N$ with $\overline{g} x_i \overline{g}^{-1} = x_1 y$ and $\overline{g} x_j \overline{g}^{-1} = x_3 z$ (note that $i < j-1$). Since $\overline{N} \leq \mathrm{Z}(\overline{F})$, it follows that

$$c_{ij} = \overline{g} c_{ij} \overline{g}^{-1} = [\overline{g x_i g^{-1}}, \overline{g x_j g^{-1}}] = [\overline{x_1 y}, \overline{x_3 z}] = \overline{[x_1, x_3]} = c_{13}.$$

Moreover,

$$c_{13}^2 a_3 = (c_{13}\overline{x_3})^2 = (\overline{x_1 x_3 x_1^{-1}})^2 = a_3$$

and thus $c_{13}^2 = 1$. A similar calculation shows

$$b_i^2 a_i^{-1} a_{i+1}^{-2} = (b_i (\overline{x_{i+1} x_i x_{i+1}})^{-1})^2 = (\overline{x_i x_{i+1} x_i})^2 = a_i^2 a_{i+1}$$

and it follows that $b_i^2 = (a_i a_{i+1})^3$. We set $d_1 := a_1$ and $d_{i+1} := b_i (a_i a_{i+1})^{-1}$. Then $d_{i+1}^2 = a_i a_{i+1}$ and

$$\overline{N} = \langle d_1, \ldots, d_{n-1}, c_{13} \rangle$$

with $c_{13}^2 = 1$. By Theorem 5.24, the free part of $\overline{N}$ has rank $n-1$. Hence $d_1, \ldots, d_{n-1}$ must have infinite order. Thus the torsion part is $\langle c_{13} \rangle$. The claim now follows from Theorem 5.24. □

Lemma 6.14. *For $n \in \mathbb{N} \setminus \{6, 7\}$ we have $|M(A_n)| \leq 2$ and $|M(A_6)|, |M(A_7)| \leq 6$.*

Proof. The cases $n \leq 7$ can be handled with GAP or GT, Corollary 11.19. Let now $n \geq 8$ and

$$A_n \cong \langle x_1, \ldots, x_{n-2} \mid x_1^3 = x_2^2 = \ldots = x_{n-2}^2 = (x_i x_{i+1})^3 = (x_i x_j)^2 = 1 \text{ for } |j-i| > 1 \rangle$$

as in Theorem 2.19. We use the notation from Lemma 6.13 with $a_1 := \overline{x_1}^3$, $a_i := \overline{x_i}^2$ $(i \geq 2)$, $b_i := (\overline{x_i x_{i+1}})^3$, $c_i := (\overline{x_1 x_i})^2$ $(i \geq 3)$ and $c_{ij} := [\overline{x_i}, \overline{x_j}]$ for $2 \leq i < j-1$. As in Lemma 6.13, it follows that $c := c_{24} = c_{ij}$, $c^2 = 1$ and $b_i^2 = (a_i a_{i+1})^3$ for $i \geq 2$. For $z_{i+1} := b_i (a_i a_{i+1})^{-1}$ we therefore have $\langle b_i, a_i a_{i+1} \rangle = \langle z_{i+1} \rangle$ for $i = 2, \ldots, n-3$. For $i \geq 3$ we have

$$a_1 = (\overline{x_i^{-1} x_1 x_i})^3 = (a_i^{-1} \overline{x_1}^{-1} c_i)^3 = a_1^{-1} (a_i^{-1} c_i)^3$$

and $\boxed{(c_i a_i^{-1})^3 = a_1^2.}$ We set $z_1 := a_1 (c_3 a_3^{-1})^{-1}$ and obtain $\langle a_1, c_3 a_3^{-1} \rangle = \langle z_1 \rangle$.

From $c_4 = \overline{x_1}^{-1}(\overline{x_1x_4})^2\overline{x_1} = (\overline{x_4x_1})^2$ and $c = a_2a_4[\overline{x_2}^{-1}, \overline{x_4}^{-1}]a_2^{-1}a_4^{-1} = [\overline{x_2}^{-1}, \overline{x_4}^{-1}]$ it follows that

$$\overline{x_4x_1x_2x_4^{-1}} = c_4\overline{x_1^{-1}x_4^{-1}x_2x_4}a_4^{-1} = c_4\overline{x_1^{-1}x_2}ca_4^{-1} = \overline{x_2(x_1x_2)^{-1}x_2^{-1}}a_2c_4ca_4^{-1}.$$

The third power yields

$$b_1 = b_1^{-1}a_2^3c_4^3ca_4^{-3}.$$

Since $(c_4a_4^{-1})^3 = a_1^2$, we obtain $b_1^2 = a_1^2a_2^3c$. With $z_2 := b_1a_1^{-1}a_2^{-1}c$ we have $z_2^2 = a_2c$ and $z_2^3 = a_1^{-1}b_1$.

For $i \geq 5$ we have

$$\overline{x_3x_1x_ix_3^{-1}} = c_3\overline{x_1^{-1}x_3^{-1}x_ix_3}a_3^{-1} = c_3\overline{x_1^{-1}x_i}c_{3i}a_3^{-1}$$

and $c_i = (c_3a_3^{-1})^2(\overline{x_1^{-1}x_i})^2$ after squaring. Since $\overline{x_1^{-1}x_i} = a_i\overline{x_i}^{-1}(\overline{x_1x_i})^{-1}\overline{x_i}$, it follows that $(c_ia_i^{-1})^2 = (c_3a_3^{-1})^2$. On the other hand, we also have $(c_ia_i^{-1})^3 = a_1^2 = (c_3a_3^{-1})^3$. This shows $c_ia_i^{-1} = c_3a_3^{-1}$ for $i \geq 5$. The same calculation with index 4 instead of 3 yields $c_4a_4^{-1} = c_6a_6^{-1}$. Hence $\langle z_1\rangle = \langle a_1, c_ia_i^{-1}\rangle$ for $i = 3, \ldots, n-2$.

Altogether,

$$\overline{N} = \langle a_i, b_i, c_i, c\rangle = \langle z_1, \ldots, z_{n-2}, c\rangle$$

and it follows that $|M(A_n)| \leq 2$. □

Theorem 6.15. *We have*

$$M(S_n) = \begin{cases} 1 & \text{if } n \leq 3, \\ C_2 & \text{if } n \geq 4. \end{cases}$$

Proof. W. l. o. g. we may assume $n \geq 4$. By Lemma 6.13 it suffices to construct a proper Schur extension of S_n. Let

$$\widehat{S}_n := \langle x_1, \ldots, x_{n-1}, z \mid z^2 = 1,\ x_i^2 = (x_ix_{i+1})^3 = [x_i, x_j] = z \text{ for } i < j-1\rangle.$$

By von Dyck's theorem, there exists an epimorphism $\widehat{S}_n \to S_n$ with kernel $\langle z\rangle$. Since $x_i^2 = [x_i, x_j] = z$, we have $z \in \mathrm{Z}(\widehat{S}_n) \cap \widehat{S}_n'$. However, we might have $z = 1$. We define matrices in $\mathrm{GL}(2^n, \mathbb{C})$ as iterated Kronecker products of 2×2 matrices:

$$A := \begin{pmatrix} 0 & \mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad B := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad C := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$I := 1_{2^n}, \qquad M_{2k-1} := C^{\otimes(n-k)} \otimes A \otimes 1_2^{\otimes(k-1)}, \qquad M_{2k} := C^{\otimes(n-k)} \otimes B \otimes 1_2^{\otimes(k-1)}$$

for $k = 1, \ldots, n$. Additionally, let $M_0 := 0_{2^n}$. Since $A^2 = B^2 = -C^2 = [A, B] = [A, C] = [B, C] = -1_2$, we have $M_k^2 = [M_k, M_l] = -I$ for $k \neq l$. We now set

$$X_k := \frac{1}{\sqrt{2k}}\big(\sqrt{k+1}M_k - \sqrt{k-1}M_{k-1}\big)$$

for $k = 1, \dots, n-1$. It follows that

$$\begin{aligned}
X_k^2 &= -\frac{1}{2k}\big((k+1)I + (k-1)I\big) = -I \qquad (\Longrightarrow X_k \in \mathrm{GL}(2^n, \mathbb{C})),\\
X_k X_l &= -X_l X_k \qquad (k < l-1),\\
X_{k+1}X_k &= \frac{1}{2\sqrt{k^2+k}}\Big(\sqrt{(k+2)(k+1)}M_{k+1}M_k - \sqrt{(k+2)(k-1)}M_{k+1}M_{k-1}\\
&\quad + \sqrt{k^2+k}I + \sqrt{k(k-1)}M_kM_{k-1}\Big)\\
&= I - X_kX_{k+1},\\
X_kX_{k+1}X_k &= X_k(I - X_kX_{k+1}) = X_k + X_{k+1} = (I - X_kX_{k+1})X_{k+1} = X_{k+1}X_kX_{k+1}\\
(X_kX_{k+1})^3 &= X_k(X_{k+1}X_kX_{k+1})X_kX_{k+1} = X_kX_kX_{k+1}X_kX_kX_{k+1} = (-I)^3 = -I.
\end{aligned}$$

By von Dyck's theorem, there exists an epimorphism $\Gamma\colon \widehat{S}_n \to \langle X_1, \dots, X_{n-1}\rangle$ with $\Gamma(z) = -I$. Hence $z \neq 1$ in $\widehat{S}_n$. □

Remark 6.16. By Theorem 5.19, S_n has at most $\gcd(|S_n/S_n'|, |M(S_n)|) = 2$ maximal Schur extensions. Besides $\widehat{S}_n$ there is also the Schur extension

$$\widetilde{S}_n := \langle x_1, \dots, x_{n-1}, z \mid z^2 = 1, x_i^2 = (x_ix_{i+1})^3 = 1, [x_i, x_j] = z \text{ for } i < j-1\rangle.$$

One easily shows $\widetilde{S}_4 \cong \mathrm{GL}(2,3)$. More generally, let q be an odd prime power, $\mathbb{F}_q^\times = \langle\zeta\rangle$ and $x := \left(\begin{smallmatrix} 0 & \zeta \\ -1 & 0\end{smallmatrix}\right) \in \mathrm{GL}(2,q)$. Then x induces by conjugation an outer automorphism α of $\mathrm{SL}(2,q)$ of order 2. We have $\mathrm{SL}(2,5) \rtimes \langle\alpha\rangle \cong \widetilde{S}_5$. By a similar construction one arrives at $\widetilde{S}_6 \cong \mathrm{SL}(2,9) \rtimes C_2$ (cf. Exercise 18). Let now $\zeta \in \mathbb{F}_{q^2}^\times$ be of order $2(q-1)$ and $x := \mathrm{diag}(\zeta, \zeta^{-1}) \in \mathrm{SL}(2,q^2)$. Then $K(q) := \mathrm{SL}(2,q)\langle x\rangle \leq \mathrm{SL}(2,q^2)$ is a nonsplit extension with $K(3) \cong \widehat{S}_4$ and $K(5) \cong \widehat{S}_5$ (without proof). An outer automorphism of S_6 yields an isomorphism $\widehat{S}_6 \cong \widetilde{S}_6$. For $n \neq 6$, on the other hand, we have $\widehat{S}_n \ncong \widetilde{S}_n$ (Exercise 21 and Remark 5.22). For larger n, the Schur extensions of S_n are not "known" groups.

```
S1:=SchurCoverOfSymmetricGroup(5,3,1); #Ŝ5 as a matrix group over F3
S2:=SchurCoverOfSymmetricGroup(5,3,-1); #S̃5
IdGroup(S1); IdGroup(S2);
IsomorphismGroups(SchurCoverOfSymmetricGroup(4,3,-1),GL(2,3));
```

Theorem 6.17. *We have*

$$M(A_n) = \begin{cases} C_6 & \text{if } n \in \{6,7\},\\ C_2 & \text{if } n = 4,5,8,9,\dots. \end{cases}$$

Proof. One easily checks that $\mathrm{SL}(2,3)$ is a Schur extension of A_4 (cf. Exercise 15). Hence $M(A_4) \cong C_2$. Let now $n \geq 5$, $\widehat{A}_n := \widehat{S}_n'$ and $Z := \mathrm{Z}(\widehat{S}_n) \leq \mathrm{Z}(\widehat{A}_n)$. Then

$$\begin{aligned}
\widehat{A}_n/Z &\cong (\widehat{S}_n/Z)' \cong S_n' \cong A_n,\\
\widehat{A}_n'Z/Z &= (\widehat{A}_n/Z)' \cong A_n' \cong A_n.
\end{aligned}$$

It follows that $|\widehat{S}_n/\widehat{A}_n'| = |\widehat{S}_n/\widehat{A}_n||\widehat{A}_n/\widehat{A}_n'| \leq 4$ and $Z \leq \widehat{S}_n' \leq \widehat{A}_n'$. Hence $\widehat{A}_n$ is a Schur extension of A_n and $2 \mid |M(A_n)|$ for $n \geq 5$. For $n \notin \{6,7\}$, $\widehat{A}_n$ is the universal Schur extension of A_n by Remark 5.22.

It therefore suffices to construct Schur extensions $\widehat{A}_6$ and $\widehat{A}_7$ with $\mathrm{Z}(\widehat{A}_6) \cong \mathrm{Z}(\widehat{A}_7) \cong C_3$. We generate $\widehat{A}_6$ by monomial matrices in $\mathrm{GL}(6,4)$. Let $\mathbb{F}_4^\times = \langle \zeta \rangle$, $\overline{\zeta} := \zeta^{-1}$ and $(a_1, \ldots, a_6; \sigma) := (a_i \delta_{i\sigma(j)})_{ij} \in \mathrm{GL}(6,4)$ for $a_1, \ldots, a_6 \in \mathbb{F}_4$ and $\sigma \in S_6$. We define

$$\begin{aligned} x_1 &:= (\overline{\zeta}, 1, \zeta, 1, \zeta, \overline{\zeta}; (145)(263)), \\ x_2 &:= (1, 1, 1, 1, \overline{\zeta}, \zeta; (13)(56)), \\ x_3 &:= (1, \zeta, \overline{\zeta}, \zeta, \overline{\zeta}, 1; (23)(45)), \\ x_4 &:= (\overline{\zeta}, 1, \zeta, 1, \zeta, \overline{\zeta}; (15)(36)). \end{aligned}$$

Let $G := \langle x_1, x_2, x_3, x_4 \rangle$. The permutations involved clearly generate A_6. A calculation shows

$$x_1^3 = x_2^2 = x_3^2 = x_4^2 = (x_1x_2)^3 = (x_2x_3)^3 = (x_3x_4)^3 = (x_1x_4)^2 = (x_2x_4)^2 = 1$$

and $z := (x_1x_3)^2 = \zeta 1_6 \in \mathrm{Z}(G)$. Moreover, $z = x_1x_3x_1x_3 = x_1x_3x_4x_1^{-1}x_4x_3 = [x_1, x_3x_4] \in G'$. By Moore, G is a Schur extension of A_6 with $\mathrm{Z}(G) = \langle z \rangle \cong C_3$. We now add

$$x_5 := \begin{pmatrix} 1 & \overline{\zeta} & 1 & \overline{\zeta} & 1 & . \\ \zeta & 1 & \zeta & . & \zeta & 1 \\ 1 & \overline{\zeta} & 1 & \overline{\zeta} & . & \overline{\zeta} \\ \zeta & . & \zeta & 1 & \zeta & 1 \\ 1 & \overline{\zeta} & . & \overline{\zeta} & 1 & \overline{\zeta} \\ . & 1 & \zeta & 1 & \zeta & 1 \end{pmatrix}.$$

Another calculation yields $x_5^2 = (x_2x_5)^2 = (x_3x_5)^2 = (x_4x_5)^3 = 1$ and $(x_1x_5)^2 = z$. Hence $\langle x_1, \ldots, x_5 \rangle$ is the desired Schur extension $\widehat{A}_7$. □

Remark 6.18.

(i) The Schur extension $\widehat{A}_6$ with $|\mathrm{Z}(\widehat{A}_6)| = 3$ constructed in the proof can be embedded into $C_3 \wr A_6$ as an imprimitive permutation group of degree 18. One calls $\widehat{A}_6$ the VALENTINER *group*.

(ii) Suppose that the (universal) Schur extension $\widehat{A}_7$ has already been constructed. Then there exists a subgroup $Z \leq H \leq \widehat{A}_7$ with $H/Z \cong A_6$. For $P \in \mathrm{Syl}_3(H)$ we also have $P \in \mathrm{Syl}_3(\widehat{A}_7)$. By Taunt (GT, Theorem 7.15), P is nonabelian and hence $Z = [P, P] \leq H'$. Thus H is a Schur extension of A_6. Hence it suffices to construct $\widehat{A}_7$. We have already done this with GAP in Remark 5.18. The group can also be constructed in other ways:

```
PerfectGroup(IsPermGroup,15120); #unique perfect group of this order
LoadPackage("atlasrep",false);
AtlasGroup("6.A7"); #requires internet access on first use
DoubleCoverOfAlternatingGroup(10,3); #\widehat{A_{10}} as a matrix group over F_3 with Z(\widehat{A_{10}}) ≅ C_2
```

(iii) Blackburn constructed all extensions of elementary abelian groups by S_n and A_n.

7 Symplectic Groups

Remark 7.1. The most complicated simple groups are the groups of *Lie type*. These are certain matrix groups over finite fields, of which we have already constructed the projective special linear groups $\mathrm{PSL}(n,q)$ in group theory. In this chapter, the (projective) symplectic groups $\mathrm{PSp}(2n,q)$ are introduced, which exist only in even dimension. For $n = 1$ we have $\mathrm{PSp}(2,q) = \mathrm{PSL}(2,q)$. For $n \geq 2$ one obtains "new" simple groups. The proof uses Iwasawa's lemma. In contrast to $\mathrm{PSL}(n,q)$, this time we try to argue "coordinate-free".

Definition 7.2. Let K be a field and V a finite-dimensional K-vector space. Let $\beta\colon V \times V \to K$, $(v, w) \mapsto [v, w]$ be a bilinear form with the following properties:

- (nondegenerate) For all $v \in V \setminus \{0\}$ there exists $w \in V$ with $[v, w] \neq 0$.
- (alternating) For all $v \in V$ we have $[v, v] = 0$.

With respect to $[.,.]$, V is called a *symplectic space*. For $v \in V$ and $U \subseteq V$ let $v^\perp := \{w \in V : [v, w] = 0\} \leq V$ and $U^\perp := \bigcap_{u \in U} u^\perp \leq V$.

Remark 7.3. In the following, V always denotes a symplectic space. From $[v, v] = 0$ we obtain the *antisymmetry* $[u, v] = [u + v, u + v] - [v, u] = -[v, u]$ for $u, v \in V$. For $\operatorname{char} K \neq 2$ the notions "alternating" and "antisymmetric" are equivalent, since $[v, v] = -[v, v] = 0$. For $\operatorname{char} K = 2$, symmetry and antisymmetry are identical.

Theorem 7.4. *For $U \leq V$ we have $\dim V = \dim U + \dim U^\perp$ (but not necessarily $U \cap U^\perp = 0$).*

Proof. We extend a basis $b_1, \ldots, b_k$ of U to a basis $b_1, \ldots, b_n$ of V. With respect to this basis we identify V with K^n. For the Gram matrix $B = ([b_i, b_j])$ of β we then have $[v, w] = vBw^{\mathrm{t}}$ for $v, w \in V$. Since β is nondegenerate, B is invertible (alternating is not needed for the proof). The first k rows of B form a matrix A of rank $k = \dim U$. Now $U^\perp$ is the solution space of the homogeneous system of equations $Ax = 0$. From linear algebra it follows that $\dim U^\perp = n - k = \dim V - \dim U$. □

Theorem 7.5. *There exists a basis $b_1, \ldots, b_n, c_1, \ldots, c_n$ of V with $[b_i, b_j] = 0 = [c_i, c_j]$ and $[b_i, c_j] = \delta_{ij}$ for $i, j = 1, \ldots, n$. In particular, $\dim V = 2n$ is even (cf. GT, Exercise 70).*

Proof. Induction on $\dim V$. Let $b_1 \in V \setminus \{0\}$. Since β is nondegenerate, there exists $c_1 \in V$ with $[b_1, c_1] \neq 0$. After scaling we have $[b_1, c_1] = 1$. Now let $U := \langle b_1, c_1 \rangle$. Clearly, $U \cap U^\perp = 0$ and Theorem 7.4 shows $V = U \oplus U^\perp$. For $u \in U^\perp$ there exists $v \in V$ with $[u, v] \neq 0$. Writing $v = v_1 + v_2$ with $v_1 \in U$ and $v_2 \in U^\perp$, we obtain $[u, v_2] = [u, v] \neq 0$. Hence the restriction of β to $U^\perp$ is also nondegenerate and alternating. The claim now follows by induction. □

Definition 7.6.

- We call $(v, w) \in V^2$ with $[v, w] = 1$ a *hyperbolic pair*. A basis as in Theorem 7.5 is called *symplectic*.
- We call

$$\begin{aligned} \operatorname{Sp}(V) &:= \big\{ f \in \operatorname{GL}(V) : [f(v), f(w)] = [v, w] \big\} \leq \operatorname{GL}(V), \\ \operatorname{PSp}(V) &:= \operatorname{Sp}(V)/\operatorname{Z}(\operatorname{Sp}(V)) \end{aligned}$$

  the (*projective*) *symplectic group* of V.

Remark 7.7.

(i) The Gram matrix of β with respect to a symplectic basis is $B = \left(\begin{smallmatrix} 0 & 1_n \\ -1_n & 0 \end{smallmatrix}\right)$. For $v, w \in K^{2n}$ with respect to this basis we have $[v, w] = vBw^{\mathrm{t}}$. This shows

$$\operatorname{Sp}(2n, K) := \{A \in \operatorname{GL}(2n, K) : A^{\mathrm{t}} B A = B\} \cong \operatorname{Sp}(V).$$

In particular, the isomorphism type of $\operatorname{Sp}(V)$ does not depend on β. For finite fields we set $\operatorname{Sp}(2n, q) := \operatorname{Sp}(2n, \mathbb{F}_q)$.

(ii) For $A = \left(\begin{smallmatrix} A_1 & A_2 \\ A_3 & A_4 \end{smallmatrix}\right) \in \mathrm{Sp}(2n, K)$ we have

$$\begin{pmatrix} A_1^{\mathrm{t}} & A_3^{\mathrm{t}} \\ A_2^{\mathrm{t}} & A_4^{\mathrm{t}} \end{pmatrix} \begin{pmatrix} A_3 & A_4 \\ -A_1 & -A_2 \end{pmatrix} = A^{\mathrm{t}} B A = B = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix},$$

so $A_1^{\mathrm{t}} A_3 = A_3^{\mathrm{t}} A_1$, $A_2^{\mathrm{t}} A_4 = A_4^{\mathrm{t}} A_2$ and $A_1^{\mathrm{t}} A_4 = 1_n + A_3^{\mathrm{t}} A_2$. For $n = 1$ the first two equations are trivial and the third says $\det(A) = 1$. This shows $\mathrm{Sp}(2, K) = \mathrm{SL}(2, K)$. For arbitrary n we have $\left(\begin{smallmatrix} A_1 & 0 \\ 0 & A_1^{-\mathrm{t}} \end{smallmatrix}\right) \in \mathrm{Sp}(2n, K)$ for all $A_1 \in \mathrm{GL}(n, K)$. Hence $\mathrm{GL}(n, K)$ is isomorphic to a subgroup of $\mathrm{Sp}(2n, K)$. For symmetric $S \in K^{n \times n}$ we also have $\left(\begin{smallmatrix} 1_n & S \\ 0 & 1_n \end{smallmatrix}\right), \left(\begin{smallmatrix} 1_n & 0 \\ S & 1_n \end{smallmatrix}\right) \in \mathrm{Sp}(2n, K)$.

(iii) We can also arrange the Gram matrix of β in the form $B = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) \oplus \ldots \oplus \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Then we obtain $\mathrm{Sp}(2k, K) \times \mathrm{Sp}(2(n-k), K) \leq \mathrm{Sp}(2n, K)$ for $0 < k < n$.

Lemma 7.8. *Let $A \in K^{n \times n}$ with $AS = SA$ for all symmetric matrices $S \in K^{n \times n}$. Then $A = \lambda 1_n$ for some $\lambda \in K$.*

Proof. For $S = E_{ii}$ (1 at position (i, i), zeros elsewhere) the equation $AE_{ii} = E_{ii}A$ shows that the i-th column and the i-th row of A vanish outside the diagonal. Since this holds for $i = 1, \ldots, n$, A is a diagonal matrix. For $S = E_{ij} + E_{ji}$ with $i \neq j$, the equation $AS = SA$ now implies $A_{ii} = A_{jj}$. Hence $A \in K 1_n$. □

Lemma 7.9. *We have $\mathrm{Z}(\mathrm{Sp}(V)) = \mathrm{Sp}(V) \cap K^\times \mathrm{id} = \langle -\mathrm{id} \rangle$.*

Proof. Since $-\mathrm{id} \in \mathrm{Sp}(V) \cap K^\times \mathrm{id} \subseteq \mathrm{Z}(\mathrm{Sp}(V))$, we only need to show $\mathrm{Z}(\mathrm{Sp}(2n, K)) \subseteq \langle -1_{2n} \rangle$. For $n = 1$ we have $\mathrm{Z}(\mathrm{Sp}(2, K)) = \mathrm{Z}(\mathrm{SL}(2, K)) = \langle -1_2 \rangle$ (in the case $\mathrm{char}\, K = 2$ we have $-1_2 = 1_2$). So let $n \geq 2$. We write $A = \left(\begin{smallmatrix} A_1 & A_2 \\ A_3 & A_4 \end{smallmatrix}\right) \in \mathrm{Z}(\mathrm{Sp}(2n, K))$ as in Remark 7.7. For symmetric $S \in K^{n \times n}$ the matrices

$$U_S := \begin{pmatrix} 1_n & S \\ 0 & 1_n \end{pmatrix} \quad \text{and} \quad L_S := \begin{pmatrix} 1_n & 0 \\ S & 1_n \end{pmatrix}$$

are contained in $\mathrm{Sp}(2n, K)$. From $AU_S = U_S A$ we obtain, by comparing blocks, $SA_3 = 0$ and $A_1 S = SA_4$. Similarly, $AL_S = L_S A$ yields the equations $A_2 S = 0$ and $A_4 S = SA_1$. For $S = 1_n$ we get $A_2 = A_3 = 0$ and $A_1 = A_4$. Hence $A_1 S = SA_1$ for all symmetric S. Lemma 7.8 shows $A_1 = A_4 = \lambda 1_n$ with $\lambda \in K^\times$. Since $A_2 = A_3 = 0$, Remark 7.7 finally yields $1_n = A_1^{\mathrm{t}} A_4 = \lambda^2 1_n$, so $\lambda = \pm 1$ and $A = \pm 1_{2n}$. □

Theorem 7.10. *We have*

$$\boxed{\begin{aligned} |\mathrm{Sp}(2n, q)| &= q^{n^2} \prod_{k=1}^{n} (q^{2k} - 1), \\ |\mathrm{PSp}(2n, q)| &= \frac{1}{\gcd(q-1, 2)} q^{n^2} \prod_{k=1}^{n} (q^{2k} - 1). \end{aligned}}$$

Proof. By Lemma 7.9 the second equation follows from the first. Every $f \in \mathrm{Sp}(V)$ maps a symplectic basis to a symplectic basis. Conversely, exactly one $f \in \mathrm{Sp}(V)$ can be defined in this way. Hence $|\mathrm{Sp}(V)|$ is the number of symplectic bases of V. For $b_1 \in V \setminus \{0\}$ there are $q^{2n} - 1$ possibilities. The vector c_1 defines exactly one coset $c_1 + b_1^\perp$. So there are $|b_1^\perp| = q^{2n-1}$ possibilities for c_1. In the case $n = 1$ we therefore have $|\mathrm{Sp}(2, q)| = q(q^2 - 1)$ as claimed. Now let $n \geq 2$ and $U := \langle b_1, c_1 \rangle$. As in

Theorem 7.5, $V = U \oplus U^\perp$ and $b_2, \dots, b_n, c_2, \dots, c_n$ form a symplectic basis of $U^\perp$. Induction on n yields

$$|\mathrm{Sp}(2n,q)| = q^{2n-1}(q^{2n}-1)|\mathrm{Sp}(2n-2,q)| = q^{2n-1+(n-1)^2}\prod_{k=1}^{n}(q^{2k}-1) = q^{n^2}\prod_{k=1}^{n}(q^{2k}-1). \qquad \square$$

Definition 7.11. For $\lambda \in K$ and $v \in V$ the linear map

$$t_{\lambda,v}\colon V \to V, \qquad x \mapsto x + \lambda[x,v]v$$

is called a (*symplectic*) *transvection*.

Remark 7.12. For $\lambda, \mu \in K$ and $v, x \in V$ we have

$$t_{\lambda,v}(t_{\mu,v}(x)) = t_{\lambda,v}(x + \mu[x,v]v) = x + \mu[x,v]v + \lambda[x + \mu[x,v]v, v]v = t_{\lambda+\mu,v}(x).$$

In particular, $t_{\lambda,v}$ is invertible with $t_{\lambda,v}^{-1} = t_{-\lambda,v}$. Because of

$$[t_{\lambda,v}(x), t_{\lambda,v}(y)] = [x,y] + \lambda[x,v][v,y] + \lambda[y,v][x,v] + \lambda^2[x,v][y,v][v,v] = [x,y]$$

we have $t_{\lambda,v} \in \mathrm{Sp}(V)$. Clearly, $v^\perp$ is the fixed-point space (eigenspace for the eigenvalue 1) of $t_{\lambda,v}$. For a hyperbolic pair (v,w) we have $t_{\lambda,v}(w) = w - \lambda v \in w + v^\perp$. This shows $\det(t_{\lambda,v}) = 1$. For $g \in \mathrm{Sp}(V)$ we have

$$(g t_{\lambda,v} g^{-1})(x) = g(g^{-1}(x) + \lambda[g^{-1}(x), v]v) = x + \lambda[x, g(v)]g(v) = t_{\lambda,g(v)}(x).$$

For $\mu \in K$ we have

$$t_{\lambda,\mu v}(x) = x + \lambda[x,\mu v]\mu v = x + \lambda\mu^2[x,v]v = t_{\lambda\mu^2,v}(x).$$

Lemma 7.13. *The symplectic transvections generate* $\mathrm{Sp}(V)$. *In particular,* $\mathrm{Sp}(V) \le \mathrm{SL}(V)$.

Proof. By Theorem 7.10 it suffices to show that $S := \langle t_{\lambda,v} : \lambda \in K, v \in V\rangle$ acts transitively on the set Ω of all symplectic bases. Let $v, w \in V \setminus \{0\}$. If $\lambda := [v,w] \neq 0$, then we have

$$t_{\lambda^{-1},v-w}(v) = v + \lambda^{-1}[v, v-w](v-w) = v - (v-w) = w.$$

Otherwise choose $x \in V \setminus (v^\perp \cup w^\perp)$. Now we can realize $v \mapsto x \mapsto w$ by two transvections. Hence S is transitive on $V \setminus \{0\}$.

Now let $u, v, w \in V$ with $[u,v] = [u,w] = 1$. Then $v - w \in u^\perp$. In the case $\lambda := [v,w] \neq 0$ we have $t_{\lambda^{-1},v-w}(u,v) = (u,w)$. Otherwise set $x := u + v$. Then $[u,v] = [u,x] = [u,w] = 1$ and $[v,x] \neq 0 \neq [x,w]$. We can therefore realize $(u,v) \mapsto (u,x) \mapsto (u,w)$ by two transvections. Thus S is transitive on the set of hyperbolic pairs.

Finally, let $b_1, \dots, b_n, c_1, \dots, c_n$ and $b_1', \dots, b_n', c_1', \dots, c_n'$ be symplectic bases of V. By what has already been shown, we may assume $(b_1', c_1') = (b_1, c_1)$. Then

$$b_2, \dots, b_n, c_2, \dots, c_n \quad \text{and} \quad b_2', \dots, b_n', c_2', \dots, c_n'$$

are symplectic bases of $\langle b_1, c_1\rangle^\perp$. By induction there exists a product of symplectic transvections on $\langle b_1, c_1\rangle^\perp$ that maps $(b_2, \dots, b_n, c_2, \dots, c_n)$ to $(b_2', \dots, b_n', c_2', \dots, c_n')$. For $x \in \langle b_1, c_1\rangle^\perp$, the vectors b_1 and c_1 are fixed by $t_{\lambda,x}$. Hence the transvections on $\langle b_1, c_1\rangle^\perp$ can be extended to $V = \langle b_1, c_1\rangle^\perp \oplus \langle b_1, c_1\rangle$ by acting trivially on $\langle b_1, c_1\rangle$. So we can map $b_1, \dots, b_n, c_1, \dots, c_n$ to $b_1', \dots, b_n', c_1', \dots, c_n'$ by means of S. $\square$

Theorem 7.14. *For every finite field K, $\mathrm{PSp}(V)$ acts faithfully and primitively on the set Ω of 1-dimensional subspaces of V.*

Proof. For $\mathrm{Sp}(2,K) = \mathrm{SL}(2,K)$ the claim is known. So let $n \geq 2$ and $q := |K|$. By GT, Lemma 10.7, $G := \mathrm{Sp}(V) \leq \mathrm{SL}(V)$ acts on Ω with kernel $\mathrm{Sp}(V) \cap K^\times \mathrm{id} = \langle -\mathrm{id} \rangle$. Hence $\overline{G} := \mathrm{PSp}(V)$ acts faithfully on Ω. By the proof of Theorem 7.10, $\overline{G}$ acts transitively on Ω. Let $U := Ku \in \Omega$ and let G_U be the stabilizer of U in G. Let (u, u') be a hyperbolic pair. As in the proof of Lemma 7.13, $G_u \subseteq G_U$ acts transitively on $u' + u^\perp$. Since every vector in $u' + u^\perp$ spans a different subspace, G_U has an orbit of length $\geq q^{2n-1}$ on Ω. Let $v, w \in u^\perp \setminus U$. In the case $[v, w] \neq 0$ we can map (u, v) to (u, w) as in the proof of Lemma 7.13. Now let $[v, w] = 0$. Because of

$$(v^\perp)^\perp = \langle v \rangle \neq U = (u^\perp)^\perp \neq (w^\perp)^\perp$$

we have $v^\perp \neq u^\perp \neq w^\perp$. Let $x \in u^\perp \setminus v^\perp$ and $y \in u^\perp \setminus w^\perp$. In the case $u^\perp \subseteq v^\perp \cup w^\perp$ we have $x \in w^\perp$, $y \in v^\perp$ and obtain the contradiction $x + y \in u^\perp \setminus (v^\perp \cup w^\perp)$. This shows $u^\perp \nsubseteq v^\perp \cup w^\perp$. So let $x \in u^\perp \setminus (v^\perp \cup w^\perp)$. Now we can realize $(u, v) \mapsto (u, x) \mapsto (u, w)$ by two transvections. Hence G_u and G_U are transitive on $u^\perp \setminus U$. This yields an orbit of length $\frac{q^{2n-1}-q}{q-1} < q^{2n-1}$. Because of

$$|\Omega \setminus \{U\}| = \frac{q^{2n} - q}{q - 1} = \frac{q^{2n-1} - q}{q - 1} + q^{2n-1}$$

we see that G_U has two orbits on $\Omega \setminus \{U\}$ with lengths q^{2n-1} and $\frac{q^{2n-1}-q}{q-1}$.

Suppose that G is imprimitive on Ω with block $\Delta \ni U$. Then one of the two nontrivial orbits of G_U must lie in Δ. However, $q^{2n-1} + 1$ and $\frac{q^{2n-1}-q}{q-1} + 1 = \frac{q^{2n-1}-1}{q-1}$ are not divisors of $|\Omega| = \frac{q^{2n}-1}{q-1} = q^{2n-1} + q^{2n-2} + \ldots + 1$. This contradiction shows that G and $\overline{G}$ act primitively. □

Lemma 7.15. *For $n \geq 2$ and $(n, q) \neq (2, 2)$, $\mathrm{Sp}(2n, q)$ is perfect.*

Proof. By Lemma 7.13 it suffices to show that the transvections are commutators. First let $q \geq 4$. Let $v \in V \setminus \{0\}$ and $\lambda \in K^\times$. Choose $\mu \in K^\times \setminus \{\pm 1\}$ and set $\alpha := \lambda(1 - \mu^2)^{-1}$. As is well known, there exists $g \in \mathrm{Sp}(V)$ with $g(v) = \mu v$. By Remark 7.12 we have

$$\mathrm{Sp}(V)' \ni [t_{\alpha,v}, g] = t_{\alpha,v} g t_{-\alpha,v} g^{-1} = t_{\alpha,v} t_{-\alpha,g(v)} = t_{\alpha,v} t_{-\alpha\mu^2,v} = t_{\alpha(1-\mu^2),v} = t_{\lambda,v}.$$

Now let $q = 3$ and let $b_1, \ldots, b_n, c_1, \ldots, c_n$ be a symplectic basis of V. We define $g, h \in \mathrm{Sp}(V)$ by

$$\begin{array}{llll} g(b_1) := b_1 + b_2, & g(c_1) := c_2, & g(b_2) := b_1, & g(c_2) := c_1 - c_2, \\ h(b_1) := b_1 - c_1 + c_2, & h(c_1) := c_1, & h(b_2) := b_2 + c_1, & h(c_2) := c_2 \end{array}$$

and $g(b_i) = h(b_i) = b_i$ as well as $g(c_i) = h(c_i) = c_i$ for $i \geq 3$. We have

$$\begin{aligned} [g,h](b_1) &= ghg^{-1}(b_1 + c_1 - c_2) = gh(b_2 + c_2) = g(b_2 + c_1 + c_2) = b_1 + c_1, \\ [g,h](c_1) &= ghg^{-1}(c_1) = gh(c_1 + c_2) = g(c_1 + c_2) = c_1, \\ [g,h](b_2) &= ghg^{-1}(b_2 - c_1) = gh(b_1 - b_2 - c_1 - c_2) = g(b_1 - b_2) = b_2, \\ [g,h](c_2) &= ghg^{-1}(c_2) = gh(c_1) = g(c_1) = c_2, \end{aligned}$$

i. e., $t_{1,c_1} = [g, h] \in \mathrm{Sp}(V)'$ and $t_{-1,c_1} = t_{1,c_1}^{-1} \in \mathrm{Sp}(V)'$. Since c_1 is arbitrary, the claim follows.

Finally, let $q = 2$ and $n \geq 3$. This time we define

$$\begin{aligned}
&g(b_1) := b_1 + b_3, && g(c_1) := c_3, && g(b_2) := b_1, && g(c_2) := c_1 + c_3,\\
&g(b_3) := b_2, && g(c_3) := c_2, && h(b_1) := b_1 + c_2, && h(c_1) := c_1,\\
&h(b_2) := b_2 + c_1 + c_2 + c_3, && h(c_2) := c_2, && h(b_3) := b_3 + c_2 + c_3, && h(c_3) := c_3
\end{aligned}$$

and $g(b_i) = h(b_i) = b_i$ as well as $g(c_i) = h(c_i) = c_i$ for $i \geq 4$. We have

$$\begin{aligned}
[g,h](b_1) &= ghg^{-1}(b_1 + c_2) = gh(b_2 + c_3) = g(b_2 + c_1 + c_2) = b_1 + c_1,\\
[g,h](c_1) &= ghg^{-1}(c_1) = gh(c_1 + c_2) = g(c_1 + c_2) = c_1,\\
[g,h](b_2) &= ghg^{-1}(b_2 + c_1 + c_2 + c_3) = gh(b_3 + c_2 + c_3) = g(b_3) = b_2,\\
[g,h](c_2) &= ghg^{-1}(c_2) = gh(c_3) = g(c_3) = c_2,\\
[g,h](b_3) &= ghg^{-1}(b_3 + c_2 + c_3) = gh(b_1 + b_2 + c_1 + c_3) = g(b_1 + b_2) = b_3,\\
[g,h](c_3) &= ghg^{-1}(c_3) = gh(c_1) = g(c_1) = c_3,
\end{aligned}$$

i. e., $t_{1,c_1} = [g,h] \in \mathrm{Sp}(V)'$. This proves the claim. □

Theorem 7.16. *We have* $\mathrm{Sp}(4,2) \cong S_6$. *In particular,* $\mathrm{Sp}(4,2) = \mathrm{PSp}(4,2)$ *is not perfect.*

Proof. Let $V := \{x \in \mathbb{F}_2^6 : \sum_{i=1}^6 x_i = 0\} \leq \mathbb{F}_2^6$ and $U := \langle (1,\dots,1)\rangle \leq V$. The symmetric group S_6 acts faithfully on $V/U \cong \mathbb{F}_2^4$ by permuting the coordinates. The "standard inner product"

$$[v + U, w + U] := \sum_{i=1}^{6} v_i w_i$$

is a well-defined alternating bilinear form on V/U. Let $v + U \neq 0$; w. l. o. g. we may assume $v_1 = v_2 = 1$ and $v_3 = 0$. Then $w + U := (0,1,1,0,0,0) + U \in V/U$ satisfies $[v + U, w + U] \neq 0$. Hence the bilinear form is nondegenerate. Finally, for $\sigma \in S_6$ we have $[\sigma(v) + U, \sigma(w) + U] = [v + U, w + U]$. This shows $S_6 \leq \mathrm{Sp}(4,2)$. On the other hand, by Theorem 7.10 we have $|\mathrm{Sp}(4,2)| = 2^4(2^2 - 1)(2^4 - 1) = 720 = |S_6|$. □

Theorem 7.17. *For* $n \geq 2$ *and* $(n,q) \neq (2,2)$, $\mathrm{PSp}(2n,q)$ *is simple.*

Proof. By the previous theorems, $G := \mathrm{PSp}(2n,q)$ is perfect and primitive on the set of 1-dimensional subspaces of V. Let $U := \langle u\rangle \leq V$ with $u \neq 0$ and $A := \{\pm t_{\lambda,u} : \lambda \in K\} \leq G_U$. Clearly, A is abelian and normal in G_U. Moreover, every transvection in $\mathrm{Sp}(V)$ is conjugate to $t_{\lambda,u}$ for some $\lambda \in K$. By Lemma 7.13 we have $\langle gAg^{-1} : g \in G\rangle = G$. By Iwasawa's lemma from group theory, G is simple. □

Remark 7.18. Another family of simple groups of Lie type can be constructed via subgroups of $\mathrm{Sp}(4,q)$ with $q = 2^{2n+1}$ and $n \in \mathbb{N}$: Let e_1, e_2, e_3, e_4 be a symplectic basis of $V = \mathbb{F}_q^4$ with $[e_1, e_3] = 1 = [e_2, e_4]$ We define a commutative multiplication on V by

$$e_i * e_j := \begin{cases} e_1 & \text{if } \{i,j\} = \{2,3\},\\ e_2 & \text{if } \{i,j\} = \{1,2\},\\ e_3 & \text{if } \{i,j\} = \{1,4\},\\ e_4 & \text{if } \{i,j\} = \{3,4\},\\ 0 & \text{otherwise} \end{cases}$$

$$\Big(\sum_{i=1}^4 \lambda_i e_i\Big) * \Big(\sum_{i=1}^4 \mu_i e_i\Big) := \sum_{i,j=1}^4 (\lambda_i \mu_j)^{2^n} e_i * e_j$$

for $\lambda_i, \mu_j \in \mathbb{F}_q$. Let $U := \{(u,v) \in V^2 : [u,v] = 0\}$ and

$$\mathrm{Sz}(q) := \{f \in \mathrm{Sp}(V) : \forall (u,v) \in U : f(u) * f(v) = u * v\} \leq \mathrm{Sp}(V).$$

$\mathrm{Sz}(q)$ is called the *Suzuki group* over $\mathbb{F}_q$. It is simple of order

$$|\mathrm{Sz}(q)| = q^2(q^2+1)(q-1) \equiv -1 \pmod 3.$$

The Suzuki groups are the only nonabelian simple groups whose order is not divisible by 3. Since $q^2+1 \equiv 16^n \cdot 4 + 1 \equiv 0 \pmod 5$, the order of every nonabelian simple group is divisible by 3 or 5 (even by 12 or $2^6 \cdot 5$ by Feit-Thompson). The subgroup structure and the character table of $\mathrm{Sz}(q)$ are very clear, similarly to $\mathrm{PSL}(2,q)$ (the Sylow subgroups for odd primes are cyclic, cf. GT, Theorem 10.13). Suzuki also discovered a sporadic group of order 448,345,497,600.

8 Unitary Groups

Remark 8.1. After the linear and symplectic groups, in this chapter we study a third family of (simple) groups of Lie type.

Definition 8.2. Let $q \neq 1$ be a prime power, $K = \mathbb{F}_{q^2}$ and V a finite-dimensional K-vector space. Let $\mathrm{Gal}(K|\mathbb{F}_q) = \langle \alpha \rangle \cong C_2$ with $\overline{x} := \alpha(x) = x^q$ for $x \in K$ (cf. complex conjugation). Let $V \times V \to K$, $(v,w) \mapsto [v,w]$ be a nondegenerate sesquilinear form, i. e., for $u,v,w \in V$ and $\lambda \in K$ we have

$$\begin{aligned} \forall u \neq 0\ \exists v : [u,v] &\neq 0, \\ [u+\lambda v, w] &= [u,w] + \lambda[v,w], \\ [u,v] &= \overline{[v,u]}. \end{aligned}$$

With respect to $[.,.]$, V is called a *unitary space*.

Remark 8.3.

(i) In the following, V always denotes a unitary space.

(ii) For $u,v,w \in V$ and $\lambda \in K$ we have

$$[u, v+\lambda w] = \overline{[v+\lambda w, u]} = \overline{[v,u]} + \overline{\lambda}\overline{[w,u]} = [u,v] + \overline{\lambda}[u,w].$$

(iii) For $K^\times = \langle \zeta \rangle$ we have $\zeta\overline{\zeta} = \zeta^{q+1} \in \mathbb{F}_q$. Hence the *norm* $N\colon K \to \mathbb{F}_q$, $x \mapsto x\overline{x}$ is surjective. Let $x \in K$ be in the kernel of the *trace* $S\colon K \to \mathbb{F}_q$, $y \mapsto y + \overline{y}$. Then $x = -x^q$, and it follows that $x = 0$ or $x^{q-1} = -1$. This shows $|\mathrm{Ker}(S)| \leq q$. As a linear map, S must also be surjective by the homomorphism theorem. For $v \in V$ we have $[v,v] = \overline{[v,v]} \in \mathbb{F}_q$. In the case $[v,v] \neq 0$ there exists a $\lambda \in K$ with $[\lambda v, \lambda v] = 1$, i. e., v can be *normalized*. In contrast to $K = \mathbb{C}$, there are vectors $v \neq 0$ with $[v,v] = 0$ (see Lemma 8.10).

(iv) As usual, one defines $v^\perp$ and $S^\perp \leq V$ for $v \in V$ and $S \subseteq V$. As in Theorem 7.4 one shows $\dim U + \dim U^\perp = \dim V$ for all subspaces $U \leq V$. From this the usual rules follow:

$$(U^\perp)^\perp = U, \qquad U \subseteq W \iff W^\perp \subseteq U^\perp, \qquad (U+W)^\perp = U^\perp \cap W^\perp.$$

If $V = U \oplus U^\perp$, then the restriction of $[.,.]$ to U is nondegenerate. Hence U is a unitary space. We will often use this for induction on $\dim V$.

Theorem 8.4. *Every unitary space V has an orthonormal basis $b_1, \ldots, b_n$, i. e., $[b_i, b_j] = \delta_{ij}$.*

Proof. Let $v \in V \setminus \{0\}$. Since $[.,.]$ is nondegenerate, there exists a $w \in V$ with $[v, w] \neq 0$. After scaling, we may assume $[v, w] + [w, v] \neq 0$ (for $\operatorname{char} K \neq 2$ one can choose $[v, w] = 1$). In the case $[v, v] = 0 = [w, w]$ we have $[v + w, v + w] \neq 0$. In any case, one finds a $b_1 \in V$ with $[b_1, b_1] \neq 0$. After normalization, $[b_1, b_1] = 1$. In the case $n = 1$ we are done. So let $n \geq 2$ and $U := b_1^{\perp}$. Clearly $V = \langle b_1 \rangle \oplus U$. By Remark 8.3, U is a unitary space of dimension $n - 1$. By induction on n we may assume that U has an orthonormal basis $b_2, \ldots, b_n$. Clearly $b_1, \ldots, b_n$ is now an orthonormal basis of V. □

Example 8.5. As over $\mathbb{C}$, one sees that $V = K^n$ with respect to the *standard inner product*

$$[v, w] := v_1\overline{w_1} + \ldots + v_n\overline{w_n} \qquad (v, w \in V)$$

is a unitary space. The standard basis $e_1, \ldots, e_n$ is an orthonormal basis of V. By Theorem 8.4, in the following we can often restrict ourselves to this special space.

Definition 8.6. A linear map between unitary spaces $f : V \to W$ is called an *isometry* if $[f(u), f(v)]_W = [u, v]_V$ for all $u, v \in V$. One defines the groups

$$\begin{aligned} \mathrm{GU}(V) &:= \left\{ f \in \mathrm{GL}(V) : \forall v, w \in V : [f(v), f(w)] = [v, w] \right\} && (\textit{general unitary group}), \\ \mathrm{SU}(V) &:= \mathrm{GU}(V) \cap \mathrm{SL}(V) && (\textit{special unitary group}), \\ \mathrm{PSU}(V) &:= \mathrm{SU}(V)/\mathrm{Z}(\mathrm{SU}(V)) && (\textit{projective special unitary group}). \end{aligned}$$

Remark 8.7. Since $[.,.]$ is nondegenerate, every isometry is injective. Let $f \in \mathrm{GL}(V)$ with matrix $A \in \mathrm{GL}(n, K)$ with respect to the standard basis. For $v, w \in V$ we have $[f(v), f(w)] = [Av, Aw] = v^{\mathrm{t}} A^{\mathrm{t}} \overline{A} \overline{w}$. By plugging in the standard basis, one sees

$$\forall v, w \in V : [f(v), f(w)] = [v, w] \iff A^{\mathrm{t}}\overline{A} = 1_n.$$

In this way one obtains the matrix groups

$$\begin{aligned} \mathrm{GU}(n, q) &:= \{A \in \mathrm{GL}(n, q^2) : A^{\mathrm{t}}\overline{A} = 1_n\},^{16} \\ \mathrm{SU}(n, q) &:= \mathrm{GU}(n, q) \cap \mathrm{SL}(n, q^2), \\ \mathrm{PSU}(n, q) &:= \mathrm{SU}(n, q)/\mathrm{Z}(\mathrm{SU}(n, q)). \end{aligned}$$

Example 8.8. For every permutation matrix $P \in \mathrm{GL}(n, q^2)$ we have $P^{\mathrm{t}}\overline{P} = P^{-1}P = 1_n$, i. e., $P \in \mathrm{GU}(n, q)$. Hence $\mathrm{GU}(n, q)$ (respectively $\mathrm{SU}(n, q)$) has a subgroup isomorphic to S_n (respectively A_n).

Lemma 8.9. *We have $\mathrm{C}_{\mathrm{GL}(V)}(\mathrm{SU}(V)) = K^{\times} \mathrm{id}_V$.*

[16] Caution: By default, GAP uses a different basis. One has to specify the Hermitian form explicitly: `GU(n, q, IdentityMat(n) * Z(q)^0)`.

Proof. W. l. o. g. we may assume $V = K^n$ with $n \geq 2$. Let $K^\times = \langle\zeta\rangle$. Since $\zeta^{q-1}\overline{\zeta^{q-1}} = \zeta^{q-1+q^2-q} = \zeta^{q^2-1} = 1$, we have $D := \operatorname{diag}(\zeta^{q-1}, \zeta^{1-q}, 1, \ldots, 1) \in \mathrm{SU}(n,q) \setminus \{1_n\}$ (also for $q = 2$). It follows that

$$\mathrm{C}_{\mathrm{GL}(V)}(\mathrm{SU}(V)) \leq \mathrm{C}_{\mathrm{GL}(V)}(D) = \left\{ \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & A \end{pmatrix} : a, b \in K^\times, A \in \mathrm{GL}(n-2, K) \right\}.$$

By permuting the coordinates, we obtain that $\mathrm{C}_{\mathrm{GL}(V)}(\mathrm{SU}(V))$ consists of diagonal matrices. Because of

$$\begin{pmatrix} 0 & \cdots & 0 & (-1)^{n-1} \\ 1 & \ddots & & 0 \\ & \ddots & \ddots & \vdots \\ & & 1 & 0 \end{pmatrix} \in \mathrm{SU}(n,q)$$

only scalar matrices can centralize $\mathrm{SU}(V)$. Conversely, of course $K^\times \operatorname{id}_V \leq \mathrm{Z}(\mathrm{GL}(V))$. □

Lemma 8.10. *For $\lambda \in \mathbb{F}_q$ let $V_\lambda := \{v \in V : [v,v] = \lambda\}$. Then*

$$z_n := |V_0| = q^{2n-1} + (-1)^n q^{n-1}(q-1),$$
$$w_n := |V_1| = q^{n-1}(q^n - (-1)^n).$$

Proof. W. l. o. g. we may assume $V = K^n$ with respect to the standard inner product.

(i) For $n = 1$ we have $V_0 = \{0\}$ and $z_1 = 1 = q-(q-1)$ as claimed. Let $n \geq 1$ and $v = (v', v_n) \in K^{n+1}$ with $v' \in K^n$ and $[v,v] = 0$. In the case $[v',v'] = 0$ we also have $v_n = 0$. There are z_n possibilities for this. Now let $[v',v'] \neq 0$ and $K^\times = \langle\zeta\rangle$. Let $a \in \mathbb{Z}$ with $-[v',v'] = \zeta^{a(q+1)}$. From

$$v_n^{q+1} = v_n\overline{v_n} = [v,v] - [v',v'] = -[v',v']$$

it follows that $v_n \in \zeta^a\langle\zeta^{q-1}\rangle$. There exist exactly $|\langle\zeta^{q-1}\rangle| = q+1$ elements with this property. This shows

$$z_{n+1} = z_n + (q^{2n} - z_n)(q+1) = q^{2n+1} + q^{2n} - qz_n.$$

The claim follows by induction.

(ii) For $\lambda \neq 0$, the map $V_\lambda \to V_{\lambda\zeta^{q+1}}$, $v \mapsto \zeta v$ is a bijection. With (i) it follows that

$$w_n = |V_1| = \frac{q^{2n} - z_n}{q-1} = q^{2n-1} - (-1)^n q^{n-1} = q^{n-1}(q^n - (-1)^n).$$ □

Theorem 8.11. *We have*

$$\boxed{\begin{aligned} |\mathrm{GU}(n,q)| &= q^{n(n-1)/2} \prod_{k=1}^{n} (q^k - (-1)^k), \\ |\mathrm{SU}(n,q)| &= q^{n(n-1)/2} \prod_{k=2}^{n} (q^k - (-1)^k), \\ |\mathrm{PSU}(n,q)| &= \frac{q^{n(n-1)/2}}{\gcd(n, q+1)} \prod_{k=2}^{n} (q^k - (-1)^k). \end{aligned}}$$

Proof. We have $A \in \mathrm{GU}(n,q)$ if and only if the rows of A form an orthonormal basis of V with respect to $[.,.]$. For the first row a_1 of A there are exactly w_n possibilities by Lemma 8.10. The second row a_2 lies in $U := a_1^\perp$. Since $V = \langle a_1 \rangle \oplus U$, U is a unitary space of dimension $n-1$. Hence there are w_{n-1} possibilities for a_2 and so on. This shows

$$|\mathrm{GU}(n,q)| = w_1 \dots w_n = \prod_{k=1}^{n} q^{k-1}(q^k - (-1)^k) = q^{n(n-1)/2} \prod_{k=1}^{n} (q^k - (-1)^k).$$

For $A \in \mathrm{GU}(n,q)$ we have $\det(A)^{q+1} = \det(A)\overline{\det(A)} = \det(A^{\mathrm{t}}\overline{A}) = 1$. Since $\mathrm{diag}(\zeta^{q-1}, 1, \dots, 1) \in \mathrm{GU}(n,q)$, the map $\det\colon \mathrm{GU}(n,q) \to \langle \zeta^{q-1} \rangle$ is surjective with kernel $\mathrm{SU}(n,q)$. This implies the formula for $|\mathrm{SU}(n,q)|$. By Lemma 8.9 we have

$$\mathrm{Z}(\mathrm{SU}(n,q)) = \mathrm{SU}(n,q) \cap K^\times 1_n = \mathrm{SU}(n,q) \cap \langle \zeta^{q-1} \rangle.$$

For $\zeta^{a(q-1)} 1_n \in \mathrm{Z}(\mathrm{SU}(n,q))$ with $0 \le a \le q$ we have $\zeta^{a(q-1)n} = 1$. Hence $a \equiv 0 \pmod{(q+1)/\gcd(n, q+1)}$ and $|\mathrm{Z}(\mathrm{SU}(n,q))| = \gcd(n, q+1)$. This implies the formula for $|\mathrm{PSU}(n,q)|$. □

Lemma 8.12. *We have* $\mathrm{SU}(2,q) \cong \mathrm{SL}(2,q)$ *and* $\mathrm{PSU}(2,q) \cong \mathrm{PSL}(2,q)$.

Proof. We first construct a basis $\{u, v\} \subseteq V_0$ of $V = K^2$ with $[u,v] = -[v,u]$. To this end, let $\mu \in K \setminus \{1\}$ with $\mu^{q+1} = \mu\overline{\mu} = -1$. For $2 \nmid q$ we set $u := (1, \mu)$, $v := (\mu, 1)$, and for $2 \mid q$ let $u := (1,1)$, $v := (1, \mu)$. In both cases u and v are linearly independent with $[u,u] = [v,v] = 0$. By Remark 8.3 it follows that $[u,v] \neq 0$. After normalizing v we obtain $[u,v] = 1$. For $2 \mid q$ we then have $[u,v] = \overline{[v,u]} = [v,u] = -[v,u]$. For $2 \nmid q$ there exists $\tau \in K$ with $\tau^{q-1} = -1$, i. e., $\overline{\tau} = -\tau$. By replacing u with τu, we obtain $[u,v] = -[v,u]$ as desired.

Let $A \in \mathrm{SU}(2,q)$ and $Au = au + cv$, $Av = bu + dv$ with $a, b, c, d \in K$. Then

$$\begin{aligned} a\overline{c} - \overline{a}c &= \frac{[Au, Au]}{[u,v]} = 0 = \frac{[Av, Av]}{[u,v]} = b\overline{d} - \overline{b}d, \\ a\overline{d} - c\overline{b} &= \frac{[Au, Av]}{[u,v]} = 1 = d\overline{a} - b\overline{c}, \\ ad - bc &= \det A = 1. \end{aligned}$$

It follows that

$$\begin{aligned} \overline{c} + bc\overline{c} &= (1+bc)\overline{c} = ad\overline{c} = d\overline{a}c = (1 + b\overline{c})c = c + bc\overline{c}, \\ b + b\overline{b}c &= b(1 + c\overline{b}) = ba\overline{d} = ad\overline{b} = (1+bc)\overline{b} = \overline{b} + b\overline{b}c \end{aligned}$$

and $b, c \in \mathbb{F}_q$. In the case $b \neq 0$ we get $d = \overline{d} \in \mathbb{F}_q$. Otherwise $a \neq 0$ and $ad = 1 = a\overline{d}$. Again $d \in \mathbb{F}_q$. Analogously we obtain $a \in \mathbb{F}_q$. Hence the change of basis $\{(1,0), (0,1)\} \mapsto \{u, v\}$ yields a monomorphism $\varphi\colon \mathrm{SU}(2,q) \to \mathrm{SL}(2,q)$. By Theorem 8.11 we have $|\mathrm{SU}(2,q)| = q(q^2-1) = |\mathrm{SL}(2,q)|$. Hence φ is an isomorphism. □

Example 8.13. Clearly $\mathrm{GU}(1,q) \cong C_{q+1}$ and $\mathrm{SU}(1,q) = 1$. By Lemma 8.12 we have $\mathrm{PSU}(2,2) \cong \mathrm{SU}(2,2) \cong \mathrm{SL}(2,2) \cong S_3$ and $\mathrm{PSU}(2,3) \cong \mathrm{PSL}(2,3) \cong A_4$. By Theorem 8.11 we have

$$\begin{aligned} |\mathrm{GU}(3,2)| &= 2^3(2+1)(2^2-1)(2^3+1) = 8 \cdot 81 = 648, \\ |\mathrm{SU}(3,2)| &= 216, \\ |\mathrm{PSU}(3,2)| &= 72. \end{aligned}$$

In particular, these groups are solvable. More precisely, $\mathrm{PSU}(3,2) \cong M_9 \cong C_3^2 \rtimes Q_8$ (Exercise 28). The following GAP code shows $\mathrm{PSU}(4,2) \cong \mathrm{SU}(4,2) \cong \mathrm{PSp}(4,3)$:

```
G:=SU(4,2);
H:=PSp(4,3);
IsomorphismGroups(G,H);
```

Remark 8.14. As in the proof of Lemma 8.12, one shows that every unitary space V of dimension ≥ 2 has a hyperbolic pair (u, v), i. e., $u, v \in V_0$ and $[u, v] = 1$ (Definition 7.6).

Theorem 8.15 (Witt)**.** *Let V be a unitary space, $U \leq V$ and $\sigma\colon U \to V$ an isometry. Then there exists a $\tau \in \mathrm{GU}(V)$ with $\tau_{|U} = \sigma$.*

Proof. Induction on $\dim U$. W. l. o. g. we may assume $U \neq 0$. Let $W \leq U$ with $\dim(U/W) = 1$. By induction there exists a $\tau \in \mathrm{GU}(V)$ with $\tau_{|W} = \sigma_{|W}$. By replacing σ with $\tau_{|U}^{-1}\sigma$, we may assume $\sigma_{|W} = \mathrm{id}_W$. W. l. o. g. we may assume $\sigma \neq \mathrm{id}_U$. Then $R := (\sigma - \mathrm{id}_U)(U) \leq V$ is 1-dimensional. For $u, v \in U$ we have

$$[\sigma(u), \sigma(v) - v] = [\sigma(u), \sigma(v)] - [\sigma(u), v] = [u, v] - [\sigma(u), v] = [u - \sigma(u), v]. \tag{8.1}$$

This shows $W \subseteq R^\perp$.

Case 1: $U \not\subseteq R^\perp$.
Let $R^\perp = W \oplus Y$. For dimension reasons, $V = U + R^\perp = U + W + Y = U \oplus Y$. For $u \in U$ and $y \in Y$ we have $[\sigma(u) - u, y] = 0$, i. e., $[\sigma(u), y] = [u, y]$. It follows easily that the map $V \to V$, $u + y \mapsto \sigma(u) + y$ is a unitary extension of σ.

Case 2: $U \subseteq R^\perp$.
From (8.1) it follows that $\sigma(U) \subseteq R^\perp$. Suppose that $\sigma(U) \neq U$. Then $W = \sigma(W) = U \cap \sigma(U)$. Choose $u \in U \setminus W$ and $v \in \sigma(U) \setminus W$. For $U_0 := \langle u + v \rangle$ we have

$$U + \sigma(U) = U \oplus U_0 = \sigma(U) \oplus U_0.$$

Let $R^\perp = (U + \sigma(U)) \oplus Y$ and $S := U_0 + Y$. Then $R^\perp = U \oplus S = \sigma(U) \oplus S$. In the case $\sigma(U) = U$ one can choose an arbitrary complement S of U in $R^\perp$. In both cases, (8.1) shows that

$$R^\perp \to R^\perp, \qquad u + s \mapsto \sigma(u) + s$$

is an isometry for $u \in U$ and $s \in S$. Hence we may assume $U = R^\perp = \sigma(U)$. Let $R = \langle x \rangle \subseteq U$. Then $[x, x] = 0$. Let $v \in V \setminus R^\perp$. Then $T := \langle x, v \rangle$ is a unitary space. We can replace v by a suitable element in T such that (x, v) is a hyperbolic pair (Remark 8.14). Since $T^\perp \subseteq x^\perp = R^\perp = U$, we have

$$U = \langle x \rangle \oplus T^\perp = \langle \sigma(x) \rangle \oplus \sigma(T^\perp).$$

It follows that $\dim(\langle v \rangle + \sigma(T^\perp)) = \dim(V) - 1$. Choose $w \in V$ with $\langle v \rangle + \sigma(T^\perp) = w^\perp$. Since $\langle v \rangle + \langle \sigma(x) \rangle + \sigma(T^\perp) = \langle v \rangle + U = V$, we have $\sigma(x) \notin w^\perp$. This gives $w \notin U$, because $U = \sigma(U) = \sigma(x^\perp) = \sigma(x)^\perp$. Since $\langle \sigma(x), w \rangle$ is a unitary space, we can modify w such that $(\sigma(x), w)$ is a hyperbolic pair. Afterwards we still have $\sigma(T^\perp) \subseteq \langle \sigma(x), w \rangle^\perp$.

We define the linear map $\tau\colon V \to V$, $u + \lambda v \mapsto \sigma(u) + \lambda w$ for $u \in U$ and $\lambda \in K$. Because of

$$[x, v] = 1 = [\sigma(x), w], \qquad [y, v] = 0 = [\sigma(y), w], \qquad [v, v] = 0 = [w, w]$$

for all $y \in T^\perp$, $\tau \in \mathrm{GU}(V)$ is an extension of σ. □

Lemma 8.16. *Let V be a unitary space and $v, w \in V$ linearly independent with $[v,v] = [w,w] = 1$. Then $V = \langle v,w\rangle \oplus \langle v,w\rangle^\perp$ if and only if $[v,w][w,v] \neq 1$.*

Proof. Let $\lambda := [v,w]$. Let $a, b \in K$ with $u := av + bw \in \langle v,w\rangle^\perp$. Then $0 = [u,v] = a + b\overline{\lambda}$ and $0 = [u,w] = a\lambda + b$. It follows that $a = -b\overline{\lambda} = a\lambda\overline{\lambda}$. This shows $u = 0$ or $\lambda\overline{\lambda} = 1$. Conversely, if $\lambda\overline{\lambda} = 1$, then $v - \lambda w \in \langle v,w\rangle^\perp$. □

Definition 8.17. Let $v \in V_0$ and $\lambda \in K$ with $\lambda + \overline{\lambda} = 0$. The linear map

$$t_{\lambda,v} \colon V \to V, \qquad x \mapsto x + \lambda[x,v]v$$

is called a (*unitary*) *transvection*.

Remark 8.18. For $v, x \in V$ and $\lambda, \mu \in K$ we have

$$t_{\lambda,v}(t_{\mu,v}(x)) = t_{\lambda,v}(x + \mu[x,v]v) = x + \mu[x,v]v + \lambda[x + \mu[x,v]v, v]v = t_{\lambda+\mu,v}(x).$$

In particular, $t_{\lambda,v}$ is invertible with $t_{\lambda,v}^{-1} = t_{-\lambda,v}$. Moreover, for $p := \operatorname{char} K$ we have $t_{\lambda,v}^p = t_{p\lambda,v} = t_{0,v} = \mathrm{id}_V$. From $\lambda + \overline{\lambda} = 0$ it follows that

$$[t_{\lambda,v}(x), t_{\lambda,v}(y)] = [x,y] + \lambda[x,v][v,y] + \overline{\lambda}\overline{[y,v]}[x,v] = [x,y]$$

and $t_{\lambda,v} \in \mathrm{GU}(V)$. As in the proof of Theorem 8.11, $\det(t_{\lambda,v})^{q+1} = 1$. On the other hand, $\det(t_{\lambda,v})^p = \det(t_{\lambda,v}^p) = \det(\mathrm{id}_V) = 1$. This shows $t_{\lambda,v} \in \mathrm{SU}(V)$. For $g \in \mathrm{GU}(V)$ we have $g t_{\lambda,v} g^{-1} = t_{\lambda,g(v)}$ as in Remark 7.12. For $\mu \in K^\times$ we have

$$t_{\lambda,\mu v}(x) = x + \lambda[x,\mu v]\mu v = x + \lambda\mu\overline{\mu}[x,v]v = t_{\lambda\mu\overline{\mu},v}(x)$$

with $\lambda\mu\overline{\mu} + \overline{\lambda}\overline{\mu}\mu = 0$.

Lemma 8.19. *Let $n \geq 2$ and $(n,q) \neq (3,2)$. Then $\mathrm{SU}(n,q)$ is generated by all transvections.*

Proof. Let $V := K^n$.
Case 1: $n = 2$.
By the proof of Lemma 8.12 there exists a basis $\{v,w\}$ of V such that $\mathrm{SU}(V)$ consists of the matrices in $\mathrm{SL}(2,q)$. By GT, Lemma 10.8, $\mathrm{SL}(2,q)$ is generated by the elementary matrices of the form $1_2 + \lambda E_{ij}$ with $\lambda \in \mathbb{F}_q$ and $i \neq j$. It suffices to show that these matrices correspond to unitary transvections. W. l. o. g. let $(i,j) = (1,2)$. Let $\mu := [w,v] = -[v,w] \in K$ (see the proof of Lemma 8.12). We have $\mu^{-1} = -\overline{\mu}^{-1} = -\mu^{-q} = -\overline{\mu^{-1}}$ and $\lambda\mu^{-1} + \overline{\lambda\mu^{-1}} = \lambda(\mu^{-1} + \overline{\mu^{-1}}) = 0$. Hence $t := t_{\lambda\mu^{-1},v}$ is a transvection with

$$t(v) = v, \qquad t(w) = w + \lambda\mu^{-1}[w,v]v = w + \lambda v.$$

Thus t corresponds to the matrix $1_2 + \lambda E_{12}$.

For the induction on n we additionally need that $\mathrm{SU}(V)$ acts transitively on V_1. For $u, v \in V_1$ there exists an $\alpha \in \mathrm{GU}(V)$ with $\alpha(u) = v$ by Witt. Let $d := \det(\alpha)$ and $u^\perp = \langle w\rangle$. We define $\beta \in \mathrm{GU}(V)$ by $\beta(u) = u$ and $\beta(w) = d^{-1}w$. Then $\alpha\beta \in \mathrm{SU}(V)$ with $\alpha\beta(u) = v$.

Case 2: $n \geq 3$.
By Example 8.13 and Theorem 7.17, $\mathrm{SU}(4,2) \cong \mathrm{PSp}(4,3)$ is simple. By Remark 8.18, the transvections generate a (nontrivial) normal subgroup. Hence we may assume $(n,q) \neq (4,2)$ (this allows the induction

step to $n-1$). Let $e_1, \dots, e_n$ be the standard basis of V. Let $\alpha \in \mathrm{SU}(V)$. Then $v := \alpha(e_1)$ is a unit vector.

Step 1: There exists a product τ of transvections with $\tau(e_1) = v$.
If e_1 and v are linearly dependent, then by the first part of the proof there exists a product τ' of transvections in $\langle e_1, e_2\rangle$ with $\tau'(e_1) = v$. By Lemma 8.16, τ' can be extended to a product of transvections on V. Hence we may assume that e_1 and v are linearly independent.

Now suppose that $U := \langle e_1, v\rangle^\perp$ contains a vector u with $[u,u] \neq 0$. After normalization, $[u,u] = 1$. By Lemma 8.16 there exist products of transvections $\tau' \in \mathrm{SU}(\langle e_1, u\rangle)$ and $\tau'' \in \mathrm{SU}(\langle u, v\rangle)$ with $\tau'(e_1) = u$ and $\tau''(u) = v$. Again we can extend τ' and τ'' to V. Then $\tau = \tau''\tau'$ has the desired property.

Hence we may assume $[u,u] = 0$ for all $u \in U$. Since $e_1 \in U^\perp \setminus U$, we have $U \subsetneq U^\perp$. From $n - 2 = \dim U < \dim U^\perp = 2$ we obtain $n = 3$. By assumption, $q \geq 3$. If $v_i\overline{v_i} = [v, e_i][e_i, v] \neq 1$ for some $i \in \{1,2,3\}$, then as before we can map e_1 to e_i and e_i to v by means of transvections. So we may assume $v_i\overline{v_i} = 1$ for $i = 1,2,3$. Since $1 = [v,v] = 1 + 1 + 1$, q is even. There exist $\lambda, \mu \in \mathbb{F}_q^\times$ with $\lambda^2 + \mu^2 = (\lambda + \mu)^2 = 1$. For $u := (0, \lambda, \mu) \in V_1$ we have

$$[u,v][v,u] = (\lambda\overline{v_2} + \mu\overline{v_3})(\lambda v_2 + \mu v_3) = \lambda^2 + \mu^2 + \lambda\mu(v_2\overline{v_3} + \overline{v_2}v_3) = 1 + \lambda\mu(v_2\overline{v_3} + \overline{v_2}v_3).$$

Suppose that $v_2\overline{v_3} + \overline{v_2}v_3 = 0$. Since $2 \mid q$, we have $v_2^2 = v_2^2 v_3\overline{v_3} = v_2\overline{v_2}v_3^2 = v_3^2$ and $v_2 = v_3$. Since we can map v to $w := \alpha(e_2)$ by means of transvections, we may transfer the conditions on v to w as well. Hence we also have $w_2 = w_3$. But now $[v,w] = v_1\overline{w_1} \neq 0$. This contradiction shows $[u,v][v,u] \neq 1$. Hence we can map e_1 to u and u to v by means of transvections.

Step 2: α is a product of transvections.
For $\beta := \tau^{-1}\alpha \in \mathrm{SU}(V)$ we have $\beta(e_1) = e_1$. Hence β acts on $U := \langle e_2, \dots, e_n\rangle = e_1^\perp$. Since $(n,q) \neq (4,2)$, we have $(n-1, q) \neq (3,2)$, so β_U is a product of transvections by induction on n, which can be extended to V. Hence α is also a product of transvections. □

Example 8.20. Every transvection in $\mathrm{SU}(3,2)$ has order $q = 2$ by Remark 8.18. By Exercise 28, the transvections in $\mathrm{SU}(3,2)$ generate a proper subgroup of order 54.

Lemma 8.21. *For $n \geq 2$, $\mathrm{PSU}(n,q)$ acts faithfully and primitively on $\Omega := \big\{\langle v\rangle : v \in V_0 \setminus \{0\}\big\}$.*

Proof. Let $V := K^n$. Clearly, $\mathrm{SU}(V)$ acts on Ω via ${}^A\langle v\rangle = \langle Av\rangle$. By Lemma 8.9, $\mathrm{Z}(\mathrm{SU}(V))$ lies in the kernel of the action. Conversely, let A be in the kernel. As in the proof of Lemma 8.12 we construct $u, v \in V_0$ with $[u,v] = -[v,u]$. Let $Au = \lambda_1 u$ and $Av = \lambda_2 v$. Since $u + v \in V_0$, we have

$$\lambda_1 u + \lambda_2 v = A(u+v) \in \langle u + v\rangle.$$

This shows $\lambda := \lambda_1 = \lambda_2$. Let $U := \langle u, v\rangle^\perp$. Since $V = U \oplus U^\perp$, U is a unitary space. For $n = 2$ we have $U = 0$ and $A = \lambda 1_2$. Now let $n \geq 3$. By induction on n there exists $\mu \in K$ with $A_{|U} = \mu\,\mathrm{id}_U$. Choose $z \in U$ with $\gamma := [z,z] \neq 0$. By Remark 8.3 there exists a $\delta \in K$ with $\delta + \overline{\delta} = \gamma$. For $\beta := \delta/[u,v]$ we have $(\overline{\beta} - \beta)[u,v] = -(\delta + \overline{\delta}) = -\gamma$ because $\overline{[u,v]} = [v,u] = -[u,v]$. For $w := u + \beta v + z$ we thus have

$$[w,w] = (\overline{\beta} - \beta)[u,v] + [z,z] = 0,$$

so $w \in V_0$. From $Aw = \lambda u + \lambda\beta v + \mu z \in \langle w\rangle$ it follows that $\lambda = \mu$ and $A = \lambda 1_n \in \mathrm{Z}(\mathrm{SU}(V))$. Hence $\mathrm{PSU}(V)$ acts faithfully on Ω.

Let $v \in V_0 \setminus \{0\}$ with w. l. o. g. $v_1 v_2 \neq 0$. As is well known, there exist $q+1$ elements $\lambda \in K$ with $\lambda^{q+1} = \lambda\overline{\lambda} = -1$. In particular, there exists such a λ with $v_1 + \lambda v_2 \neq 0$. Then $u := (1, \overline{\lambda}, 0, \dots, 0) \in V_0$

and $[v,u] \neq 0$. By scaling v we achieve $[u,v] = -[v,u]$. Now $U := \langle u,v\rangle$ is a unitary space with $\mathrm{SU}(U) \cong \mathrm{SL}(2,q)$ (Lemma 8.12). There exists an $A \in \mathrm{SU}(U)$ with $\langle Av\rangle = \langle u\rangle$. Since $V = U \oplus U^\perp$, we can extend A to $\mathrm{SU}(V)$ by acting trivially on $U^\perp$. Let $\mu \neq \lambda$ with $\mu\overline{\mu} = -1$. Then $w := (1,\overline{\mu}) \in V_0$ with $[u,w] = 1 + \overline{\lambda}\mu \neq 0$. Hence there exists a $B \in \mathrm{SU}(\langle u,w\rangle)$ with $\langle Bu\rangle = \langle w\rangle$. Again B can be extended to V. This shows that $\mathrm{PSU}(V)$ acts transitively on Ω.

To prove primitivity, let first $n = 2$ and $V = \langle u,v\rangle$ with $u,v \in V_0$ and $[u,v] = -[v,u]$. Every element in $\Omega \setminus \langle v\rangle$ is generated by a vector $w := u + \lambda v \in V_0$ with $\lambda \in K$. Then

$$0 = [w,w] = \lambda[v,u] + \overline{\lambda}[u,v] = (\lambda - \overline{\lambda})[v,u],$$

i. e., $\lambda = \overline{\lambda} \in \mathbb{F}_q$. Hence Ω consists precisely of the 1-dimensional subspaces of $\mathbb{F}_q u + \mathbb{F}_q v$. By means of the isomorphism $\mathrm{SU}(V) \cong \mathrm{SL}(2,q)$, $\mathrm{SU}(V)$ even acts 2-transitively on Ω (GT, Lemma 10.7). Now let $n \geq 3$. Suppose that there exists a block $\Delta \subseteq \Omega$. Let $\langle u\rangle, \langle v\rangle \in \Delta$ be distinct.

Case 1: $[u,v] \neq 0$.
Let also $w \in V_0$ with $[u,w] \neq 0$. After scaling we may assume $[u,v] = [u,w]$. Then there exists an isometry $\alpha\colon \langle u,v\rangle \to \langle u,w\rangle$ with $\alpha(u) = u$ and $\alpha(v) = w$. By Witt, α can be extended to $\alpha \in \mathrm{GU}(V)$. For $d := \det(\alpha)$ we have $d^{q+1} = d\overline{d} = 1$, i. e., $d = \zeta^{(q-1)a}$ for some $a \in \mathbb{Z}$. We define $U := \langle u,v\rangle$ and $\beta \in \mathrm{GL}(U)$ with $\beta(u) = \zeta^{-qa}u$ and $\beta(v) = \zeta^a v$. Because of

$$[\beta(u),\beta(v)] = \zeta^{-qa+qa}[u,v] = [u,v]$$

we have $\beta \in \mathrm{GU}(U)$ with $\det(\beta) = \zeta^{-qa+a} = d^{-1}$. Since $V = U \oplus U^\perp$, we can extend β to $\mathrm{GU}(V)$ as $\beta + \mathrm{id}_{U^\perp}$. Then $\alpha\beta \in \mathrm{SU}(V)$ with $\alpha\beta(u) \in \langle u\rangle$ and $\alpha\beta(v) \in \langle w\rangle$. Hence Δ contains all $\langle v\rangle$ with $[u,v] \neq 0$.

Now let $\langle w\rangle \in \Omega$ with $w \in u^\perp$. Since V is not the union of two proper subspaces, there exists an $x \in V \setminus (u^\perp \cup w^\perp)$. After scaling let $[x,u] = 1$. Since the trace is surjective, there exists a $\lambda \in K$ with $\lambda + \overline{\lambda} = -[x,x]$ $(\in \mathbb{F}_q)$. For $x' := x + \lambda u$ we have

$$[x',x'] = [x,x] + (\lambda + \overline{\lambda})[u,x] = 0,$$

i. e., $x' \in V_0$. Since $u \in w^\perp$, we have $x' \notin u^\perp \cup w^\perp$. From $x' \notin u^\perp$ it follows that $\langle x'\rangle \in \Delta$. From $w \notin (x')^\perp$ we obtain $\langle w\rangle \in \Delta$. This yields the contradiction $\Delta = \Omega$.

Case 2: $[u,v] = 0$.
Let $x \in u^\perp \setminus v^\perp$ with $[x,v] = 1$. Let $\lambda \in K$ with $\lambda + \overline{\lambda} = -[x,x]$. For $x' := x + \lambda v \in u^\perp \setminus v^\perp$ we have $[x',x'] = 0$ as in Case 1. Now $U := \langle x',v\rangle$ is a unitary space. There exists an $\alpha \in \mathrm{SU}(V)$ with $\alpha(U) = U$, $\alpha(v) \in \langle x'\rangle$ and $\alpha_{|U^\perp} = \mathrm{id}_{|U^\perp}$. Since $u \in U^\perp$, we have $\langle u\rangle = \langle \alpha(u)\rangle \in \Delta \cap \alpha(\Delta) = \Delta$. This shows $\langle x'\rangle = \langle \alpha(v)\rangle \in \Delta$. From Case 1 we obtain the contradiction $\Delta = \Omega$. □

Remark 8.22. By Lemma 8.10 we have $z_3 = q^5 - q^2(q-1) = (q^2-1)(q^3+1)+1$. Hence $G := \mathrm{PSU}(3,q)$ is a primitive group of degree $|\Omega| = \frac{z_3-1}{|K^\times|} = q^3 + 1$.

Lemma 8.23. *For $n \geq 3$ and $(n,q) \neq (3,2)$, $\mathrm{SU}(n,q)$ is perfect.*

Proof. By Lemma 8.19 it suffices to show that every unitary transvection $t := t_{\lambda,u}$ is a product of commutators. Let $v \in V \setminus u^\perp$ such that (u,v) is a hyperbolic pair. Then $U := \langle u,v\rangle$ is a unitary space and $t_{|U^\perp} = \mathrm{id}_{U^\perp}$ since $U^\perp \subseteq u^\perp$. By Lemma 8.12, $t_{|U} \in \mathrm{SU}(U) \cong \mathrm{SL}(2,q)$.

First let $q \geq 4$. By GT, Lemma 10.9, $\mathrm{SL}(2,q)$ is perfect. Hence $t_{|U}$ is a product of commutators $[x,y] \in \mathrm{SU}(U)$. As usual, we can extend x and y to V by acting trivially on $U^\perp$. Hence $[x,y]$ is the restriction of a commutator $c \in \mathrm{SU}(V)$ with $c_{|U^\perp} = \mathrm{id}_{U^\perp}$. Thus t is a product of commutators.

Now let $q = 3$. Together with U, also $U^\perp$ is a unitary space. Hence there exists a $w \in U^\perp$ with $[w, w] = 1$. Let $W := \langle u, v, w\rangle$. As in the first part of the proof, it suffices to show that $t_{|W} \in \mathrm{SU}(W)$ is a product of commutators. Let $\mu \in K$ with $\mu\overline{\mu} = -1$. With respect to the basis $\{u, v, w\}$ we have $d := \mathrm{diag}(\mu, -\mu, -\mu^{-2}) \in \mathrm{SU}(W)$. We have

$$t_{|W} = \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \lambda/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} d \begin{pmatrix} 1 & \lambda/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} d^{-1} = [t_{\lambda/2,\, u}, d] \in \mathrm{SU}(W)'.$$

Finally, let $q = 2$ and $n \geq 4$. We choose a hyperbolic pair (u', v') with $u', v' \in U^\perp$. W. l. o. g. let $V = U \oplus \langle u', v'\rangle$.[17] From $\lambda + \overline{\lambda} = 0$ it follows that $\lambda = \overline{\lambda} \in \mathbb{F}_2$. In the case $\lambda = 0$, $t = \mathrm{id}_V$ is a commutator. So let $\lambda = 1$. Let $K^\times = \langle\zeta\rangle$ and $x, y \in \mathrm{GL}(V)$ with

$$\begin{aligned} x(u) &:= u, & x(u') &:= u', & x(v) &:= v + \zeta^2 u', & x(v') &:= \zeta u + v', \\ y(u) &:= u, & y(u') &:= \zeta^2 u + u', & y(v) &:= v + \zeta v', & y(v') &:= v' \end{aligned}$$

Since $\zeta + \zeta^2 = 1$, we have $x, y \in \mathrm{SU}(V)$ and $x^2 = 1$, $y^2 = 1$. For $z := [x, y]$ we have

$$\begin{aligned} z(u) &= u, \\ z(u') &= xy(\zeta^2 u + u') = \zeta^2 u + \zeta^2 u + u' = u', \\ z(v) &= xyx(v + \zeta v') = xy(v + \zeta^2 u' + \zeta^2 u + \zeta v') = x(v + \zeta v' + \zeta u + \zeta^2 u' + \zeta^2 u + \zeta v') \\ &= x(u + v + \zeta^2 u') = u + v, \\ z(v') &= xy(\zeta u + v') = \zeta u + \zeta u + v' = v'. \end{aligned}$$

This shows $t = z \in \mathrm{SU}(V)'$. □

Theorem 8.24. *For $n \geq 3$ and $(n, q) \neq (3, 2)$, $\mathrm{PSU}(n, q)$ is simple.*

Proof. Let $V = K^n$. We identify the elements in $\mathrm{SU}(V)$ with their cosets in $G := \mathrm{PSU}(V)$. By Lemma 8.21, G is a primitive permutation group on $\Omega = \{\langle v\rangle : v \in V_0 \setminus \{0\}\}$. By Lemma 8.23, G is perfect. Let $v \in V_0$ and

$$K_0 := \{\lambda \in K : \lambda + \overline{\lambda} = 0\} \cong \mathbb{F}_q$$

be the kernel of the trace. By Remark 8.18, $A := \{t_{\lambda,v} : \lambda \in K_0\} \leq G$ is an elementary abelian p-group, where $p := \mathrm{char}\, K$. For $g \in G_{\langle v\rangle}$ there exists a $\mu \in K$ with $g(v) = \mu v$. By Remark 8.18 we have

$$g t_{\lambda,v} g^{-1} = t_{\lambda,\, g(v)} = t_{\lambda,\, \mu v} = t_{\lambda\mu\overline{\mu},\, v} \in A.$$

This shows $A \trianglelefteq G_{\langle v\rangle}$. Let $w \in V_0 \setminus \{0\}$ be arbitrary. Since G acts transitively on Ω, there exists a $g \in G$ with $g(v) \in \langle w\rangle$. Hence $A^G := \langle gAg^{-1} : g \in G\rangle$ contains all unitary transvections. From Lemma 8.19 it follows that $A^G = G$. By Iwasawa's lemma from group theory, G is simple. □

Example 8.25. The smallest simple group that we have not encountered so far is $\mathrm{SU}(3, 3) = \mathrm{PSU}(3, 3)$ of order $3^3(3^2 - 1)(3^3 + 1) = 2^5 \cdot 3^3 \cdot 7 = 6048$.

Remark 8.26. Wall has shown that a matrix $A \in \mathrm{GL}(n, q^2)$ is conjugate to an element of $\mathrm{GU}(n, q)$ if and only if A and $\overline{A}^{-1}$ are similar (i. e., conjugate in $\mathrm{GL}(n, q^2)$). Moreover, matrices in $\mathrm{GU}(n, q)$ are conjugate in $\mathrm{GU}(n, q)$ if and only if they are similar.

[17] One could now use $\mathrm{PSU}(4, 2) \cong \mathrm{PSp}(4, 3)$ and Lemma 7.15.

Example 8.27.

(i) Let $A \in \mathrm{GL}(n,q)$ be an upper triangular matrix with ones on the main diagonal. Since $A^{q^n} - 1_n = (A-1_n)^{q^n} = 0$, all eigenvalues of A are equal to 1. By the Jordan normal form, the similarity class of A is uniquely determined by $\mathrm{rk}((A-1_n)^k)$ for $k = 1, \ldots, n$. We show $\mathrm{Ker}((A-1_n)^k) = \mathrm{Ker}((A^{-1}-1_n)^k)$ for $k = 1, \ldots, n$. For $k = 1$ we have

$$x \in \mathrm{Ker}(A-1_n) \iff Ax = x \iff A^{-1}x = x \iff x \in \mathrm{Ker}(A^{-1}-1_n).$$

Now suppose the claim has already been proved for k. For $x \in \mathrm{Ker}((A-1_n)^{k+1})$ we have by induction

$$\begin{aligned}(A-1_n)^k(A-1_n)x = 0 &\Longrightarrow (A^{-1}-1_n)^k(A-1_n)x = 0 \Longrightarrow (A-1_n)(A^{-1}-1_n)^k x = 0\\ &\Longrightarrow (A^{-1}-1_n)(A^{-1}-1_n)^k x = 0 \Longrightarrow x \in \mathrm{Ker}((A^{-1}-1_n)^{k+1}).\end{aligned}$$

Hence A and $A^{-1} = \overline{A}^{-1}$ have the same Jordan normal form. By Wall, A is similar to a unitary matrix.

(ii) As is well known, the upper triangular matrices with ones on the main diagonal form a Sylow p-subgroup P of $\mathrm{GL}(n,q)$, where $p \mid q$. For a Sylow p-subgroup Q of $\mathrm{GU}(n,q)$ we have $|Q| = |P|$ by Theorem 8.11. Although by (i) every element of P is conjugate to an element of Q, in general P and Q are not isomorphic. For $(n,q) = (3,2)$ we have $P \cong D_8$ and $Q \cong Q_8$ by Exercise 28.

9 Orthogonal Groups

Remark 9.1. Of the so-called *classical* groups of Lie type, we are only missing the family of orthogonal groups. In this chapter we provide the necessary constructions, but omit many proof details.

Definition 9.2. Let $q \neq 1$ be a prime power, $K := \mathbb{F}_q$ and V an n-dimensional K-vector space. A map $\rho\colon V \to K$ is called a *quadratic form* if

- $\rho(\lambda v) = \lambda^2 \rho(v)$ for all $\lambda \in K$ and $v \in V$,
- the map
$$\beta\colon V \times V \to \mathbb{F}_q, \qquad (u,v) \mapsto \rho(u+v) - \rho(u) - \rho(v)$$
is a (symmetric) bilinear form.

We assume that ρ is *nondegenerate*, i. e., for all $v \in V \setminus \{0\}$ with $\rho(v) = 0$ there exists a $w \in V$ with $\beta(v,w) \neq 0$. We call (V,ρ) an *orthogonal space*. In contrast to the symplectic and unitary groups, the isomorphism type of the *general orthogonal group*

$$\mathrm{GO}(V,\rho) := \big\{ f \in \mathrm{GL}(V) : \forall v \in V : \rho(f(v)) = \rho(v) \big\}$$

depends on ρ. The corresponding matrix group is denoted by $\mathrm{GO}(n,q,\rho)$.

Remark 9.3. In the following, let V always be an orthogonal space. For odd q, ρ is uniquely determined by β, because

$$\rho(v) = \frac{1}{2}\Big(4\rho(v) - \rho(v) - \rho(v)\Big) = \frac{1}{2}\Big(\rho(v+v) - \rho(v) - \rho(v)\Big) = \frac{1}{2}\beta(v,v).$$

In this case, $K^\times \to K^\times$, $x \mapsto x^2$ is a homomorphism with kernel $\langle -1 \rangle$. With $K^\times = \langle \zeta \rangle$ we have $K^\times = \langle \zeta^2 \rangle \dot\cup \zeta \langle \zeta^2 \rangle$ (disjoint union of the squares and nonsquares). The following lemma is a variant of Sylvester's law of inertia.

Lemma 9.4. *Let q be odd. Then there exists an orthogonal basis $b_1, \dots, b_n \in V$ with respect to β such that $\rho(b_i) = 1$ for $i = 1, \dots, n-1$ and $\rho(b_n) \in \{1, \zeta\}$. In particular, β is nondegenerate.*

Proof. For $v \in V^\perp$ (with respect to β) we have $\rho(v) = \frac{1}{2}\beta(v,v) = 0$. Since ρ is nondegenerate, it follows that $v = 0$. Hence β is nondegenerate. We show the existence of $b_1, \dots, b_n$ by induction on n. Let $n = 1$, $V = \langle v \rangle$ and $\lambda := \rho(v)$. Since β is nondegenerate, we have $\lambda \neq 0$. By Remark 9.3 there exists a $\mu \in K^\times$ with $\lambda \in \{\mu^2, \zeta\mu^2\}$. Now $b_1 := \mu^{-1}v$ satisfies the claim.

Let $n \geq 2$. As in the proof of Theorem 8.4 one finds a $u \in V$ with $\rho(u) = \frac{1}{2}\beta(u,u) \neq 0$. As usual, $V = \langle u \rangle \oplus u^\perp$. Let $v \in u^\perp$ with $\rho(v) \neq 0$. The sets

$$\{\lambda^2\rho(u) : \lambda \in K\}, \qquad \{1 - \mu^2\rho(v) : \mu \in K\}$$

each have $(q+1)/2$ elements. Hence they cannot be disjoint. Therefore there exist $\lambda, \mu \in K$ (not both 0) with

$$\rho(\lambda u + \mu v) = \lambda^2\rho(u) + \mu^2\rho(v) = 1.$$

Thus we can put $b_1 := \lambda u + \mu v$. The claim now follows by induction, by restricting ρ to $b_1^\perp$. □

Remark 9.5. Let q be odd.

(i) Lemma 9.4 shows that there are at most two nonequivalent quadratic forms ρ_+ and ρ_-. Here let $\rho_+(b_n) = 1$ and $\rho_-(b_n) = \zeta$. Let $B_\zeta = \operatorname{diag}(1, \dots, 1, \zeta)$ be the Gram matrix of ρ_-. Every form equivalent to ρ_- has Gram matrix $B' := SB_\zeta S^{\mathrm{t}}$ with $S \in \mathrm{GL}(n,q)$. Because of $\det(B') = \det(S)^2\zeta \notin \langle \zeta^2 \rangle$, ρ_+ and ρ_- cannot be equivalent.

(ii) Clearly

$$\mathrm{GO}(n, q, \rho_+) = \left\{A \in \mathrm{GL}(n,q) : A^{\mathrm{t}}A = 1_n\right\}$$

is the set of orthogonal matrices and

$$\mathrm{GO}(n, q, \rho_-) = \left\{A \in \mathrm{GL}(n,q) : A^{\mathrm{t}}B_\zeta A = B_\zeta\right\}.$$

If $n = 2k+1$ is odd, then ρ_+ and $\zeta\rho_-$ are equivalent by (i). Since scaling with ζ does not change the isomorphism type of the orthogonal group, one sets

$$\mathrm{GO}(2k+1, q) := \mathrm{GO}(2k+1, q, \rho_+).$$

(iii) For even $n = 2k$ one defines

$$\mathrm{GO}^\pm(2k, q) := \mathrm{GO}(n, q, \rho_{\pm\epsilon}), \qquad \epsilon \equiv q^k \pmod 4.$$

These two groups already differ in their orders (Theorem 9.7 and Remark 9.10).

Definition 9.6. Let q be odd.

(i) By Remark 9.5(ii) we have $\det(f) = \pm 1$ for all $f \in \mathrm{GO}(V, \rho)$. Hence the *special orthogonal group* $\mathrm{SO}(V, \rho)$ has index 2 in $\mathrm{GO}(V, \rho)$. As in Lemma 7.9 one shows that $\mathrm{Z}(\mathrm{GO}(V, \rho))$ consists only of scalar matrices. For $\lambda \operatorname{id}_V \in \mathrm{GO}(V, \rho)$ we have $\rho(b_1) = \rho(\lambda b_1) = \lambda^2\rho(b_1) \neq 0$ and $\lambda^2 = 1$. Hence $\mathrm{Z}(\mathrm{GO}(V, \rho)) = \langle -\operatorname{id}_V \rangle$ (and $\mathrm{Z}(\mathrm{SO}(2k+1, q, \rho)) = 1$ for $k \geq 1$). As usual, let

$$\mathrm{PSO}(V, \rho) := \mathrm{SO}(V, \rho)/\mathrm{Z}(\mathrm{SO}(V, \rho))$$

be the *projective special orthogonal group*. In contrast to PSL, PSp and PSU, these groups are in general not simple.

(ii) For $v \in V$ with $\rho(v) \neq 0$ one defines the *reflection* (an involution) $\sigma_v \in \mathrm{GO}(V,\rho)$ by

$$\sigma_v(x) := x - \frac{\beta(x,v)}{\rho(v)} v \qquad (v \in V).$$

Every $f \in \mathrm{GO}(V,\rho)$ can be written as a product of reflections $f = \sigma_{v_1} \dots \sigma_{v_r}$. The map

$$\mathrm{GO}(V,\rho) \to K^\times/(K^\times)^2, \qquad f \mapsto \rho(v_1)\dots\rho(v_r)(K^\times)^2$$

is a well-defined homomorphism, which is called the *spinor norm*. Its kernel is denoted by $\Omega(V,\rho)$. The corresponding quotients $\mathrm{P}\Omega(2k-1,q)$ and $\mathrm{P}\Omega^\epsilon(2k,q)$, respectively, have index 2 and $\frac{1}{2}\gcd(4,q^k-\epsilon) \leq 2$, respectively, in $\mathrm{PSO}(n,q,\rho)$. Apart from $\mathrm{P}\Omega^+(4,q)$, these are simple groups for $k \geq 2$. In particular, $\mathrm{P}\Omega = \mathrm{PSO}'$ in this case.

Theorem 9.7. *For odd q we have* $\mathrm{GO}^\pm(2,q) \cong D_{2(q\mp1)}$.

Proof. By Remark 9.5 we have $\mathrm{GO}^\epsilon(2,q) = \{A \in \mathrm{GL}(2,q) : A^{\mathrm{t}}BA = B\}$ with $B = \mathrm{diag}(1,\gamma)$ and $\gamma \in \{1,\zeta\}$. We have $\gamma = 1$ if and only if $q \equiv \epsilon \pmod 4$. For $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ we have $\delta := \det(A) = \pm 1$ by Definition 9.6, and the condition $A^{\mathrm{t}}B = BA^{-1}$ is equivalent to

$$\begin{pmatrix} a & \gamma c \\ b & \gamma d \end{pmatrix} = \delta \begin{pmatrix} d & -b \\ -\gamma c & \gamma a \end{pmatrix},$$

i. e., $d = \delta a$ and $b = -\delta\gamma c$. Clearly $\mathrm{diag}(1,-1) \in \mathrm{GO}^\epsilon(2,q)$ with $\delta = -1$. On the other hand, $\mathrm{SO}^\epsilon(2,q)$ consists of the matrices $M(a,b) := \left(\begin{smallmatrix} a & -\gamma b \\ b & a \end{smallmatrix}\right)$ with $a^2+\gamma b^2 = 1$. Clearly $\mathrm{diag}(1,-1)$ inverts these matrices by conjugation. Hence we have to show $\mathrm{SO}^\epsilon(2,q) \cong C_{q-\epsilon}$. We have

$$M(a,b)\,M(a',b') = M(aa' - \gamma bb', ab' + a'b).$$

There exists a $\tau \in \mathbb{F}_{q^2}$ with $\tau^2 = -\gamma$. Then the map

$$\Gamma\colon \mathrm{SO}^\epsilon(2,q) \to \mathbb{F}_{q^2}^\times, \qquad M(a,b) \mapsto a + \tau b$$

is a homomorphism.

Case 1: $\epsilon = 1$.
For $q \equiv 1 \pmod 4$, -1 is a square, and otherwise $-\zeta$ is a square as a product of two nonsquares. In any case, $-\gamma$ is a square and $\tau \in K$. For $x \in K^\times$ let $a := \frac{x+x^{-1}}{2}$ and $b := \frac{x-x^{-1}}{2\tau}$. Because of

$$a^2 + \gamma b^2 = \frac{x^2 + x^{-2} + 2}{4} - \frac{x^2 + x^{-2} - 2}{4} = 1$$

we have $M(a,b) \in \mathrm{SO}^+(2,q)$. This yields a map

$$\Delta\colon \mathbb{F}_q^\times \to \mathrm{SO}^+(2,q), \qquad x \mapsto M(a,b).$$

with $\Gamma\circ\Delta = \mathrm{id}$. Since $(a+\tau b)(a-\tau b) = a^2+\gamma b^2 = 1$, we also have $\Delta\circ\Gamma = \mathrm{id}$. Hence $\mathrm{SO}^+(2,q) \cong C_{q-1}$.

Case 2: $\epsilon = -1$.
Here $-\gamma$ is not a square and $\tau \notin \mathbb{F}_q$. Since $\{1,\tau\}$ is linearly independent over $\mathbb{F}_q$, Γ is injective. Let $N\colon \mathbb{F}_{q^2}^\times \to \mathbb{F}_q^\times$, $x \mapsto x^{q+1}$ be the norm from Remark 8.3. Since $\tau^q = \tau(-\gamma)^{(q-1)/2} = -\tau$, we have

$$N(a+\tau b) = (a+\tau b)(a+\tau^q b) = (a+\tau b)(a-\tau b) = a^2 + \gamma b^2 = 1$$

for all $a, b \in K$. Hence Γ maps onto $\mathrm{Ker}(N) \cong C_{q+1}$. Conversely, for every $a + \tau b \in \mathrm{Ker}(N)$ we have $M(a,b) \in \mathrm{SO}^-(2,q)$. This shows $\mathrm{SO}^-(2,q) \cong C_{q+1}$. □

Example 9.8. We construct a family of *exceptional* groups of Lie type as subgroups of $\mathrm{GO}(7,q)$ for odd q. To this end, let V be a K-vector space with basis $b_0,\dots,b_7$. On $\{b_0,\dots,b_7\}$ one defines a (nonassociative) multiplication $*$ with identity element $1 := b_0$ and multiplication table

$*$	b_1	b_2	b_3	b_4	b_5	b_6	b_7
b_1	-1	b_3	$-b_2$	b_5	$-b_4$	b_7	$-b_6$
b_2	$-b_3$	-1	b_1	b_6	b_7	$-b_4$	$-b_5$
b_3	b_2	$-b_1$	-1	b_7	$-b_6$	b_5	$-b_4$
b_4	$-b_5$	$-b_6$	$-b_7$	-1	b_1	b_2	b_3
b_5	b_4	$-b_7$	b_6	$-b_1$	-1	$-b_3$	b_2
b_6	$-b_7$	b_4	$-b_5$	$-b_2$	b_3	-1	b_1
b_7	b_6	b_5	b_4	$-b_3$	$-b_2$	$-b_1$	-1

Note that $\langle b_1,b_2\rangle = \{\pm 1,\pm b_1,\pm b_2,\pm b_3\} \cong Q_8$. By extending $*$ distributively to V, one obtains the *octonions* $\mathbb{O} = (V,+,*)$ (also called Cayley *algebra*, although formally it is not even a ring). Now the group

$$\mathrm{G}_2(q) := \big\{ f \in \mathrm{GL}(V) : \forall v,w \in V : f(v*w) = f(v)*f(w) \big\}$$

is simple. As usual one shows that 1 is the only identity element in $\mathbb{O}$. In particular, $f(1) = 1$ for $f \in \mathrm{G}_2(q)$. Let $U := \langle b_1,\dots,b_7\rangle \le V$ and $i \ge 1$. Let $f(b_i) = \lambda 1 + u$ with $\lambda \in K$ and $u \in U$. Then

$$-1 = f(-1) = f(b_i * b_i) = f(b_i) * f(b_i) = \lambda^2 1 + 2\lambda u + u * u.$$

Because of $b_i * b_j = -b_j * b_i$ for $1 \le i < j \le 7$ we have $u * u \in K1$. Comparing coefficients shows $\lambda = 0$ or $u = 0$. In the second case, f would not be injective because of $f(b_i) = f(\lambda 1)$. Hence $f(b_i) = u \in U$ and $f(U) = U$. Clearly $\rho(u) = -u * u = \lambda_1^2 + \ldots + \lambda_7^2$ for $u = \sum \lambda_i b_i$ defines a quadratic form on U. The corresponding bilinear form $(u,v) \mapsto \frac{1}{2}\big(\rho(u+v) - \rho(u) - \rho(v)\big)$ is the standard inner product with respect to $b_1,\dots,b_7$, and hence nondegenerate. Because of

$$\rho(f(u)) = -f(u) * f(u) = f(-u * u) = f(\rho(u)) = \rho(u)$$

we now have $f_{|U} \in \mathrm{GO}(U)$. In this way one can regard $\mathrm{G}_2(q)$ as a subgroup of $\mathrm{GO}(7,q)$. The smallest group of this kind is $\mathrm{G}_2(3)$ of order 4,245,696. For even q, $\mathrm{G}_2(q)$ can be constructed differently, where $\mathrm{G}_2(2)' \cong \mathrm{SU}(3,3)$ has index 2 in $\mathrm{G}_2(2)$.

Remark 9.9. Finally, let q be even. Here β is obviously alternating.

(i) First let $n = 2k+1$ be odd. By Theorem 7.5, β is degenerate, i. e., $V^\perp \ne 0$. Let $u,v \in V^\perp \setminus \{0\}$. After scaling we may assume $\rho(u) = \rho(v)$, since squaring is an automorphism of K. From $\rho(u+v) = \beta(u,v) + \rho(u) + \rho(v) = 0$ we obtain $u = v$, since ρ is nondegenerate. That is, $V^\perp = \langle v\rangle$ and $\overline{V} := V/V^\perp$ is a symplectic space of dimension $2k$. For $f \in \mathrm{GO}(V,\rho)$ and $w \in V$ we have $\beta(f(v),f(w)) = \beta(v,w) = 0$. Hence $V^\perp$ is f-invariant. Suppose that f is the identity on $\overline{V}$. Then for all $w \in V$ there exists a $\lambda \in K$ with $f(w) = w + \lambda v$. Because of

$$\rho(w) = \rho(f(w)) = \rho(w+\lambda v) = \beta(w,\lambda v) + \rho(w) + \rho(\lambda v) = \rho(w) + \lambda^2\rho(v)$$

we have $\lambda^2\rho(v) = 0$ and hence $\lambda = 0$. This shows $f = \mathrm{id}_V$. It turns out that the corresponding monomorphism

$$\Gamma\colon \mathrm{GO}(2k+1,q) \to \mathrm{Sp}(\overline{V}) \cong \mathrm{Sp}(2k,q)$$

is even an isomorphism. Hence $\mathrm{GO}(2k+1,q) \cong \mathrm{Sp}(2k,q)$ is simple for $(k,q) \ne (1,2)$ by Lemma 7.9 and Theorem 7.17.

(ii) Finally, let $n = 2k$ be even. By Theorem 7.5 there exists a symplectic basis $b_1, \ldots, b_n$ with respect to β. However, there are two nonequivalent quadratic forms ρ with the same bilinear form:

$$\rho_+(\sum \lambda_i b_i) := \sum_{i=1}^{k} \lambda_{2i-1}\lambda_{2i},$$
$$\rho_-(\sum \lambda_i b_i) := \sum_{i=1}^{k} \lambda_{2i-1}\lambda_{2i} + \lambda_{2k-1}^2 + \zeta\lambda_{2k}^2,$$

where $\zeta \in K$ is chosen such that $X^2 + X + \zeta \in K[X]$ is irreducible. For odd k one can choose $\zeta = 1$. One defines

$$\mathrm{GO}^{\pm}(2k, q) := \mathrm{GO}(2k, q, \rho_{\pm})$$

(in contrast to Remark 9.5, the sign here does not depend on k and q).

(iii) The computations from Definition 9.6 yield $\mathrm{GO}(V, \rho) = \mathrm{SO}(V, \rho) \cong \mathrm{PSO}(V, \rho)$. For $f \in \mathrm{GO}(V)$ let $\mathrm{Fix}(f) := \{v \in V : f(v) = v\} \leq V$. One shows that the so-called *quasideterminant* (or *Dickson invariant*)

$$\mathrm{GO}(V, \rho) \to \{\pm 1\}, \quad f \mapsto (-1)^{\dim \mathrm{Fix}(f)}$$

is a homomorphism. The kernel $\Omega(n, q, \rho) \leq \mathrm{GO}(n, q, \rho)$ is simple apart from small exceptions (see Remark 9.10). For $n = 2k$ we have $|\mathrm{GO}^{\pm}(2k, q) : \Omega^{\pm}(2k, q)| = 2$. In particular, $\Omega = \mathrm{GO}'$.

Remark 9.10. For all q the following order formulas hold:

$$|\Omega(2k+1, q)| = \frac{q^{k^2}}{\gcd(2, q-1)} \prod_{i=1}^{k} (q^{2i} - 1),$$
$$|\Omega^{\pm}(2k, q)| = \frac{q^{k(k-1)}(q^k \mp 1)}{\gcd(4, q \mp 1)} \prod_{i=1}^{k-1} (q^{2i} - 1).$$

These groups yield "new" families of simple groups only for $n \geq 7$, because

$$\mathrm{P\Omega}(3, q) \cong \mathrm{PSL}(2, q), \qquad \mathrm{P\Omega}^+(4, q) \cong \mathrm{PSL}(2, q)^2, \qquad \mathrm{P\Omega}^-(4, q) \cong \mathrm{PSL}(2, q^2),$$
$$\mathrm{P\Omega}(5, q) \cong \mathrm{PSp}(4, q), \qquad \mathrm{P\Omega}^+(6, q) \cong \mathrm{PSL}(4, q), \qquad \mathrm{P\Omega}^-(6, q) \cong \mathrm{PSU}(4, q).$$

The smallest groups of this kind are $\mathrm{P\Omega}^+(8, 2)$, $\mathrm{P\Omega}^-(8, 2)$ and $\mathrm{P\Omega}(7, 3)$ with the orders 174,182,400, 197,406,720 and 4,585,351,680.

10 Sporadic Groups

Remark 10.1. By the classification of finite simple groups (CFSG), besides C_p, A_n and the groups of Lie type there are 26 further *sporadic* groups, which do not belong to any of the families. In this chapter we construct five sporadic groups, which are at the same time interesting multiply transitive groups. We first choose a path that is as direct as possible, using a "tailor-made" lemma of Witt. Afterwards we describe a combinatorial approach with which other sporadic groups can be constructed as well.

Definition 10.2. A k-transitive action $G \to \mathrm{Sym}(\Omega)$ is called *sharply* k-transitive if for any two k-tuples $(\alpha_1, \ldots, \alpha_k)$ and $(\beta_1, \ldots, \beta_k)$ with pairwise distinct entries there exists *exactly* one $g \in G$ with ${}^g\alpha_i = \beta_i$ for $i = 1, \ldots, k$.

Remark 10.3. Every sharply k-transitive action is faithful. The sharply 1-transitive actions are exactly the regular actions. Let $G \to \mathrm{Sym}(\Omega)$ be transitive, $\omega \in \Omega$ and $k \geq 2$. As in GT, Lemma 6.33 one shows that G acts sharply k-transitively on Ω if and only if G_ω acts sharply $(k-1)$-transitively on $\Omega \setminus \{\omega\}$. Inductively one obtains that a k-transitive action of degree n is sharply k-transitive if and only if $|G| = n(n-1)\dots(n-k+1)$ holds (cf. GT, Lemma 6.34).

Example 10.4.

(i) The natural action of S_n on $\{1, \dots, n\}$ is sharply n-transitive and sharply $(n-1)$-transitive if $n \geq 2$.

(ii) The natural action of A_n is sharply $(n-2)$-transitive if $n \geq 2$ (see GT, Example 6.32).

(iii) Let $n \in \mathbb{N}$, let p be a prime and let $S \leq \mathrm{GL}(n,p)$ be a Singer cycle (GT, Example 6.23). Then $\mathbb{F}_p^n \rtimes S \leq \mathrm{AGL}(n,p)$ is a sharply 2-transitive permutation group on $\mathbb{F}_p^n$.

(iv) By GT, Exercise 49, $\mathrm{SL}(2,2^n)$ acts 3-transitively on the set Ω of all 1-dimensional subspaces of $\mathbb{F}_{2^n}^2$. Since $|\Omega| = 2^n + 1$ and $|\mathrm{SL}(2,2^n)| = (2^{2n}-1)2^n = (2^n+1)2^n(2^n-1)$, this action is sharply 3-transitive.

Lemma 10.5. *Let $\alpha, \beta \in \Omega$, $H \leq G \leq \mathrm{Sym}(\Omega)$ and $a, x \in G$ with the following properties:*

- *$\alpha \neq \beta$, ${}^a\alpha \neq \alpha$ and ${}^x\beta \neq \beta$,*
- *$x \in H$ and $G = \langle H, a \rangle$,*
- *$aH_\beta a = H_\beta$,*
- *H acts k-transitively on $\Omega \setminus \{\alpha\}$ with $k \geq 2$,*
- *$a^2 = x^2 = (ax)^3 = 1$.*

Then G acts $(k+1)$-transitively on Ω and $G_\alpha = H$.

Proof. Since $H \subseteq G_\alpha$, G_α is k-transitive on $\Omega \setminus \{\alpha\}$. Since ${}^a\alpha \neq \alpha$, G is clearly also transitive on Ω. By GT, Lemma 6.33, G acts $(k+1)$-transitively. It remains to show: $G_\alpha \subseteq H$.

For $K := H \cup HaH$ we have $K^{-1} := \{g^{-1} : g \in K\} = K$ since $a^{-1} = a$. Let $z \in H \setminus H_\beta$. Since $k \geq 2$, H_β acts transitively on $\Omega \setminus \{\alpha, \beta\}$. Hence there exists an $h \in H_\beta$ with ${}^{hz}\beta = {}^x\beta$. It follows that $x^{-1}hz \in H_\beta$ and $z \in H_\beta x H_\beta$. This shows $H = H_\beta \cup H_\beta x H_\beta$. The relations $a^2 = x^2 = (ax)^3 = 1$ imply $axa = xax$. By assumption we obtain

$$\begin{aligned} aHa &= aH_\beta a \cup aH_\beta x H_\beta a = H_\beta \cup H_\beta axa H_\beta \\ &= H_\beta \cup H_\beta xax H_\beta \subseteq H \cup HaH = K. \end{aligned}$$

For $g, g' \in HaH$ we therefore have $gg' \in HaHaH \subseteq HKH \subseteq K$. This shows $K \leq G$. Since $a \in K$, we even have $G = \langle H, a \rangle = K$. For every $g \in G \setminus H \subseteq HaH$ we therefore have ${}^g\alpha \neq \alpha$. This proves the claim. □

Lemma 10.6 (Witt)**.** *Let H be a 2-transitive permutation group on $\Delta := \{4, \dots, n\} \ni \omega$ and let $x \in H \setminus H_\omega$ be an involution. Let $a, b, c \in \mathrm{N}_{S_n}(H_\omega)$ be involutions with*

$$a = (1,\omega)(2)(3)\dots, \qquad b = (1,2)(3)(\omega)\dots, \qquad c = (2,3)(1)(\omega)\dots$$

and

$$(ax)^3 = (ba)^3 = (cb)^3 = 1, \qquad (xb)^2 = (xc)^2 = (ac)^2 = 1.$$

Then $G := \langle H, a, b, c\rangle$ *is* 5*-transitive on* $\{1, \dots, n\}$ *and* $G_1 \cap G_2 \cap G_3 = H$.

Proof. By Lemma 10.5 with $(\alpha, \beta) = (1, \omega)$, $K := \langle H, a\rangle$ is 3-transitive on $\Delta \cup \{1\}$ and $K_1 = H$. By GT, Theorem 6.35, H acts primitively on Δ. In particular, $H_\omega < H$ is maximal and $H = \langle H_\omega, x\rangle$. From $x^2 = (xb)^2 = b^2 = 1$ we obtain $xb = bx$. Hence

$$bK_1b = bHb = \langle bH_\omega b, x\rangle = \langle H_\omega, x\rangle = H = K_1.$$

A further application of Lemma 10.5 with $(b, a, 2, 1)$ in place of (a, x, α, β) shows that $L := \langle K, b\rangle$ acts 4-transitively on $\Delta \cup \{1, 2\}$ with $L_2 = K$. From the relations we now again obtain $ac = ca$ and $xc = cx$. Hence

$$cL_2c = cKc = \langle cHc, a\rangle = \langle cH_\omega c, x, a\rangle = \langle H_\omega, x, a\rangle = K = L_2.$$

Finally, a third application of Lemma 10.5 with $(c, b, 3, 2)$ in place of (a, x, α, β) yields that $G = \langle L, c\rangle$ acts 5-transitively on $\{1, \dots, n\}$ with $G_3 = L$. Thus also $G_1 \cap G_2 \cap G_3 = G_1 \cap L_2 = G_1 \cap K = K_1 = H$. □

Theorem 10.7 (Mathieu)**.** *Let*

$$\begin{aligned} a &= (1,4)(7,8)(9,11)(10,12), & b &= (1,2)(7,10)(8,11)(9,12),\\ c &= (2,3)(7,12)(8,10)(9,11), & d &= (4,5,6)(7,8,9)(10,11,12),\\ e &= (4,7,10)(5,8,11)(6,9,12), & f &= (5,7,6,10)(8,9,12,11),\\ g &= (5,8,6,12)(7,11,10,9). \end{aligned}$$

Then $M_{12} := \langle a, b, c, d, e, f, g\rangle \le S_{12}$ *is sharply* 5*-transitive of degree* 12 *and* $M_{11} := \langle a, b, d, e, f, g\rangle$ *is sharply* 4*-transitive of degree* 11.

Proof. Since d permutes the three cycles of e, $E := \langle d, e\rangle$ is elementary abelian of order 9. Moreover, E acts regularly on $\Delta := \{4, \dots, 12\}$. Clearly $f^2 = g^2$ is an involution and $fgf^{-1} = g^{-1}$. For $Q := \langle f, g\rangle$ we therefore have $\langle g\rangle \trianglelefteq Q$ and $|Q : \langle g\rangle| = 2$. Hence $|Q| = 8$ (Q is a quaternion group). A computation shows

$$\begin{aligned} fdf^{-1} &= e, & gdg^{-1} &= (4,8,12)(11,6,7)(9,10,5) = de,\\ fef^{-1} &= d^{-1}, & geg^{-1} &= (4,11,9)(8,6,10)(12,7,5) = de^{-1}. \end{aligned}$$

Hence $Q \subseteq \mathrm{N}_{S_{12}}(E)$ and $H := EQ \le S_{12}$. For order reasons, $E \cap Q = 1$ and thus $|H| = |E||Q| = 9 \cdot 8$. Since E acts regularly on Δ, we have $H_4 = E_4Q = Q$. One easily sees that Q acts transitively on $\Delta \setminus \{4\}$. Hence H is 2-transitive on Δ by GT, Lemma 6.33. Since $|H| = 9 \cdot 8$ and $|\Omega| = 9$, the action is even sharply 2-transitive. We now choose $\omega = 4$ and

$$x := df^2d^{-1} = d(5,6)(7,10)(8,12)(9,11)d^{-1} = (4,6)(7,12)(8,11)(9,10) \in H \setminus H_4$$

To apply Witt's lemma, we first have to show $a, b, c \in \mathrm{N}_{S_{12}}(H_4) = \mathrm{N}_{S_{12}}(Q)$:

$$\begin{aligned} afa^{-1} &= g, & aga^{-1} &= a^2fa^{-2} = f,\\ bfb^{-1} &= f^{-1}, & bgb^{-1} &= (5,11,6,9)(7,12,10,8) = gf,\\ cfc^{-1} &= g^{-1}, & cgc^{-1} &= c^2f^{-1}c^{-2} = f^{-1}. \end{aligned}$$

The relations from Lemma 10.6 are checked as follows:

$$\begin{aligned} ax &= (1,4,6)(7,10,11)(8,9,12), & ba &= (1,4,2)(7,11,12)(8,10,9),\\ cb &= (1,3,2)(7,8,9)(10,12,11), & xb &= (1,2)(4,6)(7,9)(10,12),\\ xc &= (2,3)(4,6)(8,9)(10,11), & ac &= (1,4)(2,3)(7,10)(8,12). \end{aligned}$$

Hence $G := \langle H, a, b, c\rangle = M_{12}$ is 5-transitive on $\Omega = \{1, \dots, 12\}$ and $G_1 \cap G_2 \cap G_3 = H$. Since H acts sharply 2-transitively on Ω, we have $G_1 \cap G_2 \cap G_3 \cap G_4 \cap G_5 = H_4 \cap H_5 = 1$. This shows that G is even sharply 5-transitive. In the proof of Lemma 10.6 we obtained $G_3 = M_{11}$. By Remark 10.3, M_{11} is therefore sharply 4-transitive of degree 11. □

Remark 10.8. In group theory we have shown that the simple group $H = \mathrm{PSL}(3,4)$ of order

$$|H| = \frac{(4^3-1)(4^3-4)(4^3-4^2)}{(4-1)\gcd(3,4-1)} = 2^6 \cdot 3^2 \cdot 5 \cdot 7 = 20{,}160$$

acts 2-transitively on the set Δ of the 21 1-dimensional subspaces of $\mathbb{F}_4^3$ (proof of GT, Lemma 10.7). To distinguish vectors from permutations, we write the elements of $\mathbb{F}_4^3$ in the form $[r,s,t]$ with $r,s,t \in \mathbb{F}_4$. Moreover, let $\mathbb{F}_4^\times = \langle\zeta\rangle$ and $[[r,s,t]] = \mathbb{F}_4[r,s,t] \in \Delta$.

Lemma 10.9. *With the notation from Remark 10.8, the following maps are involutions in* $\mathrm{Sym}(\Delta)$*:*

$$^{\alpha}[[r,s,t]] := [[r^2+st, s^2, t^2]], \qquad ^{\beta}[[r,s,t]] := [[r^2, s^2, t^2\zeta]], \qquad ^{\gamma}[[r,s,t]] := [[r^2, s^2, t^2]]$$

for $r,s,t \in \mathbb{F}_4$.

Proof. Since the values of α, β and γ are homogeneous polynomials of degree 2 in r,s,t, the images of $[[r,s,t]] = \mathbb{F}_4[r,s,t]$ do not depend on the choice of the representative $[r,s,t]$ (i. e., α, β, γ are well-defined). As is well known, $\mathbb{F}_4 \to \mathbb{F}_4$, $x \mapsto x^2$ is the Frobenius automorphism (of order 2). Hence we have

$$\begin{aligned} ^{\alpha^2}[[r,s,t]] &= {}^{\alpha}[[r^2+st, s^2, t^2]] = [[r + s^2t^2 + s^2t^2, s, t]] = [[r,s,t]],\\ ^{\beta^2}[[r,s,t]] &= {}^{\beta}[[r^2, s^2, t^2\zeta]] = [[r, s, t\zeta^2\zeta]] = [[r,s,t]],\\ ^{\gamma^2}[[r,s,t]] &= [[r,s,t]]. \end{aligned}$$

Consequently α, β, γ are invertible and have order 2. □

Theorem 10.10 (Mathieu)**.** *Let* $\mathrm{PSL}(3,4) \cong H \leq \mathrm{Sym}(\Delta)$ *be as in Remark 10.8 and let* $\Omega = \Delta \dot\cup \{1,2,3\}$*. With the notation from Lemma 10.9, let* $a := (1, [[1,0,0]])\alpha$*,* $b := (1,2)\beta$ *and* $c := (2,3)\gamma$*. Then:*

(i) $M_{22} := \langle H, a\rangle$ *is 3-transitive on* $\Delta \cup \{1\}$*.*

(ii) $M_{23} := \langle H, a, b\rangle$ *is 4-transitive on* $\Omega \setminus \{3\}$*.*

(iii) $M_{24} := \langle H, a, b, c\rangle$ *is 5-transitive on* Ω*.*

Proof. We use Witt's lemma with $\omega = [[1,0,0]] \in \Delta$ and

$$x := \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbb{F}_4^\times \in \mathrm{PSL}(3,4).$$

By Lemma 10.9, a, b, c, x are involutions. Clearly ${}^b\omega = {}^c\omega = \omega$. Let F be the Frobenius automorphism on H (or $\mathrm{SL}(3,4)$). For $h \in H$ and $[[r,s,t]] \in \Delta$ we have ${}^{chc}[[r,s,t]] = {}^c[h[r^2,s^2,t^2]] = [F(h)[r,s,t]]$. This shows $cHc^{-1} = H$ and $c \in \mathrm{N}_{S_{24}}(H_\omega)$. Let $y := \mathrm{diag}(1,1,\zeta) \in \mathrm{GL}(3,4)$. Then

$${}^{bhb}[[r,s,t]] = {}^b[h[r^2,s^2,t^2\zeta]] = [yF(h)y^{-1}[r,s,t]].$$

Since $yF(h)y^{-1} \in H$, we also have $b \in \mathrm{N}_{S_{24}}(H_\omega)$. Moreover, ${}^{aha}\omega = {}^{ah}1 = {}^a1 = \omega$. An element $h \in H_\omega$ has the form ${}^h[[r,s,t]] = [[r+\sigma s+\tau t, \rho_1 s+\rho_2 t, \rho_3 s+\rho_4 t]]$ with $\rho_1\rho_4+\rho_2\rho_3 = \det(h) = 1$. A computation shows

$$\begin{aligned} {}^{aha}[[r,s,t]] &= {}^{ah}[[r^2+st, s^2, t^2]] = {}^a[[r^2+st+\sigma s^2+\tau t^2, \rho_1 s^2+\rho_2 t^2, \rho_3 s^2+\rho_4 t^2]] \\ &= [[r+s^2t^2+\sigma^2 s+\tau^2 t+(\rho_1 s^2+\rho_2 t^2)(\rho_3 s^2+\rho_4 t^2), \rho_1^2 s+\rho_2^2 t, \rho_3^2 s+\rho_4^2 t]] \\ &= [[r+(\sigma^2+\rho_1\rho_3)s+(\tau^2+\rho_2\rho_4)t, \rho_1^2 s+\rho_2^2 t, \rho_3^2 s+\rho_4^2 t]]. \end{aligned}$$

Hence aha corresponds to the matrix

$$\begin{pmatrix} 1 & \sigma^2+\rho_1\rho_3 & \tau^2+\rho_2\rho_4 \\ 0 & \rho_1^2 & \rho_2^2 \\ 0 & \rho_3^2 & \rho_4^2 \end{pmatrix} \in H_\omega,$$

because $\det(aha) = \rho_1^2\rho_4^2 + \rho_2^2\rho_3^2 = \det(h)^2 = 1$. Again it follows that $a \in \mathrm{N}_{S_{24}}(H_\omega)$.

Clearly $xb = bx$, $xc = cx$ and $ac = ca$. This shows $(xb)^2 = (xc)^2 = (ac)^2 = 1$. It remains to show: $(ax)^3 = (ba)^3 = (cb)^3 = 1$. This follows from the following computations:

$$\begin{aligned} {}^{cbc}[[r,s,t]] &= {}^c[[r,s,t\zeta]] = [[r^2,s^2,t^2\zeta^2]] = {}^b[[r,s,t\zeta^2]] = {}^{bcb}[[r,s,t]], \\ {}^{bab}[[r,s,t]] &= {}^b[[r+s^2t^2\zeta, s, t\zeta^2]] = [[r^2+st\zeta^2, s^2, t^2\zeta^2]] \overset{1+\zeta=\zeta^2}{=} {}^a[[r+s^2t^2, s, t\zeta]] = {}^{aba}[[r,s,t]], \\ {}^{axa}1 &= {}^a[[0,1,0]] = [[0,1,0]] = {}^{xax}1, \\ {}^{axa}[[1,0,0]] &= [[1,0,0]] = {}^{xax}[[1,0,0]], \\ {}^{axa}[[0,1,0]] &= 1 = {}^{xax}[[1,0,0]], \\ {}^{axa}[[r,s,t]] &= {}^a[[s^2, r^2+st, t^2]] = [[s+r^2t^2+st^3, r+s^2t^2, t]] \\ &= \left\{ \begin{matrix} [[r^2, s^2+rt, t^2]] \text{ if } t \neq 0, \\ [[s,r,0]] \text{ if } t = 0 \neq rs \end{matrix} \right\} = {}^x[[s^2+rt, r^2, t^2]] = {}^{xax}[[r,s,t]]. \end{aligned}$$

Witt's lemma now shows that M_{24} acts 5-transitively on Ω. From the proof of Witt's lemma we obtain $(M_{24})_3 = M_{23}$ and $M_{22} = (M_{23})_2$. Thus the remaining statements follow. □

Definition 10.11. The groups M_{11}, M_{12}, M_{22}, M_{23} and M_{24} are called the *Mathieu groups* of degree 11, 12, 22, 23 and 24, respectively. From Remark 10.3 we obtain

$$\begin{aligned} |M_{11}| &= 11 \cdot 10 \cdot 9 \cdot 8 = 7{,}920 = 2^4 \cdot 3^2 \cdot 5 \cdot 11, \\ |M_{12}| &= 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95{,}040 = 2^6 \cdot 3^3 \cdot 5 \cdot 11, \\ |M_{22}| &= 22|\mathrm{PSL}(3,4)| = 443{,}520 = 2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11, \\ |M_{23}| &= 23|M_{22}| = 10{,}200{,}960 = 2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 23, \\ |M_{24}| &= 24|M_{23}| = 244{,}823{,}040 = 2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23. \end{aligned}$$

Lemma 10.12. *Let $G \leq \mathrm{Sym}(\Omega)$ be 3-transitive of degree $d = |\Omega| \geq 5$. If G_ω is simple for some $\omega \in \Omega$, then G is simple or d is a power of 2.*

Proof. The proof works as for the simplicity of the alternating groups. Let $1 \neq N \trianglelefteq G$. By GT, Theorem 6.20, N acts transitively. The Frattini argument yields $G = G_\omega N$. We may assume indirectly that $G_\omega \not\subseteq N$. From the simplicity of G_ω it follows that $N_\omega = G_\omega \cap N = 1$. Hence N is a regular normal subgroup and $|N| = d \geq 5$. By GT, Lemma 6.19, G_ω acts 2-transitively on $N \setminus \{1\}$ by conjugation. From GT, Theorem 6.36 it follows that $d = |N| = 2^k$ for some $k \in \mathbb{N}$. □

Example 10.13. Let $d = 2^n \geq 8$ and $\Omega := \mathbb{F}_2^n$. By linear algebra, the simple group $\mathrm{GL}(n, 2)$ acts 2-transitively on $\Omega \setminus \{0\}$. Hence the nonsimple group $\mathrm{AGL}(n, 2) = \mathbb{F}_2^n \rtimes \mathrm{GL}(n, 2)$ acts 3-transitively on Ω with stabilizer $G_0 = \mathrm{GL}(n, 2)$ (GT, Example 6.23).

Theorem 10.14. *The Mathieu groups M_{11}, M_{12}, M_{22}, M_{23} and M_{24} are simple.*

Proof (Chapman). First let $G = M_{11}$ and $P \in \mathrm{Syl}_{11}(G)$. In S_{11} there are 10! elements of order 11, which are distributed among 9! Sylow subgroups. Hence $|\mathrm{N}_{S_{11}}(P)| = 11 \cdot 10$ and $|\mathrm{N}_G(P) : P|$ divides 10. By Sylow,

$$5 \equiv 10 \cdot 9 \cdot 8 = \frac{|G|}{11} = |G : \mathrm{N}_G(P)||\mathrm{N}_G(P) : P| \equiv |\mathrm{N}_G(P) : P| \pmod{11}$$

and thus $|\mathrm{N}_G(P) : P| = 5$ and $|G : \mathrm{N}_G(P)| = 16 \cdot 9$. Now let $1 \neq N \trianglelefteq G$. Then N is transitive and all 11-Sylow subgroups of G lie in N. In particular, $|N : \mathrm{N}_N(P)| = 16 \cdot 9$. This shows $|G : N| \leq 5$. Let us suppose $|G : N| = 5$. Then N has exactly $16 \cdot 9 \cdot 10$ elements of order 11. The remaining $|N|/11$ elements must then form the stabilizer N_1. In particular, $N_1 = \ldots = N_{11}$. But then N cannot act faithfully. Hence $G = N$ and G is simple.

The simplicity of M_{12} now follows from Lemma 10.12, since M_{11} is a stabilizer of M_{12}. Since $\mathrm{PSL}(3, 4)$ is simple, one can also show the simplicity of M_{22}, M_{23} and M_{24} using Lemma 10.12. □

Remark 10.15.

(i) Occasionally one defines the Mathieu groups M_9, M_{10}, M_{20} and $M_{21} \cong \mathrm{PSL}(3, 4)$ as suitable stabilizers of the larger Mathieu groups. However, they are not sporadic simple groups. In the proof of Theorem 10.7 we constructed $M_9 \cong C_3^2 \rtimes Q_8$ as a sharply 2-transitive group of degree 9. In particular, M_9 is a Frobenius group.

(ii) One can show (elementarily) that S_k, S_{k+1}, A_{k+2}, M_{11} and M_{12} are the only sharply k-transitive permutation groups with $k \geq 4$.[18] Using the CFSG, one was able to show that S_n, A_n, M_{11}, M_{12}, M_{23} and M_{24} are the only 4-transitive permutation groups.

(iii) One can also generate each of the Mathieu groups with just two permutations, but then it is difficult to determine the structure (let alone the order). In fact, Mathieu's work of 1861 met with incomprehension for a long time. For instance, Miller claimed in 1898 that M_{24} did not exist. By the CFSG, we know that every finite simple group can be generated by an involution and one further element of prime order.

(iv) The following definition can be motivated as follows: How many lottery tickets does one have to buy if one wants to be guaranteed four "correct numbers"? In the best case, every combination of 4 numbers appears on only one lottery ticket (whether this is possible, we will see in Example 10.19).

[18] See the lecture notes on permutation groups.

Definition 10.16. A $((t,k,v)$-$)$*Steiner system* is a pair $S = (\Omega, \mathcal{B})$ with the following properties:

- Ω is a v-element set of "points" (vertices).
- $\mathcal{B}$ is a set of k-element subsets of Ω called "blocks".[19]
- Every t-element subset of Ω lies in exactly one block of $\mathcal{B}$.

We call

$$\mathrm{Aut}(S) := \{\sigma \in \mathrm{Sym}(\Omega) : \forall B \in \mathcal{B} : \sigma(B) \in \mathcal{B}\} \leq \mathrm{Sym}(\Omega)$$

the *automorphism group* of S.

Example 10.17.

(i) (t,k,v)-Steiner systems can clearly exist only for $t \leq k \leq v$. In the case $t = k$, $\mathcal{B}$ is the set of all k-element subsets of Ω and $\mathrm{Aut}(S) = \mathrm{Sym}(\Omega)$. In the case $k = v$ we have $\mathcal{B} = \{\Omega\}$ and $\mathrm{Aut}(S) = \mathrm{Sym}(\Omega)$. These Steiner systems are called *trivial*. Hence we may assume $t < k < v$.

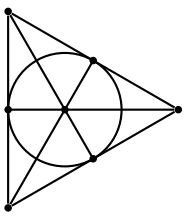
(ii) In the case $t = 1$, $\mathcal{B}$ is a partition of Ω and $k \mid v$ follows. The number of these Steiner systems is $\frac{v!}{(k!)^u u!}$ with $u := v/k$.[20] We see in Lemma 10.18 that there are in general strong restrictions on t, k, v.

(iii) Let q be a prime power, $n \geq 2$, $\Omega = \mathbb{F}_q^n$ and

$$\mathcal{B} = \{\mathbb{F}_q v + w : v, w \in \Omega, v \neq 0\}$$

the set of "lines" on Ω (think of $\mathbb{F}_q$ as being replaced by $\mathbb{R}$). Since any two distinct points lie on exactly one line, $S = (\Omega, \mathcal{B})$ is a $(2, q, q^n)$-Steiner system. We call S the *affine plane* over $\mathbb{F}_q^n$.[21]

(iv) Let $n \geq 3$, $\Omega = \{\mathbb{F}_q v : v \in \mathbb{F}_q^n \setminus \{0\}\}$ be the set of 1-dimensional subspaces of $\mathbb{F}_q^n$ and $\mathcal{B}$ the set of 2-dimensional subspaces of $\mathbb{F}_q^n$ (formally: every block consists of the 1-dimensional subspaces of a 2-dimensional space). Since any two linearly independent vectors span a 2-dimensional subspace, $S = (\Omega, \mathcal{B})$ is a $(2, q+1, \frac{q^n-1}{q-1})$-Steiner system. We call S the *projective plane* over $\mathbb{F}_q^n$. For $(q,n) = (2,3)$ we obtain the *Fano plane*:



For $(q,n) = (2,4)$ we obtain a solution of Kirkman's *schoolgirl problem*: 15 girls are to walk in groups of three on seven consecutive days such that any two girls are in the same group of three only once.

Lemma 10.18. *Let $S = (\Omega, \mathcal{B})$ be a (t,k,v)-Steiner system and $0 \leq s \leq t$. Then:*

(i) The number of blocks containing a given s-element subset of Ω is

$$\gamma_s := \frac{(v-s)(v-s-1)\dots(v-t+1)}{(k-s)(k-s-1)\dots(k-t+1)}.$$

[19] This has nothing to do with the blocks of an imprimitive action.
[20] See the lecture notes on Discrete Mathematics.
[21] See the lecture notes on Synthetic Geometry.

(ii) $|\mathcal{B}| = \gamma_0 = \frac{v\gamma_1}{k}$.

(iii) (FISHER*'s inequality*) $v \leq |\mathcal{B}|$ *and* $k \leq \gamma_1$.

Proof.

(i) Every s-element subset lies in exactly $\binom{v-s}{t-s}$ t-element subsets of Ω. By definition, every t-element subset lies in exactly one block. Every block contains exactly $\binom{k-s}{t-s}$ s-element subsets. This shows

$$\gamma_s = \frac{\binom{v-s}{t-s}}{\binom{k-s}{t-s}} = \frac{(v-s)(v-s-1)\dots(v-t+1)}{(k-s)(k-s-1)\dots(k-t+1)}.$$

(ii) Follows from (i).

(iii) Let $\Omega = \{\omega_1, \dots, \omega_v\}$ and $\mathcal{B} = \{B_1, \dots, B_r\}$. Let $M := (m_{ij}) \in \mathbb{Z}^{v\times r}$ be the incidence matrix of S, i. e., $m_{ij} = 1$ if $\omega_i \in B_j$ and $m_{ij} = 0$ otherwise. Then

$$MM^t = \Big(\sum_{l=1}^{r} m_{il}m_{jl}\Big)_{ij} = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_2 \\ \gamma_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \gamma_2 \\ \gamma_2 & \cdots & \gamma_2 & \gamma_1 \end{pmatrix} = (\gamma_1 - \gamma_2)1_v + \gamma_2 J,$$

where $J = (1) \in \mathbb{Z}^{v\times v}$ is the matrix consisting entirely of ones. By (i), $\gamma_1 = \frac{v-1}{k-1}\gamma_2 > \gamma_2$. The eigenvalues of J are well known to be 0 with multiplicity $v-1$ and v with multiplicity 1. Hence $\gamma_1 - \gamma_2 > 0$ and $\gamma_1 - \gamma_2 + \gamma_2 v$ are the eigenvalues of MM^t. In particular, MM^t is invertible. This shows

$$v = \mathrm{rk}(MM^{\mathrm{t}}) \leq \mathrm{rk}(M) \leq r = |\mathcal{B}|.$$

From (ii) we obtain $k = \frac{v\gamma_1}{r} \leq \gamma_1$. □

Example 10.19. For the optimal solution of the lottery problem from Remark 10.15 one needs a $(4, 6, 49)$-Steiner system. Since $\gamma_3 = \frac{46}{3} \notin \mathbb{N}$, this does not exist. In any case, one needs at least

$$\gamma_0 = \frac{49 \cdot 48 \cdot 47 \cdot 46}{6 \cdot 5 \cdot 4 \cdot 3} > 14{,}125$$

lottery tickets.[22]

Remark 10.20.

(i) For the (alternative) construction of the Mathieu groups, one starts with the affine plane S over $\mathbb{F}_3^2$ as a $(2, 3, 9)$-Steiner system. Then one extends S by adding three points to a $(5, 6, 12)$-Steiner system $\hat{S}$ (the details are extremely laborious[23]). Afterwards one defines $M_{12} := \mathrm{Aut}(\hat{S})$. The larger Mathieu groups can also be obtained in this way. In general, MENDELSOHN proved that every finite group is the automorphism group of a $(2, 3, v)$-Steiner system and of a $(3, 4, v)$-Steiner system.

[22] One can show that at least 14,749 lottery tickets are necessary. However, it is open whether this actually suffices. See `https://ljcr.dmgordon.org/show_cover.php?v=49&k=6&t=4`.

[23] See [Dixon-Mortimer, *Permutation Groups*, Springer, New York, 1996]. The errata mention that the proof of Theorem 6.3B is incomplete.

(ii) For M_{22} one uses a $(3,6,22)$-Steiner system $S = (\Omega, \mathcal{B})$ (we have $\mathrm{Aut}(S) \cong M_{22} \rtimes C_2$). By Lemma 10.18 we have

$$|\mathcal{B}| = \frac{22 \cdot 21 \cdot 20}{6 \cdot 5 \cdot 4} = 77.$$

Let Γ be a graph with vertex set $\Gamma_E := \Omega \cup \mathcal{B} \cup \{\zeta\}$, so $|\Gamma_E| = 22 + 77 + 1 = 100$. The vertex ζ is adjacent to all $\omega \in \Omega$. Two blocks are adjacent in Γ if and only if they are disjoint. All other edges have the form $(\omega, B) \in \Omega \times \mathcal{B}$ with $\omega \in B$. By Exercise 33, Γ is a 22-regular graph. It is called the *Higman-Sims graph*. We have $\mathrm{Aut}(\Gamma) = HS \rtimes C_2$, where HS is the sporadic simple *Higman-Sims group* of order 44,352,000.

(iii) The Mathieu groups can also be constructed using methods from coding theory. The (extended binary) *Golay code* is the 12-dimensional subspace $C \leq \mathbb{F}_2^{24}$ spanned by the rows of the matrix

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

One can show that

$$\mathrm{Aut}(C) := \{\sigma \in S_{24} : \forall (c_i) \in C : (c_{\sigma(i)}) \in C\}$$

is isomorphic to M_{24}. The vectors (*codewords*) in C with exactly eight ones correspond to the blocks of a $(5,8,24)$-Steiner system (the positions of the ones are the blocks). The Golay code was used for the data transmission of the Voyager space probes.

(iv) The automorphism groups and Schur multipliers are

G	M_{11}	M_{12}	M_{22}	M_{23}	M_{24}
$\mathrm{Out}(G)$	1	C_2	C_2	1	1
$M(G)$	1	C_2	C_{12}	1	1

(without proof). For all sporadic simple groups S we have $|\mathrm{Out}(S)| \leq 2$. Three attempts were needed for the correct computation of $M(M_{22})$. "Worse" is only $M(M_{21}) = M(\mathrm{PSL}(3,4)) \cong C_{12} \times C_4$. In GAP one can, for example, do the following:

```
G:=MathieuGroup(24);;
AllPrimitiveGroups(NrMovedPoints,24,Transitivity,5); #gives only one
PrimitiveIdentification(G); #G=PrimitiveGroup(24,1)
IsSimple(G);
H:=Stabilizer(G,24); #= M23

LoadPackage("guava",false); #package for codes
C:=ExtendedBinaryGolayCode();
Display(GeneratorMat(C));
A:=AutomorphismGroup(C);
Transitivity(A); #5-transitiv
c:=Positions(Basis(C)[1],Z(2)^0); #codeword with exactly eight ones
B:=Orbit(A,c,OnSets);; #blocks of the (5,8,24)-Steiner system
Size(B); #759 = (24·23·22·21·20)/(8·7·6·5·4)
```

(v) Closely related to the Golay code is the *Leech lattice* $L \leq \mathbb{Z}^{24}$, which consists of the integer linear combinations of the rows of the following matrix:

$$\begin{pmatrix}
8 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & . & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & . & . & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & . & . & . & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & . & . & . & . & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & . & . & . & . & . & . & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & . & . & . & . & . & . & . & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & . & . & . & . & . & . & . & . & 4 & . & . & . & . & . & . & . & . & . & . & . & . & . \\
2 & 2 & 2 & 2 & . & . & . & . & 2 & 2 & 2 & 2 & . & . & . & . & . & . & . & . & . & . & . & . \\
4 & . & . & . & . & . & . & . & . & . & . & . & 4 & . & . & . & . & . & . & . & . & . & . & . \\
2 & 2 & . & . & 2 & 2 & . & . & 2 & 2 & . & . & 2 & 2 & . & . & . & . & . & . & . & . & . & . \\
2 & . & 2 & . & 2 & . & 2 & . & 2 & . & 2 & . & 2 & . & 2 & . & . & . & . & . & . & . & . & . \\
2 & . & . & 2 & 2 & . & . & 2 & 2 & . & . & 2 & 2 & . & . & 2 & . & . & . & . & . & . & . & . \\
4 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & 4 & . & . & . & . & . & . & . \\
2 & . & 2 & . & 2 & . & . & 2 & 2 & 2 & . & . & . & . & . & . & 2 & 2 & . & . & . & . & . & . \\
2 & . & . & 2 & 2 & 2 & . & . & 2 & . & 2 & . & . & . & . & . & 2 & . & 2 & . & . & . & . & . \\
2 & 2 & . & . & 2 & . & 2 & . & 2 & . & . & 2 & . & . & . & . & 2 & . & . & 2 & . & . & . & . \\
. & 2 & 2 & 2 & 2 & . & . & . & 2 & . & . & . & 2 & . & . & . & 2 & . & . & . & 2 & . & . & . \\
. & . & . & . & . & . & . & . & 2 & 2 & . & . & 2 & 2 & . & . & 2 & 2 & . & . & 2 & 2 & . & . \\
. & . & . & . & . & . & . & . & 2 & . & 2 & . & 2 & . & 2 & . & 2 & . & 2 & . & 2 & . & 2 & . \\
-3 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{pmatrix}$$

Let $A := \mathrm{Aut}(L) := \{f \in \mathrm{O}(\mathbb{R}^{24}) : f(L) = L\}$, where $\mathrm{O}(\mathbb{R}^n)$ is the orthogonal group (matrices x with $xx^{\mathrm{t}} = 1_n$). By Exercise 35, A is finite. We have $\mathrm{Z}(A) = \langle -1_{24} \rangle$ and $Co_1 := A/\mathrm{Z}(A)$ is called the first *Conway group* (A is a Schur extension of Co_1). The following vectors lie in L:

$$\begin{aligned} u &:= (4, 4, 0, \ldots, 0), \\ v &:= (4, -4, 0, \ldots, 0), \\ w &:= (5, 1, \ldots, 1). \end{aligned}$$

We call $Co_2 := \mathrm{C}_A(u)$ and $Co_3 := \mathrm{C}_A(w)$ the second and third Conway group, respectively. Moreover, $McL := \mathrm{C}_A(u, w)$ is the *McLaughlin group* and $\mathrm{C}_A(v, w) \cong HS$. All of them are sporadic simple groups with orders

$$\begin{aligned} |Co_1| &= 2^{21} \cdot 3^9 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 23 = 4{,}157{,}776{,}806{,}543{,}360{,}000, \\ |Co_2| &= 2^{18} \cdot 3^6 \cdot 5^3 \cdot 7 \cdot 11 \cdot 23 = 42{,}305{,}421{,}312{,}000, \\ |Co_3| &= 2^{10} \cdot 3^7 \cdot 5^3 \cdot 7 \cdot 11 \cdot 23 = 495{,}766{,}656{,}000, \\ |McL| &= 2^7 \cdot 3^6 \cdot 5^3 \cdot 7 \cdot 11 = 898{,}128{,}000. \end{aligned}$$

(vi) Only more than 100 years after Mathieu's work did Janko discover the next sporadic simple group J_1. It is most easily defined by an $\mathbb{F}_{11}$-representation:

$$J_1 := \left\langle \begin{pmatrix} . & 1 & . & . & . & . & . \\ . & . & 1 & . & . & . & . \\ . & . & . & 1 & . & . & . \\ . & . & . & . & 1 & . & . \\ . & . & . & . & . & 1 & . \\ . & . & . & . & . & . & 1 \\ 1 & . & . & . & . & . & . \end{pmatrix}, \begin{pmatrix} -3 & 2 & -1 & -1 & -3 & -1 & -3 \\ -2 & 1 & 1 & 3 & 1 & 3 & 3 \\ -1 & -1 & -3 & -1 & -3 & -3 & 2 \\ -1 & -3 & -1 & -3 & -3 & 2 & -1 \\ -3 & -1 & -3 & -3 & 2 & -1 & -1 \\ 1 & 3 & 3 & -2 & 1 & 1 & 3 \\ 3 & 3 & -2 & 1 & 1 & 3 & 1 \end{pmatrix} \right\rangle \leq \mathrm{GL}(7, 11).$$

We have $|J_1| = 2^3 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19 = 175{,}560$.

(vii) The second Janko group J_2 (also called the *Hall-Janko group* HJ) of order 604,800 is constructed similarly to HS as a permutation group of degree $100 = 63+36+1$. Let C be the set of involutions

and $\mathcal{U}$ the set of subgroups of order 168 of $G = \mathrm{SU}(3,3)$ (both are conjugacy classes). We have $|C| = 63$ and $|\mathcal{U}| = |G : U| = 36$ as well as $U \cong \mathrm{GL}(3,2)$ for $U \in \mathcal{U}$. Let Γ be a graph with vertex set $\Gamma_E := C \cup \mathcal{U} \cup \{\gamma\}$, where γ is adjacent to every $U \in \mathcal{U}$. Two involutions $x, y \in C$ are adjacent if and only if $|\langle xy \rangle| = 4$ (i. e., $\langle x, y \rangle \cong D_8$). Two subgroups $U_1, U_2 \in \mathcal{U}$ are adjacent if and only if $|U_1 \cap U_2| = 24$. Finally, $(x, U) \in C \times \mathcal{U}$ form an edge if and only if $x \in U$. Now $\mathrm{Aut}(\Gamma) \cong J_2 \rtimes C_2$ (see Exercise 34). Thus we have "constructed" all simple groups of order $\leq 10^7$ (see GT, Appendix). The sporadic groups J_3 and J_4 are also named after Janko, where J_4 is the most recently constructed simple group of all (construction in 1977 by Norton).

(viii) Presentations, permutation and matrix representations of all sporadic groups can be looked up at `https://brauer.maths.qmul.ac.uk/Atlas/v3/` or obtained directly with GAP:

```
LoadPackage("atlasrep");
DisplayAtlasInfo("M24"); #overview of representations, needs internet
AtlasGroup("M24",Dimension,11,Ring,GF(2));
prog:=AtlasProgram("M24","presentation");;
slp:=StraightLineProgramFromStraightLineDecision(prog.program);;
F:=FreeGroup(2);;
rels:=ResultOfStraightLineProgram(slp,GeneratorsOfGroup(F));;
G:=F/rels;;
PresentationFpGroup(G);

ct:=CharacterTable("M"); #character table of the Monster
PrintFactorsInt(Size(ct)); #prime factorization of the order
NrConjugacyClasses(ct);
SizesConjugacyClasses(ct);
OrdersClassRepresentatives(ct); #orders of the elements up to conjugation
```

(ix) Only in 2014 was Keevash able to show that nontrivial (t, k, v)-Steiner systems with $t \geq 6$ exist. In fact, the divisibility conditions from Lemma 10.18 are sufficient for the existence of a corresponding Steiner system in "most" cases. The proof is probabilistic and not constructive.[24] The number of $(2, 3, v)$-Steiner systems for $v \leq 19$ can be looked up at OEIS.

11 Coxeter Groups

Remark 11.1. The groups generated by two involutions are well known to be exactly the dihedral groups. In this chapter we study groups generated by finitely many involutions. These occur as symmetry groups of higher-dimensional spaces and lead to the classification of the (finite) reflection groups.

Definition 11.2. A group of the form

$$G = \langle x_1, \ldots, x_n \mid (x_i x_j)^{m_{ij}} = 1, 1 \leq i \leq j \leq n \rangle$$

with $2 \leq m_{ij} \leq \infty$ for $i < j$ and $m_{ii} = 1$ for $i = 1, \ldots, n$ is called a *Coxeter group* of *rank* n.

Remark 11.3.

(i) In the following, let $G = \langle x_1, \ldots, x_n \rangle$ always be a Coxeter group (we identify the generators $x_1, \ldots, x_n$ of the free group with the corresponding cosets in G). The relations of the form $(x_i x_j)^\infty = 1$ have no meaning and can be ignored. From $m_{ii} = 1$ it follows that $x_i^2 = 1$ for

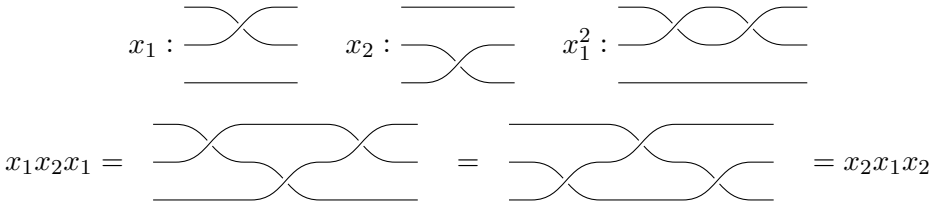
[24] On this, a popular science article: wired.com

$i = 1, \ldots, n$. Because of $(x_j x_i)^{m_{ij}} = (x_i x_j)^{-m_{ij}} = 1$ we define $m_{ji} := m_{ij}$ for $i < j$. By von Dyck's theorem there exists a homomorphism $f\colon G \to \{\pm 1\}$ with $f(x_i) = -1$ for $i = 1, \ldots, n$. Hence $x_1, \ldots, x_n$ are involutions in G. The equation $m_{ij} = 2$ says that x_i and x_j commute.

(ii) The relation $(x_i x_j)^{m_{ij}} = 1$ is equivalent to $x_i x_j x_i \ldots = x_j x_i x_j \ldots$, where there are exactly m_{ij} factors on both sides. Let $(x_i x_j)_n := x_i x_j x_i \ldots$ be an alternating product of n factors. If one omits the condition $m_{ii} = 1$ in the definition of Coxeter groups, one obtains *Artin groups*:

$$\langle x_1, \ldots, x_n \mid (x_i x_j)_{m_{ij}} = (x_j x_i)_{m_{ij}}, 1 \le i, j \le n \rangle$$

with $2 \le m_{ij} \le \infty$. The choice $m_{ij} = \infty$ for all i, j yields the free group F_n. By von Dyck's theorem, every Coxeter group is a quotient group of an Artin group. In the special case $m_{ij} = 3$ for $|i - j| = 1$ and $m_{ij} = 2$ for $|i - j| > 1$ one speaks of a *braid group*. In contrast to S_n (Theorem 2.18), we have $m_{ii} = \infty$ instead of $m_{ii} = 1$. One can realize the elements of the braid group by "braids" with n "strands". The composition is the "gluing" of the strands at their ends:



However, we will not go into these groups any further.

Example 11.4.

(i) $G = \langle x, y \mid x^2 = y^2 = (xy)^m = 1 \rangle \cong D_{2m}$ is a Coxeter group.

(ii) $G = \langle x_1, \ldots, x_n \mid x_i^2 = (x_i x_j)^2 = 1, i < j \rangle \cong C_2^n$ is a Coxeter group.

(iii) By Theorem 2.18, the symmetric groups are Coxeter groups with $x_i = (i, i+1)$ for $i = 1, \ldots, n-1$.

(iv) In the case $m_{ij} = \infty$ for all $i < j$, one calls G the *universal* Coxeter group of rank n. For $n = 2$ we obtain D_∞. Every Coxeter group is a quotient group of a universal Coxeter group.

(v) A *reflection* $\sigma \in \mathrm{GL}(\mathbb{R}^d)$ is a map of the form $\sigma_b(v) := v - 2[v, b]b$ for a unit vector $b \in \mathbb{R}^d$ (here $[v, b]$ is the standard inner product). A *reflection group* is a finite subgroup $S \le \mathrm{GL}(\mathbb{R}^d)$ generated by reflections $\sigma_1, \ldots, \sigma_n$. By von Dyck's theorem, S is a quotient group of a Coxeter group. In Theorem 11.41 we will show that S is in fact isomorphic to a Coxeter group (i. e., all further relations in S follow from the relations $(\sigma_i \sigma_j)^{m_{ij}} = 1$).

(vi) Let S be a finite nonabelian simple group. By the Feit-Thompson theorem, S has an involution s. Clearly, S is generated by all conjugates of s. Hence S is a quotient group of a (possibly infinite) Coxeter group.

Definition 11.5. Every $g \in G$ can be written in the form $g = x_{i_1} \ldots x_{i_k}$. If k is as small as possible, this expression is called *reduced* (in contrast to free groups, reduced expressions are not necessarily unique). Moreover, let $l(g) := k$ be the *length* of g.

Lemma 11.6. *For $g, h \in G$ and $1 \le i \le n$ we have*

(i) $l(gh) \le l(g) + l(h)$.

(ii) $l(g^{-1}) = l(g)$.

(iii) $l(gx_i) = l(g) \pm 1$.

Proof. The first two statements are trivial. They show

$$l(g) - 1 = l(gx_ix_i) - l(x_i) \leq l(gx_i) \leq l(g) + l(x_i) = l(g) + 1.$$

The homomorphism $f\colon G \to \{\pm 1\}$ from Remark 11.3 yields

$$(-1)^{l(gx_i)} = f(gx_i) = -f(g) = -(-1)^{l(g)}$$

and $l(gx_i) \neq l(g)$. □

Remark 11.7. The kernel of the homomorphism $f\colon G \to \{\pm 1\}$ is the set of elements of even length. It is called the *alternating subgroup*. In S_n we have $\operatorname{sgn}(g) = (-1)^{l(g)}$ for $g \in S_n$.

Definition 11.8. Let V be an $\mathbb{R}$-vector space with basis $b_1, \ldots, b_n$.

- We define a symmetric bilinear form on V by

$$[b_i, b_j]_G := [b_i, b_j] = -\cos\frac{\pi}{m_{ij}},$$

  where $-\cos\frac{\pi}{\infty} = -1$ is set.

- For $i = 1, \ldots, n$ let

$$\sigma_i\colon V \to V, \qquad v \mapsto v - 2[v, b_i]b_i.$$

Theorem 11.9. *There exists exactly one homomorphism $\sigma\colon G \to \mathrm{GL}(V)$ with $\sigma(x_i) = \sigma_i$ for $i = 1, \ldots, n$. Moreover, $\sigma_i\sigma_j$ has order m_{ij} for $1 \leq i, j \leq n$. With the notation ${}^g v := \sigma(g)(v)$ we have $[v, w]_G = [{}^g v, {}^g w]_G$.*

Proof. For $i = 1, \ldots, n$ we have $[b_i, b_i] = 1$. Hence $V_i := b_i^\perp := \operatorname{Ker}(v \mapsto [v, b_i]) \leq V$ is a hyperplane and $V = V_i \oplus \mathbb{R}b_i$. We have $\sigma_i(b_i) = -b_i$ and $\sigma_i(v) = v$ for $v \in V_i$. Hence σ_i is a "reflection" in V_i (in contrast to Euclidean space, $[.,.]_G$ is not necessarily positive definite). In particular, σ_i has order 2. Now let $i < j$ and $W := \langle b_i, b_j \rangle$. The definition of σ_i shows that $\langle \sigma_i, \sigma_j \rangle$ acts on W.

First let $m_{ij} = \infty$. Then we have

$$(\sigma_i\sigma_j)^k(b_i) = (\sigma_i\sigma_j)^{k-1}(\sigma_i(b_i + 2b_j)) = (\sigma_i\sigma_j)^{k-1}(3b_i + 2b_j) = \ldots = (2k+1)b_i + 2kb_j$$

for $k \geq 1$. In particular, $\sigma_i\sigma_j$ has infinite order.

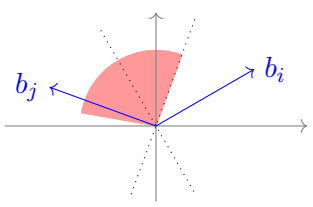
Now let $m_{ij} < \infty$ and $\varphi := \pi/m_{ij}$. For $v = \lambda b_i + \mu b_j \in W$ we have

$$[v, v] = \lambda^2 - 2\lambda\mu\cos\varphi + \mu^2 = (\lambda - \mu\cos\varphi)^2 + \mu^2\sin(\varphi)^2 > 0.$$

Hence $[.,.]_G$ is positive definite on W and therefore coincides with the standard inner product up to the choice of basis. Moreover,

$$[b_i, b_j] = -\cos\varphi = \cos(\pi - \varphi),$$

i. e., the "angle" between b_i and b_j is $\pi - \varphi$. Hence $(\sigma_i\sigma_j)_{|W}$ is a "rotation" by 2φ (since the map is linear, it suffices to consider the images of the two linearly independent reflection axes):



In particular, $(\sigma_i\sigma_j)_{|W}$ has order m_{ij}. Since $[.,.]_G$ is positive definite on W, we have $V = W \oplus W^\perp = W \oplus (V_i \cap V_j)$. Since $\sigma_i\sigma_j$ acts trivially on $V_i \cap V_j$, $\sigma_i\sigma_j$ also has order m_{ij} on V. By von Dyck's theorem, σ is now a homomorphism. For the last statement we compute

$$\begin{aligned}[\sigma_i(v), \sigma_i(w)] &= [v - 2[v, b_i]b_i, w - 2[w, b_i]b_i] \\ &= [v, w] - 2[v, b_i][b_i, w] - 2[w, b_i][v, b_i] + 4[v, b_i][w, b_i] = [v, w].\end{aligned}$$ □

Definition 11.10.

- We call
$$\Phi := \{{}^g b_i : 1 \leq i \leq n, g \in G\} \subseteq V$$
the *root system* of G and its elements are called *roots*.
- A root v can be written uniquely in the form $v = \sum_{i=1}^n v_i b_i$. We call v *positive* (or *negative*) if $v_1, \ldots, v_n \geq 0$ (or $v_1, \ldots, v_n \leq 0$). In this case, one writes $v > 0$ (or $v < 0$; the case $v = 0$ is excluded). Let Π be the set of positive roots.
- For $v \in \Phi$ let
$$\sigma_v \colon V \to V, \qquad w \mapsto w - 2[w, v]v.$$

Remark 11.11.

(i) By Theorem 11.9, all roots of G are unit vectors with respect to $[.,.]_G$. In particular, $\sigma_v(v) = -v$, i. e., σ_v is the reflection in the hyperplane $v^\perp$.

(ii) Because of $\sigma_i(b_i) = -b_i$ we have $-\Phi = \Phi$ and $\sigma_v = \sigma_{-v}$. Let $g \in G$ and $1 \leq i \leq n$ with $v = g(b_i)$. Then
$${}^{gx_ig^{-1}}w = {}^g({}^{g^{-1}}w - 2[{}^{g^{-1}}w, b_i]b_i) = w - 2[w, {}^g b_i]{}^g b_i = w - 2[w, v]v.$$
This shows $\sigma(gx_ig^{-1}) = \sigma_v$. We will also call $x_v := gx_ig^{-1}$ a reflection.

Example 11.12.

(i) Let $G = \langle x, y\rangle \cong D_{2m}$, $\varphi = \pi - \frac{\pi}{m}$, $b_1 = (1, 0)$ and $b_2 = (\cos\varphi, \sin\varphi)$. Then $[.,.]_G$ is the standard inner product on $\mathbb{R}^2$ with respect to b_1, b_2 and $\sigma \colon G \to \mathrm{GL}(\mathbb{R}^2)$ is the well-known representation as the symmetry group of the regular m-gon. The $2m$ roots correspond to the m reflection axes. The positive roots lie geometrically "between" b_1 and b_2. In particular, $\Phi = \Pi \cup (-\Pi)$.

(ii) Let $G = \langle x_1, \ldots, x_{n-1}\rangle \cong S_n$. Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$ and $b_i := \frac{1}{\sqrt{2}}(e_i - e_{i+1})$ for $i = 1, \ldots, n-1$. One easily checks that $[.,.]_G$ is the standard inner product on $V := \langle b_1, \ldots, b_{n-1}\rangle$ with respect to $b_1, \ldots, b_{n-1}$. The homomorphism $\sigma \colon G \to \mathrm{GL}(V)$ arises from the permutation action of S_n on the n coordinates, since

$$\begin{aligned}{}^{x_i}b_i = \sigma_i(b_i) = -b_i = \frac{1}{\sqrt{2}}(e_{i+1} - e_i) = \frac{1}{\sqrt{2}}(e_{x_i(i)} - e_{x_i(i+1)}),\\ {}^{x_i}b_{i+1} = b_{i+1} - 2[b_{i+1}, b_i]b_i = b_{i+1} + b_i = \frac{1}{\sqrt{2}}(e_i - e_{i+2}) = \frac{1}{\sqrt{2}}(e_{x_i(i+1)} - e_{x_i(i+2)}),\\ {}^{x_i}b_j = b_j = \frac{1}{\sqrt{2}}(e_{x_i(j)} - e_{x_i(j+1)}) \qquad (|j - i| > 1).\end{aligned}$$

Clearly, $\Phi = \{\frac{1}{\sqrt{2}}(e_i - e_j) : i \neq j\}$ and $\Pi = \{\frac{1}{\sqrt{2}}(e_i - e_j) : i < j\}$, since $\frac{1}{\sqrt{2}}(e_i - e_j) = b_i + b_{i+1} + \ldots + b_{j-1}$. In particular, $|\Phi| = n(n-1)$ and $|\Pi| = n(n-1)/2$.

Lemma 11.13. *For $g \in G$ and $1 \leq i \leq n$ we have ${}^g b_i > 0$ if $l(gx_i) > l(g)$ and ${}^g b_i < 0$ if $l(gx_i) < l(g)$. In particular, every root is positive or negative, i. e., $\Phi = \Pi \cup (-\Pi)$.*

Proof. Since $g = gx_i x_i$ and ${}^{gx_i} b_i = -{}^g b_i$, it suffices to consider the case $l(gx_i) > l(g)$. We prove the first statement by induction on $l(g)$. In the case $l(g) = 0$ we have $g = 1$, $l(x_i) = 1$ and $b_i > 0$. Now let $g \neq 1$ and $1 \leq j \leq n$ with $l(gx_j) = l(g) - 1$ (the last factor of a reduced expression of g). By assumption, $x_j \neq x_i$. We consider the dihedral group $H := \langle x_i, x_j \rangle \leq G$. Let $l_H \colon H \to \mathbb{N}_0$ be the length function with respect to the generators x_i, x_j of H.[25] Let

$$A := \{y \in gH : l(y) + l_H(y^{-1}g) = l(g)\}.$$

Since $g \in A$, we have $A \neq \varnothing$. Choose $y \in A$ with $l(y)$ minimal. Since $l(gx_j) + l_H(x_j) = l(g) - 1 + 1 = l(g)$, we have $gx_j \in A$. The choice of y shows $l(y) \leq l(gx_j) < l(g)$.

Suppose that $l(yx_i) = l(y) - 1$. Then

$$l(g) \leq l(yx_i) + l_H(x_i y^{-1} g) \leq l(y) - 1 + l_H(y^{-1}g) + 1 = l(y) + l_H(y^{-1}g) = l(g)$$

and $l(g) = l(yx_i) + l_H(x_i y^{-1} g)$. This shows $yx_i \in A$, contradicting $l(yx_i) < l(y)$. Hence $l(yx_i) > l(y)$ and completely analogously $l(yx_j) > l(y)$. Induction shows ${}^y b_i, {}^y b_j > 0$. Let $h := y^{-1} g \in H$. Since $g = yh$, it suffices to show that ${}^h b_i$ is a nonnegative linear combination of b_i and b_j.

Suppose that $l_H(hx_i) < l_H(h)$. Then we would have

$$l(gx_i) = l(yhx_i) \leq l(y) + l_H(hx_i) < l(y) + l_H(h) = l(g).$$

Hence every reduced expression of h with respect to x_i, x_j must end in x_j. In the case $m_{ij} = \infty$ we have ${}^{x_j} b_i = b_i + 2b_j$, ${}^{x_i x_j} b_i = {}^{x_i}(b_i + 2b_j) = 3b_i + 2b_j$ and so on (cf. proof of Theorem 11.9). Hence we may assume $m := m_{ij} < \infty$. Then b_i and b_j form the angle $\pi - \frac{\pi}{m}$ and $x_i x_j$ is a "rotation" by $\frac{2\pi}{m}$. If $l_H(h) = m$, then $h = (x_i x_j)^{m/2}$ if m is even and $h = x_j (x_i x_j)^{(m-1)/2}$ if m is odd. In both cases there would be a reduced expression ending with x_i (namely $h = (x_j x_i)^{m/2}$ or $x_i (x_j x_i)^{(m-1)/2}$, respectively). Hence $l_H(h) < m$ and $h \in \{(x_i x_j)^k, x_j (x_i x_j)^k\}$ with $k < m/2$. In the case $k = (m-1)/2$ we have $h = (x_i x_j)^k$ and ${}^h b_i = b_j$. In all other cases, $(x_i x_j)^k$ is a rotation by less than $\frac{(m-1)\pi}{m}$, i. e., ${}^{(x_i x_j)^k} b_i$ lies strictly "between" b_i and b_j. Another application of x_j does not change this. Hence ${}^h b_i$ is a nonnegative linear combination of b_i and b_j. □

Theorem 11.14. *The homomorphism $\sigma \colon G \to \mathrm{GL}(V)$ from Theorem 11.9 is injective.*

Proof. Let $g \in \mathrm{Ker}(\sigma) \setminus \{1\}$. Then there exists $1 \leq i \leq n$ with $l(gx_i) < l(g)$. Lemma 11.13 yields the contradiction $b_i = {}^g b_i < 0$. □

Definition 11.15. For $I \subseteq \{1, \ldots, n\}$ one calls $G_I := \langle x_i : i \in I \rangle \leq G$ a *parabolic* subgroup of G.

Theorem 11.16. *For $I \subseteq \{1, \ldots, n\}$ we have*

$$G_I \cong \langle \{y_i : i \in I\} \mid \{(y_i y_j)^{m_{ij}} : i, j \in I\} \rangle,$$

i. e., G_I is itself a Coxeter group.

[25] By Lemma 11.17, l_H is the restriction of l to H. This is not used here, however.

Proof. Applying the construction of σ to the Coxeter group on the right-hand side yields exactly $\sigma(G_I) \cong G_I$. □

Lemma 11.17. *Let $g = x_{i_1} \dots x_{i_k} \in G_I$ be reduced in G. Then $i_1, \dots, i_k \in I$. In particular, $\{x_1, \dots, x_n\} \cap G_I = \{x_i : i \in I\}$.*

Proof. Induction on k. W. l. o. g. let $k \geq 1$. By Lemma 11.13 we have ${}^g b_{i_k} < 0$. Moreover, let $g = x_{j_1} \dots x_{j_l}$ with $j_1, \dots, j_l \in I$. By the definition of the σ_i we have

$$ {}^g b_{i_k} = b_{i_k} + \sum_{a=1}^{l} \lambda_a b_{j_a} $$

with $\lambda_a \in \mathbb{R}$. Since ${}^g b_{i_k} < 0$, we must have $i_k = j_s \in I$ for some $1 \leq s \leq l$. In particular, $h := x_{i_1} \dots x_{i_{k-1}} = g x_{i_k} \in G_I$ is reduced. The claim now follows by induction. □

Corollary 11.18. *If $g = x_{i_1} \dots x_{i_k} = x_{j_1} \dots x_{j_k}$ are two reduced expressions of $g \in G$, then $\{i_1, \dots, i_k\} = \{j_1, \dots, j_k\}$.*

Example 11.19. In the situation of Corollary 11.18, $\{i_1, \dots, i_k\}$ and $\{j_1, \dots, j_k\}$ need not coincide as multisets. For example, in S_3 we have the reduced expressions $x_1 x_2 x_1 = (1,3) = x_2 x_1 x_2$.

Theorem 11.20. *We consider S_n as a Coxeter group with generators $x_i := (i, i+1)$. For $\sigma \in S_n$ there exist uniquely determined numbers $l \geq 0$ and $1 \leq a_1, \dots, a_l \leq n-1$ with the following properties:*

(i) $\sigma = x_{a_1} \dots x_{a_l}$.

(ii) $a_i \neq a_{i-1} \leq a_i + 1$ *for* $i = 2, \dots, l$.

(iii) The sequence $(a_1, \dots, a_l)$ has no segment of the form $(a, a-1, \dots, a-r, a)$ with $r \geq 1$.

In this case, $\sigma = x_{a_1} \dots x_{a_l}$ is a reduced word, i. e., $l = l(\sigma)$.

Proof. We apply the following algorithm to an arbitrary expression $\sigma = x_{a_1} \dots x_{a_k}$:

(1) If $a_i = a_{i+1}$ for some i, delete $x_{a_i} x_{a_{i+1}} = 1$ from the expression.

(2) If $a_{i-1} > a_i + 1$, swap $x_{a_{i-1}}$ and x_{a_i} and start again at (1). This is allowed since $x_{a_{i-1}}$ and x_{a_i} are disjoint.

(3) If $(a_1, \dots, a_k)$ has a segment of the form $(a, a-1, \dots, a-r, a)$, replace it by $(a-1, a, a-1, \dots, a-r)$ and start again at (1). This does not change σ, since

$$ x_a x_{a-1} \dots x_{a-r} x_a = x_a x_{a-1} x_a x_{a-2} x_{a-3} \dots x_{a-r} = x_{a-1} x_a x_{a-1} x_{a-2} \dots x_{a-r} $$

by the braid relation.

By (1) and (3), $\sum_{i=1}^{k} a_i$ is decreased, while (2) decreases the lexicographic order of the sequence $(a_1, \ldots, a_k)$. Hence the algorithm must terminate after finitely many steps. In the end, the conditions (i)–(iii) are satisfied.

We call a sequence $(a_1, \ldots, a_l)$ that satisfies (ii) and (iii) *regular*. For $n = 2$ there are only the regular sequences $a = ()$ (with $l = 0$) and $a = (1)$. Inductively, we assume that there are exactly $(n-1)!$ regular sequences $(a_1, \ldots, a_l)$ with $1 \leq a_1, \ldots, a_l \leq n-2$. Now let $(a_1, \ldots, a_l)$ with $a_k = n-1$. Then $(a_k, a_{k+1}, \ldots, a_l) = (n-1, n-2, \ldots, n-r)$ for some $r \geq 1$ by (ii) and (iii). For $(a_1, \ldots, a_{k-1})$ there are inductively exactly $(n-1)!$ possibilities, while there are exactly $n-1$ possibilities for r. Hence there exist $(n-1)!(n-1)$ regular sequences containing $n-1$. Together with the $(n-1)!$ sequences not containing $n-1$, we obtain exactly $n! = |S_n|$ regular sequences. Hence every permutation can be represented by only one regular sequence.

If one applies the above algorithm to a reduced word, the length l cannot decrease. Hence $l = l(\sigma)$. □

Example 11.21. In S_5 we have

$$\begin{aligned} x_4x_3(x_2x_1x_2)x_3x_4 &= x_4x_3(x_1x_2x_1)x_3x_4 = x_1x_4(x_3x_2x_1x_3)x_4 \\ &= x_1x_4(x_2x_3x_2x_1)x_4 = x_1x_2(x_4x_3x_2x_1x_4) = x_1x_2x_3x_4x_3x_2x_1. \end{aligned}$$

A reduced expression can be found somewhat faster by proceeding recursively. Let $\sigma \in S_n$ and $a := \sigma^{-1}(n)$. Then $\tau := \sigma x_a x_{a+1} \ldots x_{n-1} \in S_{n-1}$. By induction, τ has an expression of the desired form. Hence $\sigma = \tau x_{n-1} x_{n-2} \ldots x_a$ also has this property.

Theorem 11.22. *The map $I \mapsto G_I$ is an isomorphism of lattices, i. e., we have*

(i) $I \subseteq J \iff G_I \leq G_J$.

(ii) $G_{I \cup J} = \langle G_I, G_J \rangle$.

(iii) $G_{I \cap J} = G_I \cap G_J$.

Proof.

(i) Clearly, $I \subseteq J$ implies $G_I \leq G_J$. If $G_I \leq G_J$, then

$$\{x_i : i \in I\} = \{x_1, \ldots, x_n\} \cap G_I \subseteq \{x_1, \ldots, x_n\} \cap G_J = \{x_j : j \in J\}$$

by Lemma 11.17.

(ii) Trivial.

(iii) Clearly, $G_{I \cap J} \leq G_I \cap G_J$. The reverse inclusion follows from Lemma 11.17. □

Theorem 11.23. *For $g \in G$, $l(g)$ is the number of positive roots that are mapped to negative roots under g, i. e., $l(g) = |\Pi \cap g^{-1}(-\Pi)| \leq |\Pi|$. In particular, this number is finite.*

Proof. Induction on $l(g)$: W. l. o. g. we may assume $g \neq 1$. First let $g = x_i$ for some $1 \leq i \leq n$. Since ${}^g b_i = -b_i$, we have to show ${}^g v > 0$ for all $v \in \Pi \setminus \{b_i\}$. Since all roots are normalized, v is not a multiple of b_i. So let $v = \sum_{j=1}^{n} \lambda_j b_j$ with $\lambda_k > 0$ for at least one $k \neq i$. In ${}^g v = \sigma_i(v) = v - 2[v, b_i]b_i$, the coefficient of b_k is still the positive number λ_k. Hence ${}^g v > 0$.

Now let $l(g) \geq 2$ and $1 \leq i \leq n$ with $l(gx_i) = l(g)-1$. By Lemma 11.13, ${}^g b_i < 0$, i. e., $b_i \in \Pi \cap g^{-1}(-\Pi)$. With what we have just proved, it follows that

$$\Pi \cap (gx_i)^{-1}(-\Pi) = x_i(x_i(\Pi) \cap g^{-1}(-\Pi)) = x_i(\Pi \cap g^{-1}(-\Pi) \setminus \{b_i\}).$$

Induction shows

$$|\Pi \cap g^{-1}(-\Pi)| = |\Pi \cap (gx_i)^{-1}(-\Pi)| + 1 = l(gx_i) + 1 = l(g). \qquad \square$$

Remark 11.24. If the length in G is bounded, say by m, then $|G| \leq \sum_{k=0}^{m} n^k < \infty$. Hence in infinite Coxeter groups there must be infinitely many (positive) roots. In particular, $-\operatorname{id}_V \notin \sigma(G)$ (otherwise $-\operatorname{id}$ would have infinite length).

Corollary 11.25. *If G is finite, then there is exactly one element $g \in G$ of maximal length.*

Proof. Suppose that $g, h \in G$ have maximal length $l(g) = l(h)$. Then $l(gx_i) < l(g)$ and ${}^g b_i < 0$ for $i = 1, \ldots, n$. Since every positive root is a nonnegative linear combination of the b_i, g (and h) must map all positive roots to negative roots, i. e., $g(\Pi) = -\Pi$. Hence $g^2(\Pi) = \Pi = gh(\Pi)$ and $l(g^2) = 0 = l(gh)$. This shows $h = g^{-1} = g$. $\square$

Example 11.26.

(i) For $G = \langle x, y \rangle \cong D_{2m}$, the element $z := xy\ldots = yx\ldots$ (with m factors each) is the element of maximal length (cf. Example 11.12). If m is even, then $\sigma(z) = -\operatorname{id}$, i. e., the rotation by π.

(ii) For $g \in G = S_n$ and $i < j$ we have

$${}^g\big(\frac{1}{\sqrt{2}}(e_i - e_j)\big) < 0 \iff \frac{1}{\sqrt{2}}(e_{g(i)} - e_{g(j)}) < 0 \iff g(i) > g(j)$$

(Example 11.12). Hence $l(g)$ is the number of *inversions* of g, i. e., pairs $i < j$ with ${}^g i > {}^g j$. Therefore, the element of maximal length is

$$g = \begin{pmatrix} 1 & 2 & \cdots & n \\ n & n-1 & \cdots & 1 \end{pmatrix} = (1,n)(2,n-1)\ldots = t_1 \ldots t_{n-1}$$

with $t_i = (i+1, i)\ldots(2,1)$ and $l(g) = 1 + \ldots + n - 1 = n(n-1)/2 = |\Pi|$.

Lemma 11.27 (Tits)**.** *Let $g \in G$ and $v \in \Pi$. Then $l(gx_v) > l(g) \iff {}^g v > 0$.*

Proof. As in Lemma 11.13, it suffices to show $l(gx_v) > l(g) \Rightarrow {}^g v > 0$. Induction on $l(g)$. The case $g = 1$ is clear. Now let $l(g) > 0$ and $l(x_i g) < l(g)$. Since

$$l(x_i g x_v) \geq l(gx_v) - 1 > l(g) - 1 = l(x_i g)$$

we have ${}^{x_i g} v > 0$ by the induction hypothesis. Suppose that ${}^g v < 0$. By Theorem 11.23, we then have ${}^g v = -b_i$. This shows ${}^{g^{-1} x_i g} v = {}^{g^{-1}} b_i = -v$ and $g^{-1} x_i g = x_v$ by Remark 11.11. But this contradicts $l(gx_v) > l(g) > l(x_i g) = l(gx_v)$. $\square$

Remark 11.28. In the following, we use the notation $x_1 \ldots \check{x}_i \ldots x_k := x_1 \ldots x_{i-1} x_{i+1} \ldots x_k$.

Theorem 11.29.

(i) (exchange condition) *Let $g = x_{i_1} \dots x_{i_k} \in G$ and $v \in \Phi$ with $l(gx_v) < l(g)$. Then there exists $1 \leq s \leq k$ with $gx_v = x_{i_1} \dots \check{x}_{i_s} \dots x_{i_k}$. If $l(g) = k$, then s is uniquely determined.*

(ii) (deletion condition) *Let $g = x_{i_1} \dots x_{i_k} \in G$ with $l(g) < k$. Then there exist $1 \leq s < t \leq k$ with $g = x_{i_1} \dots \check{x}_{i_s} \dots \check{x}_{i_t} \dots x_{i_k}$.*

Proof.

(i) Since $x_v = x_{-v}$, we may assume $v > 0$. Lemma 11.27 implies ${}^g v < 0$. Since $v > 0$, there exists an s with ${}^{x_{i_{s+1}} \dots x_{i_k}} v > 0$ and ${}^{x_{i_s} \dots x_{i_k}} v < 0$. Theorem 11.23 implies ${}^{x_{i_{s+1}} \dots x_{i_k}} v = b_{i_s}$. This shows $(x_{i_{s+1}} \dots x_{i_k}) x_v (x_{i_{s+1}} \dots x_{i_k})^{-1} = x_{i_s}$, hence $gx_v = x_{i_1} \dots \check{x}_{i_s} \dots x_{i_k}$. Now let $l(g) = k$. Suppose that there exist $s < t$ with $x_{i_1} \dots \check{x}_{i_s} \dots x_{i_k} = gx_v = x_{i_1} \dots \check{x}_{i_t} \dots x_{i_k}$. This yields $x_{i_{s+1}} \dots x_{i_t} = x_{i_s} \dots x_{i_{t-1}}$ and $x_{i_s} \dots x_{i_t} = x_{i_{s+1}} \dots x_{i_{t-1}}$. But then we would have $g = x_{i_1} \dots \check{x}_{i_s} \dots \check{x}_{i_t} \dots x_{i_k}$ and $l(g) < k$.

(ii) Since $l(g) < k$, there exists a t with $l(x_{i_1} \dots x_{i_t}) < l(x_{i_1} \dots x_{i_{t-1}})$. By (i) we obtain $x_{i_1} \dots x_{i_t} = x_{i_1} \dots \check{x}_{i_s} \dots x_{i_{t-1}}$ for some $s < t$. □

Remark 11.30.

(i) We have

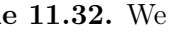
$$l(x_v g) < l(g) \Longrightarrow l(g^{-1} x_v) < l(g^{-1}) \Longrightarrow g^{-1} x_v = x_{i_k} \dots \check{x}_{i_s} \dots x_{i_1} \Longrightarrow x_v g = x_{i_1} \dots \check{x}_{i_s} \dots x_{i_k}.$$

(ii) One can show that every group that satisfies the exchange condition (or deletion condition) with respect to a generating set of involutions is a Coxeter group.

(iii) The deletion condition shows that one obtains a reduced expression from an arbitrary expression $g = x_{i_1} \dots x_{i_k}$ by suitable deletion.

Definition 11.31.

- We call G *irreducible*, if there is no partition $\{1, \dots, n\} = I \dot\cup J$ with $G = G_I \times G_J$.
- The *Coxeter graph* $C(G)$ consists of the vertices $e_1, \dots, e_n$ and the edges (e_i, e_j) with $m_{ij} \geq 3$. In the case $m_{ij} > 3$, the edges are labeled with m_{ij}. Clearly, G is uniquely determined by $C(G)$ up to the order of the x_i.

Example 11.32. We have $C(S_n)$:  and $C(D_{2m})$:  by Example 11.4.

Theorem 11.33. *G is irreducible if and only if $C(G)$ is connected.*

Proof. If $G = G_I \times G_J$, then $m_{ij} = 2$ for $i \in I$ and $j \in J$. Hence there is no path between e_i and e_j in $C(G)$. Conversely, if $C(G)$ is disconnected, then there exists a partition $\{1, \dots, n\} = I \cup J$ with $[x_i, x_j] = 1$ (commutator) for all $i \in I$ and $j \in J$. In particular, $[G_I, G_J] = 1$. By Theorem 11.22, $G = G_{I \cup J} = \langle G_I, G_J \rangle$ and $G_I \cap G_J = G_{I \cap J} = 1$. This shows $G = G_I \times G_J$. □

Remark 11.34. Note: The isomorphism type of G does not determine whether G is irreducible as a Coxeter group. For example, $G = D_{12}$ is irreducible with two generators, but also reducible with three generators $G \cong D_6 \times S_2$.

Lemma 11.35. *Let G be irreducible and $V_0 := \{v \in V : [v, V]_G = 0\} \leq V$. Then G acts trivially on V_0 and every proper G-invariant subspace of V lies in V_0.*

Proof. For $v \in V_0$ we have ${}^{x_i}v = \sigma_i(v) = v - 2[v, b_i]_G b_i = v$. Since G is generated by $x_1, \ldots, x_n$, G acts trivially on V_0.

Now let $W < V$ be G-invariant. Suppose that there exists $v \in W \setminus V_0$. Then there exists an i with $[v, b_i]_G \neq 0$. It follows that $b_i = \frac{v - \sigma_i(v)}{2[v, b_i]} \in W$. For b_j with $m_{ij} \geq 3$ we have $[b_i, b_j]_G \neq 0$ and $b_j = \frac{b_i - \sigma_j(b_i)}{2[b_i, b_j]} \in W$. Since $C(G)$ is connected, we obtain $b_1, \ldots, b_n \in W$, contradicting $W < V$. □

Theorem 11.36. *Let G be a finite irreducible Coxeter group and $z \in G$ of maximal length. Then $\mathrm{Z}(G) \leq \langle z \rangle$. In particular, $|\mathrm{Z}(G)| \leq 2$.*

Proof. Let $g \in \mathrm{Z}(G) \setminus \{1\}$. From ${}^{x_i}({}^g b_i) = {}^{gx_i} b_i = -{}^g b_i$ we get ${}^g b_i = \pm b_i$ (Theorem 11.23). Hence $b_i \in E_1(\sigma(g)) \cup E_{-1}(\sigma(g))$ (eigenspaces for the eigenvalues 1 and -1, respectively). Since $g \in \mathrm{Z}(G)$, $E_1(\sigma(g))$ and $E_{-1}(\sigma(g))$ are G-invariant and $E_1(\sigma(g)) < V$, since $g \neq 1$. In the case $E_{-1}(\sigma(g)) < V$ we would have $b_i \in V_0$ by Lemma 11.35. However, $[b_i, b_i]_G = 1$. This shows $\sigma(g) = -\mathrm{id}$ and ${}^g v < 0$ for all $v \in \Pi$. Theorem 11.23 implies $g = z$. □

Lemma 11.37. *Let $H \leq \mathrm{GL}(n, \mathbb{R})$ be finite and irreducible as a matrix group. Then:*

(i) There exists an H-invariant positive definite bilinear form on $\mathbb{R}^n$.

(ii) Suppose that there exist $h \in H$ and $\lambda \in \mathbb{R}$ such that the eigenspace $E_\lambda(h)$ has odd dimension. Then $\mathrm{C}_{\mathrm{GL}(n,\mathbb{R})}(H) = \mathbb{R}^\times 1_n$.[26] In particular, this holds if n is odd.

(iii) If $\mathrm{C}_{\mathrm{GL}(n,\mathbb{R})}(H) = \mathbb{R}^\times 1_n$, then any two nondegenerate H-invariant bilinear forms on $\mathbb{R}^n$ differ only by a constant.

Proof.

(i) For $v, w \in \mathbb{R}^n$, the formula

$$[v, w]_H := \sum_{h \in H} [hv, hw]$$

defines an H-invariant positive definite bilinear form, where $[.,.]$ is the standard inner product.

(ii) Let $f \in \mathrm{C}_{\mathrm{GL}(n,\mathbb{R})}(H)$. Then f acts on $E_\lambda(h)$. Since $\dim E_\lambda(h)$ is odd, f has a real eigenvalue μ on $E_\lambda(h)$. Now $0 \neq E_\mu(f) \leq \mathbb{R}^n$ is H-invariant, and it follows that $E_\mu(f) = \mathbb{R}^n$ and $f = \mu 1_n$, since H is irreducible. The second statement follows with $h = 1$.

(iii) Let $[.,.]_1$ and $[.,.]_2$ be two nondegenerate H-invariant bilinear forms on $V := \mathbb{R}^n$. Let $V^* := \mathrm{Hom}(V, \mathbb{R})$ be the dual space of V. Then $\varphi_i \colon V \to V^*$, $v \mapsto [v, .]_i$ for $i = 1, 2$ are isomorphisms of vector spaces (note that $\dim V = \dim V^*$). Hence $f := \varphi_2^{-1} \circ \varphi_1 \colon V \to V$ is an isomorphism with $[v, w]_1 = [f(v), w]_2$ for all $v, w \in V$. For $h \in H$ it follows that

$$[f(hv), hw]_2 = [hv, hw]_1 = [v, w]_1 = [f(v), w]_2 = [hf(v), hw]_2.$$

This shows $fh = hf$, i. e., $f \in \mathrm{C}_{\mathrm{GL}(n,\mathbb{R})}(H) = \mathbb{R}^\times 1_n$. □

[26] One says that H is *absolutely* irreducible.

Remark 11.38. We consider the dual space $V^* := \mathrm{Hom}(V, \mathbb{R})$ with the dual basis $\beta_1, \ldots, \beta_n$, where $\beta_i(b_j) = \delta_{ij}$. Via ${}^g\varphi(v) := \varphi({}^{g^{-1}}v)$ for $v \in V$, $\varphi \in V^*$ and $g \in G$, G acts on V^*. The map $\Gamma\colon V \to V^*$, $v \mapsto [v, .]_G$ induces an isomorphism between the actions on V and V^*, since

$$({}^g\Gamma(v))(w) = \Gamma(v)({}^{g^{-1}}w) = [v, {}^{g^{-1}}w]_G = [{}^gv, w]_G = \Gamma({}^gv)(w)$$

for $v, w \in V$ and $g \in G$. Let $\sigma^*\colon G \to \mathrm{GL}(V^*)$ be the corresponding monomorphism. We define

$$C := \{\varphi \in V^* : \forall i : \varphi(b_i) > 0\} \subseteq V^*.$$

With respect to the dual basis, $C = \mathbb{R}^n_{>0}$ is an open set in Euclidean space $\mathbb{R}^n$. Since the determinant $\mathbb{R}^{n\times n} \to \mathbb{R}$ is continuous, $\mathrm{GL}(V^*) = \det^{-1}(\mathbb{R} \setminus \{0\})$ is open in $\mathbb{R}^{n\times n}$.

Theorem 11.39. *The image $\sigma^*(G)$ is a discrete subgroup of $\mathrm{GL}(V^*)$, i. e., for all $a \in \sigma^*(G)$ there exists an open neighborhood $U(a) \subseteq \mathrm{GL}(V^*)$ with $U(a) \cap \sigma^*(G) = \{a\}$. In particular, $\sigma^*(G)$ is closed.*

Proof. We identify G with $\sigma^*(G)$. Let $c \in C$ and $F\colon \mathrm{GL}(V^*) \to V^*$, $a \mapsto {}^ac = a(c)$. Since matrix-vector multiplication is continuous, F is continuous and $D := F^{-1}(C) \subseteq \mathrm{GL}(V^*)$ is an open neighborhood of $1 \in \mathrm{GL}(V^*)$. For $g \in G \setminus \{1\}$ there exists x_i with $l(g^{-1}x_i) < l(g)$. Tits' lemma implies ${}^{g^{-1}}b_i < 0$ and ${}^gc(b_i) = c({}^{g^{-1}}b_i) < 0$. This shows $D \cap G = \{1\}$. For an arbitrary $g \in G$, gD is an open neighborhood of g with $gD \cap G = g(D \cap G) = \{g\}$.

Now let $(a_i) \subseteq G$ be a convergent sequence. Then there exists $k \in \mathbb{N}$ with $a_i \in U(a_k) \cap G = \{a_k\}$ for all $i \geq k$. Hence the sequence becomes constant and the limit lies in G. This shows that G is closed. □

Theorem 11.40. *G is finite if and only if $[., .]_G$ is positive definite on V.*

Proof. Let G be finite. We argue by induction on n. In the case $n = 1$ we have $[b_1, b_1]_G = 1$ and we are done. Let $n > 1$. Suppose that $G = G_I \times G_J$ is reducible. Then $V = V_I \oplus V_J$ with $V_I := \langle b_i : i \in I\rangle$ and analogously V_J. By induction, $[., .]_{G_I}$ and $[., .]_{G_J}$ are positive definite on V_I and V_J, respectively. For $i \in I$ and $j \in J$ we have $[b_i, b_j]_G = -\cos\frac{\pi}{2} = 0$. For $v = v_I + v_J \neq 0$ it follows that

$$[v, v]_G = [v_I, v_I]_{G_I} + [v_J, v_J]_{G_J} > 0.$$

Hence $[., .]_G$ is positive definite on V.

Now let G be irreducible and $V_0 = \{v \in V : [v, V]_G = 0\} \leq V$. By Maschke, V_0 has a G-invariant complement $W \leq V$. Lemma 11.35 shows $W = V$ and $V_0 = 0$, i. e., $[., .]_G$ is nondegenerate. The same argument also shows that $\sigma(G)$ is irreducible as a matrix group. Since $\dim E_{-1}(\sigma(x_1)) = \dim E_{-1}(\sigma_1) = 1$, $[., .]_G$ is uniquely determined up to a constant by Lemma 11.37. But Lemma 11.37 also states that a G-invariant positive definite bilinear form exists. Hence $[., .]_G$ is positive definite.

Conversely, let $[., .]_G$ be positive definite. The isomorphism $\Gamma\colon V \to V^*$, $v \mapsto [v, .]_G$ from Remark 11.38 yields a G-invariant positive definite bilinear form on V^* (namely $[\varphi, \mu] := [\Gamma^{-1}(\varphi), \Gamma^{-1}(\mu)]_G$ for $\varphi, \mu \in V^*$). By Sylvester's law of inertia, V^* has an orthonormal basis Δ with respect to this bilinear form. If we write $g \in \sigma^*(G)$ as a matrix with respect to Δ, then the columns have norm 1. Since all norms on $\mathbb{R}^n$ are equivalent, $\sigma^*(G)$ is bounded (with respect to the Euclidean norm). By Theorem 11.39, $\sigma^*(G)$ is additionally closed and hence compact. For $a \in \sigma^*(G)$ we choose, according to Theorem 11.39, an open neighborhood $U(a) \subseteq \mathbb{R}^{n\times n}$ with $U(a) \cap \sigma^*(G) = \{a\}$. By Heine-Borel, a finite selection of these neighborhoods already covers $\sigma^*(G)$. Hence $\sigma^*(G) \cong G$ must be finite. □

Theorem 11.41. *The finite Coxeter groups are exactly the reflection groups.*

Proof. Let G be a finite Coxeter group. By Theorem 11.9 and Theorem 11.14, $G \cong \sigma(G) \leq \mathrm{GL}(V)$ is a reflection group.

Now conversely let $V := \mathbb{R}^n$ and $S \leq \mathrm{GL}(V)$ be a reflection group. Let $\Phi \subseteq V$ be the set of unit vectors b such that the reflection σ_b in $b^\perp$ lies in S. We show that Φ has the properties of a root system. From $|S| < \infty$ we get $|\Phi| < \infty$. Since $b^\perp = (-b)^\perp$, we have $-\Phi = \Phi$. We order the $b \in \Phi$ lexicographically according to the coefficients with respect to the standard basis of $\mathbb{R}^n$. Let $\Pi := \{b \in \Phi : b > 0\}$ be the set of *positive roots* ($b > 0$ means that the first nonzero component of b is positive). Then $\Phi = \Pi \cup (-\Pi)$. Let $\Delta \subseteq \Phi$ be a minimal subset such that for all $b \in \Pi$ there exist numbers $\lambda_s \geq 0$ with $b = \sum_{s \in \Delta} \lambda_s s$. In the following, let $[.,.]$ be the standard inner product on $\mathbb{R}^n$.

Step 1: $[b, c] \leq 0$ for all distinct $b, c \in \Delta$.
Suppose that $[b, c] > 0$. For $\mu := 2[b, c] > 0$ we have $\sigma_b(c) = c - \mu b$. Since $\sigma_b \sigma_c \sigma_b = \sigma_{\sigma_b(c)} \in S$, we have $\sigma_b(c) \in \Phi$. First let $\sigma_b(c) \in \Pi$. Then there exist $\lambda_s \geq 0$ with $\sigma_b(c) = \sum_{s \in \Delta} \lambda_s s$. In the case $\lambda_c < 1$ we obtain

$$(1 - \lambda_c)c = \sigma_b(c) + \mu b - \lambda_c c = \mu b + \sum_{s \neq c} \lambda_s s.$$

Now one could remove c from Δ, contradicting the minimality of Δ. Hence $\lambda_c \geq 1$ and

$$(\lambda_c - 1)c + \mu b + \sum_{s \neq c} \lambda_s s = 0.$$

Since $\mu > 0$, this contradicts the definition of Π. This shows $-\sigma_b(c) \in \Pi$. With $-\sigma_b(c) = \sum_{s \in \Delta} \lambda_s s$ it follows that

$$(\mu - \lambda_b)b = -\sigma_b(c) + c - \lambda_b b = c + \sum_{s \neq b} \lambda_s s.$$

In the case $\lambda_b < \mu$ one could remove b from Δ. Hence $\lambda_b \geq \mu$. But then we would have

$$(\lambda_b - \mu)b + c + \sum_{s \neq b} \lambda_s s = 0$$

which contradicts the definition of Π. Altogether, $[b, c] \leq 0$ must hold.

Step 2: Δ is linearly independent.
Let $\sum_{s \in \Delta} \lambda_s s = 0$ with $\lambda_s \in \mathbb{R}$ for $s \in \Delta$. Separating the positive and negative coefficients yields $b := \sum_{\lambda_s \geq 0} \lambda_s s = -\sum_{\lambda_t < 0} \lambda_t t$. From Step 1 we obtain

$$0 \leq [b, b] = \sum_{\lambda_s \geq 0} \sum_{\lambda_t < 0} \lambda_s(-\lambda_t)[s, t] \leq 0$$

and $b = 0$. Since b is a nonnegative linear combination of positive roots, we get $\lambda_s = 0$ for all $s \in \Delta$.

Step 3: $S = \langle \sigma_b : b \in \Delta \rangle$.
By Step 2, every root in Φ can be written uniquely as a linear combination of Δ, where either all coefficients are nonnegative or all coefficients are nonpositive. Let $T := \langle \sigma_b : b \in \Delta \rangle \leq S$ and $b \in \Pi$. Among all elements in the orbit ${}^T b$ we choose $c = \sum_{s \in \Delta} \lambda_s s \in \Pi$ such that $h(c) := \sum_{s \in \Delta} \lambda_s$ is as small as possible. Suppose that $c \notin \Delta$. Since $1 = [c, c] = \sum_{s \in \Delta} \lambda_s [c, s]$, there exists a $t \in \Delta$ with $[c, t] > 0$. We have $\sigma_t(c) = c - 2[c, t]t$. Since $c \in \Pi \setminus \Delta$, c is not a multiple of t. By Step 2, the representation of $\sigma_t(c)$ with respect to Δ is unique. Since $\Phi = \Pi \cup (-\Pi)$, there cannot be coefficients with different signs. This shows $\lambda_t \geq 2[c, t]$ and we get the contradiction $h(\sigma_t(c)) < h(c)$ to the choice of c. Hence $c \in \Delta$ and there exists $t \in T$ with $b = t(c)$. As noted before, it follows that $\sigma_b = t\sigma_c t^{-1} \in T$. Hence $S = \langle \sigma_b : b \in \Pi \rangle = T$.

Step 4: S is a Coxeter group.
Let $\Delta = \{b_1, \dots, b_n\}$ and $\sigma_i = \sigma_{b_i}$ for $i = 1, \dots, n$. Let $|\langle \sigma_i \sigma_j \rangle| = m_{ij}$ for $1 \le i, j \le n$. By Step 2, $m_{ij} \ge 2$ for $i \ne j$. In this case, $\sigma_i \sigma_j$ is a rotation by the angle $\varphi := 2\pi/m_{ij} \le \pi$ in the plane $\langle b_i, b_j \rangle$. It follows that

$$\cos\varphi = [b_j, \sigma_i\sigma_j(b_j)] = -[b_j, \sigma_i(b_j)] = -[b_j, b_j - 2[b_j, b_i]b_i] = -1 + 2[b_i, b_j]^2,$$
$$[b_i, b_j]^2 = \frac{\cos\varphi + 1}{2} = \frac{\cos(\varphi/2)^2 - \sin(\varphi/2)^2 + 1}{2} = \cos(\varphi/2)^2.$$

From Step 1 we get $[b_i, b_j] = -\cos(\varphi/2) = -\cos(\pi/m_{ij})$. Let $G = \langle x_1, \dots, x_n \rangle$ be the Coxeter group with parameters m_{ij}. With the notation from Definition 11.8, $[.,.]_G$ is now the standard inner product and the monomorphism σ from Theorem 11.9 maps G to S. By Step 3, σ is surjective and $S \cong G$ is a Coxeter group. □

Remark 11.42.

(i) G is finite if and only if the matrix $(-\cos(\pi/m_{ij}))_{ij}$ is positive definite. For $G = S_n$ we obtain the matrix

$$\begin{pmatrix} 1 & -1/2 & & 0 \\ -1/2 & \ddots & \ddots & \\ & \ddots & 1 & -1/2 \\ 0 & & -1/2 & 1 \end{pmatrix}.$$

(ii) In the following, we call G *positive semidefinite* if $[.,.]_G$ is positive semidefinite. In particular, every finite Coxeter group is positive semidefinite.

Lemma 11.43. *Let $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ be symmetric, positive semidefinite and indecomposable (i. e., for every partition $\{1, \dots, n\} = I \dot\cup J$ there exist $i \in I$, $j \in J$ with $a_{ij} \ne 0$). Moreover, let $a_{ij} \le 0$ for $i \ne j$. Then we have*

(i) $\mathrm{Ker}(A) = \{v \in \mathbb{R}^n : vAv^{\mathrm{t}} = 0\}$ *and* $\dim \mathrm{Ker}(A) \le 1$.

(ii) The eigenspace for the smallest eigenvalue of A is spanned by a positive eigenvector (i. e., all components are positive).

Proof.

(i) Clearly, $\mathrm{Ker}(A)$ lies in $N := \{v \in \mathbb{R}^n : vAv^{\mathrm{t}} = 0\}$. Conversely, let $v \in N$. By the spectral theorem, there exists an orthogonal matrix S with $D := SAS^{\mathrm{t}} = \mathrm{diag}(d_1, \dots, d_n)$ and $d_1, \dots, d_n \ge 0$. For $w := vS^{\mathrm{t}}$ it follows that

$$\sum_{i=1}^{n} w_i^2 d_i = wDw^{\mathrm{t}} = vAv^{\mathrm{t}} = 0.$$

Hence $w_i = 0$ or $d_i = 0$ for all i. This shows $0 = S^{\mathrm{t}}Dw^{\mathrm{t}} = Av^{\mathrm{t}}$ and $N \subseteq \mathrm{Ker}(A)$.

Now let $x \in N \setminus \{0\}$ and $z = (|x_1|, \dots, |x_n|)$. Since $a_{ij} \le 0$ for $i \ne j$, we have

$$0 \le zAz^{\mathrm{t}} = \sum_{i,j=1}^{n} a_{ij}|x_i x_j| \le \sum_{i,j=1}^{n} a_{ij} x_i x_j = xAx^{\mathrm{t}} = 0,$$

so $z \in N$. Let $I := \{1 \le i \le n : z_i = 0\}$ and $J = \{1, \dots, n\} \setminus I$. From $z \in \mathrm{Ker}(A)$ it follows that $\sum_{j \in J} a_{ij} z_j = 0$ for $i = 1, \dots, n$. Since $z_j > 0$, this is only possible if $a_{ij} = 0$ for all $i \in I$ and

$j \in J$. Since A is indecomposable, $I = \varnothing$ must hold, i. e., all components of z are positive and all components of x are nonzero. Since $x \neq 0$ was arbitrary, we conclude $\dim N \leq 1$ (otherwise one could combine two linearly independent vectors into a linear combination with a 0 component).

(ii) Let $d := \min\{d_1, \ldots, d_n\} \geq 0$. Then $B := A - d1_n$ also satisfies the hypotheses of the lemma. The claim now follows from the proof of (i) for B. □

Lemma 11.44. *Let G be irreducible and positive semidefinite. We construct a proper subgraph D from $C(G)$ by removing vertices or edges or by reducing edge labels. Then the Coxeter group of D is finite.*

Proof. We number the vertices of $C(G)$ such that D is formed by the first k vertices. Let $A = (a_{ij}) = ([b_i, b_j]_G) \in \mathbb{R}^{n\times n}$ and $B = (b_{ij}) = (-\cos(\pi/m'_{ij})) \in \mathbb{R}^{k\times k}$ be the corresponding matrices. Then $b_{ij} \geq -\cos(\pi/m_{ij}) = a_{ij}$. Suppose that B is not positive definite. Let $v \in \mathbb{R}^k \setminus \{0\}$ with $vBv^{\mathrm{t}} \leq 0$ and $w = (|v_1|, \ldots, |v_k|, 0, \ldots, 0) \in \mathbb{R}^n$. Since G is positive semidefinite, it follows that

$$0 \leq wAw^{\mathrm{t}} = \sum_{i,j=1}^{k} a_{ij}|v_i||v_j| \leq \sum_{i,j=1}^{k} b_{ij}|v_i||v_j| \leq \sum_{i,j=1}^{k} b_{ij}v_iv_j = vBv^{\mathrm{t}} \leq 0$$

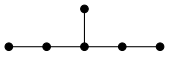
and $wAw^{\mathrm{t}} = 0$. By Lemma 11.43, all components of w are positive, i. e., $k = n$ and $v_i \neq 0$ for $i = 1, \ldots, k$. This yields $a_{ij} = b_{ij}$. But now D is no longer a *proper* subgraph. Contradiction. □

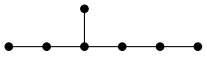
Theorem 11.45 (Coxeter)**.** *Every finite irreducible Coxeter group G belongs to one of the following families:*[27]

(A_n) *with $n \geq 1$.*

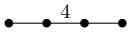
(B_n)  *with $n \geq 2$.*

(D_n)  *with $n \geq 4$.*

(E_6) 

(E_7) 

(E_8) 

(F_4) 

(H_3) 

(H_4) 

$(I_2(m))$  *with $m \geq 5$.*

Proof.

Existence: Since the given graphs are connected, the corresponding groups are irreducible (Theorem 11.33). We show that they are all finite. All graphs are trees. If one removes a suitable leaf, one obtains a tree which is again of one of the given types. We can therefore argue by induction and only need to show that the matrix $M := 2([b_i, b_j])$ has positive determinant (Sylvester's criterion). For $n = 2$ we obtain

$$\det M = 4(1 - \cos(\pi/m)^2) = 4\sin(\pi/m)^2 > 0$$

since $m = m_{12} \geq 3$. Now let $n \geq 3$. We number the vertices $e_1, \ldots, e_n$ such that e_n is a leaf and the corresponding edge $\{e_{n-1}, e_n\}$ has label $m = 3$ or 4. Laplace expansion along the last column shows

$$\det M = 2\det M_{n-1} - 4\cos(\pi/m)^2 \det M_{n-2} = 2\det M_{n-1} - \lambda \det M_{n-2}$$

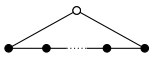
[27] These special graphs are called *Dynkin diagrams*.

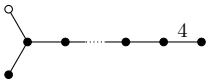
with $\lambda \in \{1, 2\}$, since $\cos(\pi/3) = 1/2$ and $\cos(\pi/4) = 1/\sqrt{2}$. Using $\sin(\pi/5) = \sqrt{10 - 2\sqrt{5}}/4$ one computes inductively

$$
\begin{aligned}
(A_n)&\colon \det M = 2n - (n-1) = n + 1,\\
(B_n)&\colon \det M = 2 \cdot 2 - 2 = 2,\\
(D_n)&\colon \det M = 2 \cdot 4 - 4 = 4,\\
(E_6)&\colon \det M = 2 \cdot 4 - 5 = 3,\\
(E_7)&\colon \det M = 2 \cdot 3 - 4 = 2,\\
(E_8)&\colon \det M = 2 \cdot 2 - 3 = 1,\\
(F_4)&\colon \det M = 2 \cdot 2 - 3 = 1,\\
(H_3)&\colon \det M = 2 \cdot 4\sin(\pi/5)^2 - 2 = \frac{10 - 2\sqrt{5} - 4}{2} = 3 - \sqrt{5} > 0,\\
(H_4)&\colon \det M = 2(3 - \sqrt{5}) - 4\sin(\pi/5)^2 = \frac{12 - 4\sqrt{5} - 5 + \sqrt{5}}{2} = \frac{7 - 3\sqrt{5}}{2} > 0.
\end{aligned}
$$

Auxiliary graphs: We add one more vertex to some of the graphs and obtain the following Coxeter graphs[28] (the number of vertices is now $n + 1$):

$(\tilde{A}_1)$ 

$(\tilde{A}_n)$  with $n \geq 2$.

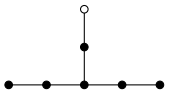
$(\tilde{B}_n)$ 

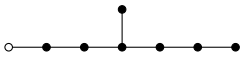
with $n \geq 3$.

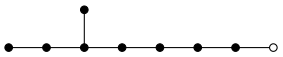
$(\tilde{C}_n)$ 

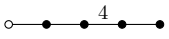
with $n \geq 2$.

$(\tilde{D}_n)$  with $n \geq 4$.

$(\tilde{E}_6)$ 

$(\tilde{E}_7)$ 

$(\tilde{E}_8)$ 

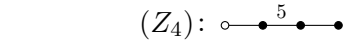
$(\tilde{F}_4)$ 

$(\tilde{G}_2)$ 

We show that the corresponding bilinear form is positive semidefinite, but not positive definite. For $\tilde{A}_n$, the row sums of M are all $2 - 1 - 1 = 0$, i. e., 0 is an eigenvalue and $\det M = 0$. In all other cases we can apply the above recursion formula (note that $\sin(\pi/6) = 1/2$):

$$
\begin{aligned}
(\tilde{B}_n)&\colon \det M = 2 \cdot 2 - 2 \cdot 2 = 0,\\
(\tilde{C}_n)&\colon \det M = 2 \cdot 2 - 2 \cdot 2 = 0,\\
(\tilde{D}_n)&\colon \det M = 2 \cdot 4 - 2 \cdot 4 = 0,\\
(\tilde{E}_6)&\colon \det M = 2 \cdot 3 - 6 = 0,\\
(\tilde{E}_7)&\colon \det M = 2 \cdot 2 - 4 = 0,\\
(\tilde{E}_8)&\colon \det M = 2 \cdot 1 - 2 = 0,\\
(\tilde{F}_4)&\colon \det M = 2 \cdot 1 - 2 = 0,\\
(\tilde{G}_2)&\colon \det M = 2 \cdot 1 - 2 = 0.
\end{aligned}
$$

If we add a vertex to H_3 and H_4, we obtain the following Coxeter graphs:

$(Z_4)\colon$ 

$(Z_5)\colon$ 

[28]The notation is not uniform in the literature.

Since

$$(Z_4)\colon \det M = 2(3-\sqrt{5}) - 3 = 3 - 2\sqrt{5} < 0,$$
$$(Z_5)\colon \det M = 7 - 3\sqrt{5} - (3-\sqrt{5}) = 4 - 2\sqrt{5} < 0$$

the corresponding bilinear forms are no longer positive semidefinite.

Uniqueness: Now let G be a finite irreducible Coxeter group of rank n. Let m be the largest edge label. Suppose that $C(G)$ does not belong to the positive (semi)definite graphs described above.

(1) Since all connected graphs with $n = 2$ have already been listed, we have $n \geq 3$.

(2) By Lemma 11.44, $\tilde{A}_1$ is not a subgraph of $C(G)$ and hence $m < \infty$.

(3) Since $\tilde{A}_k$ is not a subgraph, $C(G)$ must be a tree. First suppose $m = 3$.

(4) Since $C(G) \neq A_n$, $C(G)$ has a branch vertex.

(5) Since $\tilde{D}_k$ is not a subgraph, there is exactly one branch vertex e.

(6) Since $\tilde{D}_4$ is not a subgraph, e has exactly three branches. Let $a \leq b \leq c$ be the numbers of vertices of the three branches.

(7) Since $\tilde{E}_6$ is not a subgraph, $a = 1$.

(8) Since $\tilde{E}_7$ is not a subgraph, $b \leq 2$.

(9) Since $C(G) \neq D_n$, we have $b = 2$.

(10) Since $\tilde{E}_8$ is not a subgraph, $c \leq 4$.

(11) Since $C(G)$ is neither E_6, E_7 nor E_8, the case $m = 3$ cannot occur. So let $m \geq 4$.

(12) Since $\tilde{C}_k$ is not a subgraph, there can be only one edge with label > 3.

(13) Since $\tilde{B}_k$ is not a subgraph, there are no branch vertices, i. e., $C(G)$ is a path. Now suppose $m = 4$.

(14) Since $C(G) \neq B_n$, the two outer edges have label 3.

(15) Since $\tilde{F}_4$ is not a subgraph, we must have $n = 4$.

(16) Since $C(G) \neq F_4$, $m = 4$ cannot hold. So let $m \geq 5$.

(17) Since $\tilde{G}_2$ is not a subgraph, $m = 5$.

(18) Since Z_4 is not a subgraph, an outer edge must have label m.

(19) Since Z_5 is not a subgraph, we must have $n \leq 4$.

(20) But now $C(G)$ would be equal to H_3 or H_4. Contradiction. □

Theorem 11.46. *For the groups in Theorem 11.45 we have*

(A_n) $G \cong S_{n+1}$

(B_n) $G \cong C_2 \wr S_n$.

(D_n) $G \cong C_2^{n-1} \rtimes S_n$.

(E_6) $G \cong \mathrm{Aut}(\mathrm{SU}(4,2)) \cong \mathrm{SU}(4,2) \rtimes C_2$ *with* $|G| = 51{,}840 = 2^7 3^4 5$.

(E_7) $G \cong \mathrm{Sp}(6,2) \times C_2$ *with* $|G| = 2{,}903{,}040 = 2^{10}3^4 5 \cdot 7$.

(E_8) $G = 2.\mathrm{GO}^+(8,2)$ *with* $|G| = 696{,}729{,}600 = 2^{14}3^5 5^2 7$.

(F_4) $G \cong (C_2^3 \rtimes S_4) \rtimes S_3$.

(H_3) $G \cong A_5 \times C_2$. $(I_2(m))$ $G \cong D_{2m}$.

(H_4) $|G| = 14{,}400 = 2^6 3^2 5^2$.

Proof sketch. For (A_n) and $(I_2(m))$ the claim is known from Example 11.32. Since $[.,.]_G$ is positive definite, $[.,.]_G$ coincides with the standard inner product on $\mathbb{R}^n$ up to a choice of basis. We choose concrete vectors $b_1, \ldots, b_n$ with the prescribed values $[b_i, b_j]$ and determine $\sigma(G)$ explicitly. Let $e_1, \ldots, e_n$ be the standard basis of $\mathbb{R}^n$.

(B_n) Define $b_i := \frac{1}{\sqrt{2}}(e_i - e_{i+1})$ for $i = 1, \ldots, n-1$ and $b_n = e_n$. We have $[b_i, b_i] = 1$, $[b_i, b_j] = 0$ for $|j - i| > 1$, $[b_i, b_{i+1}] = -1/2 = -\cos(\pi/3)$ for $i = 1, \ldots, n-2$ and $[b_{n-1}, b_n] = -1/\sqrt{2} = -\cos(\pi/4)$ as desired. We have $b_i^\perp = \langle e_j, e_i + e_{i+1} : i \neq j \neq i+1 \rangle$. With respect to $e_1, \ldots, e_n$, $\sigma(x_i)$ corresponds to the permutation matrix of the transposition $(i, i+1)$ for $i = 1, \ldots, n-1$ and $\sigma(x_n) = \mathrm{diag}(1, \ldots, 1, -1)$. Hence $\sigma(G)$ is the group of signed permutation matrices. This shows $G \cong C_2 \wr S_n$.

(D_n) Define $b_i := \frac{1}{\sqrt{2}}(e_i - e_{i+1})$ for $i = 1, \ldots, n-1$ and $b_n = \frac{1}{\sqrt{2}}(e_{n-1} + e_n)$. We have $[b_{n-2}, b_{n-1}] = [b_{n-2}, b_n] = -1/2$ and $[b_{n-1}, b_n] = 0$. Again $\sigma(x_i)$ for $i < n$ are permutation matrices and

$$\sigma(x_n) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & -1 \\ & & & -1 & 0 \end{pmatrix}.$$

Hence $\sigma(G)$ consists of the signed permutation matrices with an even number of entries -1. The corresponding diagonal matrices form a normal subgroup C_2^{n-1} with complement S_n in G.

(H_3) For $x := x_3 x_2$, $y := x_2 x_1$ we have $|\langle x \rangle| = 3$, $|\langle y \rangle| = 5$ and $|\langle xy \rangle| = 2$. By Example 2.16 there exists an epimorphism $A_5 \to \langle x, y \rangle =: H$. Since A_5 is simple, it follows that $H \cong A_5$. From $\sigma(H) \leq \mathrm{SL}(3, \mathbb{R})$ and $\det(\sigma_1) = -1$ we obtain $|G| = 2|H| = 120$. By Remark 6.3 we have $G \in \{S_5, A_5 \times C_2\}$. In the case $G \cong S_5$ we would have $D_{10} \cong \langle x_1, x_2 \rangle \leq A_5$, in contradiction to $\det(x_1) = -1$.

The other cases are rather laborious and can be handled with GAP:

```
for para in [["E",6],["E",7],["E",8],["F",4]] do
  L:=SimpleLieAlgebra(para[1],para[2],Rationals);
  R:=RootSystem(L);;
  W:=WeylGroup(R);
  Print(para," ",Order(W),"\n");
od;
F:=FreeGroup("a","b","c","d");;
AssignGeneratorVariables(F);;
H4:=F/[a^2,b^2,c^2,d^2,(a*b)^5,(b*c)^3,(c*d)^3,Comm(a,c),Comm(a,d),Comm(b,d)];
    #Comm(a,b)=[a,b]
Size(H4);
```

□

Remark 11.47.

(i) The matrix -1_n clearly lies in the center of the Coxeter group $G \cong C_2^n \rtimes S_n$ of type (B_n) (signed permutation matrices). By Theorem 11.36, $\mathrm{Z}(G) = \langle -1_n \rangle$. If n is odd, then $G = \mathrm{Z}(G) \times H$, where

H is isomorphic to the Coxeter group of type (D_n) (permutation matrices with an even number of entries -1). Nevertheless, G is irreducible. For $n = 3$ we obtain

$$G \cong C_2 \times (C_2^2 \rtimes S_3) \cong C_2 \times S_4.$$

This group permutes the eight vertices of the cube $(\pm 1, \pm 1, \pm 1)$ and hence is its symmetry group (and that of the octahedron with vertices $(\pm 1, 0, 0)$, $(0, \pm 1, 0)$, $(0, 0, \pm 1)$). For $n \geq 4$ we obtain the symmetry group of the n-dimensional hypercube. The Coxeter group $H \cong C_2^2 \rtimes S_3 \cong S_4$ of type (D_3) permutes the four vertices $(1, 1, 1)$, $(1, -1, -1)$, $(-1, 1, -1)$, $(-1, -1, 1)$ of the tetrahedron and hence is its symmetry group. As is well known, S_4 is also the Coxeter group of type (A_3).

(ii) One can show that the Coxeter group of type (H_3) is the symmetry group of the dodecahedron (or of the icosahedron). The Coxeter group of type (F_4) is an extension of the group of type (D_4) by S_3. The Coxeter group of type (E_8) is a Schur extension of the orthogonal group $\mathrm{GO}^+(8, 2) \cong \Omega^+(8, 2).2$ with center C_2.

(iii) A (complex) *Lie algebra* is a finite-dimensional $\mathbb{C}$-vector space L with an alternating bilinear map $L \times L \to L$, $(v, w) \mapsto [v, w]$, such that the *Jacobi identity*

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$$

holds. For example, $L = \mathbb{C}^{n \times n}$ with $[v, w] := vw - wv$. One calls L *simple* if $[L, L] \neq 0$ and if there is no subspace $0 < U < L$ with $[L, U] \subseteq U$. The classification of the simple Lie algebras leads to the diagrams A_n, B_n, C_n, D_n, E_6, E_7, E_8, F_4 and G_2.[29] These correspond more or less to the simple groups of Lie type (for example, A_n corresponds to the family $\mathrm{PSL}(n + 1, q)$ and C_n to the family $\mathrm{PSp}(2n, q)$).

(iv) One calls $\sigma \in \mathrm{GL}(\mathbb{C}^d)$ a *complex reflection* if $|\langle \sigma \rangle| < \infty$ and $\dim \mathrm{Ker}(\sigma - \mathrm{id}) = d - 1$. A *complex reflection group* is a finite subgroup $G \leq \mathrm{GL}(\mathbb{C}^d)$ that is generated by complex reflections. Clearly, every reflection group is also a complex reflection group. Shephard and Todd have completely classified the irreducible complex reflection groups. Apart from 34 exceptions, every such group has the form

$$G(m, d, n) := \big\langle (x_1, \ldots, x_n; \sigma) \in C_m \wr S_n : (x_1 \ldots x_n)^{m/d} = 1 \big\rangle,$$

where $n, m, d \in \mathbb{N}$ and $d \mid m$. Clearly, $G(m, 1, 1) \cong C_m$, $G(1, 1, n) \cong S_n$, $G(2, 1, n) \cong C_2 \wr S_n$ (type (B_n)), $G(2, 2, n)$ is of type (D_n) and $G(m, m, 2) \cong D_{2m}$ (type $(I_2(m))$). The 34 exceptional groups are unfortunately only available in the old GAP version 3.[30]

Definition 11.48. For $a, b, c \in \mathbb{Z} \setminus \{0\}$ one calls

$$D(a, b, c) := \langle x, y \mid x^a = y^b = (xy)^c = 1 \rangle$$

a *von Dyck group*.

Lemma 11.49. *$D(a, b, c)$ does not depend on the order and the signs of a, b, c.*

Proof. The independence of the signs is obvious. So let $a, b, c \geq 1$. Since $(xy)^c = 1 \iff (yx)^c = 1$, we can swap a and b. Now let $x' := xy$ and $y' := y^{-1}$. Then $x^a = y^b = (xy)^c = 1$ is equivalent to $(x')^c = (y')^b = (x'y')^a = 1$. So we can also swap a and c. □

[29] See the Algebra lecture notes

[30] `https://github.com/jmichel7/gap3-jm`

Theorem 11.50. *$D(a,b,c)$ is finite if and only if $\frac{1}{a}+\frac{1}{b}+\frac{1}{c}>1$. In this case, one of the following cases occurs:*

(i) $D(1,b,c)\cong C_{\gcd(b,c)}$.

(ii) $D(2,2,c)\cong D_{2c}$.

(iii) $D(2,3,3)\cong A_4$.

(iv) $D(2,3,4)\cong S_4$.

(v) $D(2,3,5)\cong A_5$.

Proof. By Lemma 11.49 we may assume $1\le a\le b\le c$. First let $a=1$. By Euclid there exist $\alpha,\beta\in\mathbb{Z}$ with $\gcd(b,c)=\alpha b+\beta c$. Hence $y^{\gcd(b,c)}=(y^b)^\alpha(y^c)^\beta=1$ in $D(1,b,c)$. It follows that $|D(1,b,c)|\le\gcd(b,c)$. Conversely, $C_{\gcd(b,c)}$ also satisfies the relations of $D(1,b,c)$. Hence (i) holds.

In the case $a=b=2$ the claim follows from Example 1.17. The case $(a,b,c)=(2,3,3)$ was treated in Exercise 2.

Let $(a,b,c)=(2,3,4)$ and $G:=\langle x,y\mid x^4=y^2=(xy)^3=1\rangle\cong D(2,3,4)$ and $H:=\langle x\rangle\le G$. We consider the cosets

$$H, yH, xyH, x^2yH, x^3yH, yx^2yH.$$

Since $xyx=yx^{-1}y$ and $xyx^2y=yx^{-1}yxy=yx^2xyxy=yx^2yx^{-1}$, the cosets are permuted by left multiplication by x. Since $yxy=x^{-1}yx^{-1}=x^3yx^3$, also y permutes these cosets and we see that every element of G is contained in one of these cosets. Hence $|G|\le|G:H||H|\le 24$. On the other hand, one sees that $x':=(1,2,3,4)$ and $y':=(1,2)$ in S_4 also satisfy the relations of G. This yields (iv). The case $(a,b,c)=(2,3,5)$ was already treated several times (Remark 1.18, Example 2.16). One easily sees that $\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\le 1$ holds in all other cases.

Let

$$H:=\langle u,v,w\mid u^2=v^2=w^2=(uv)^a=(vw)^b=(uw)^c=1\rangle$$

be a Coxeter group. Clearly $N:=\langle uv,vw\rangle\le H$ with $|H:N|=2$. Moreover, there exists an epimorphism $D(a,b,c)\to N$. So it suffices to show that H is infinite for $\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\le 1$. In the case $a\ge 3$, $C(H)$ is a cycle and $|H|=\infty$ by Theorem 11.45. So let $a=2$ and $\frac{1}{b}+\frac{1}{c}\le\frac{1}{2}$. Then $b,c\ge 3$ and H is irreducible. For $b=3$ we have $c\ge 6$ and the claim follows in any case from Theorem 11.45. □

Remark 11.51. A finite group $G\neq 1$ is called a *Hurwitz group* if $G=\langle x,y\rangle$ with $x^2=y^3=(xy)^7=1$. These are thus exactly the finite nontrivial quotients of the infinite von Dyck group $D(2,3,7)$. By Exercise 37, all Hurwitz groups are perfect and indeed many simple groups are Hurwitz groups, among others A_n for $n\ge 168$, $\mathrm{GL}(3,2)$ and the sporadic *Monster group*.

Theorem 11.52. *For $n\in\mathbb{N}$ there are infinitely many primes $p\equiv 1 \pmod n$.*[31]

[31] This is a special case of Dirichlet's theorem on primes in arithmetic progressions: The primes are equally distributed among the reduced residue classes modulo n, i. e., for $\gcd(k,n)=1$ the proportion of primes in $k+n\mathbb{Z}$ is exactly $1/\varphi(n)$.

Proof. Let $p_1, \ldots, p_s$ be primes with $p_i \equiv 1 \pmod{n}$ for $i = 1, \ldots, s$ (the case $s = 0$ is allowed). Let $m := np_1 \ldots p_s$. The cyclotomic polynomial Φ_m, being monic, induces an unbounded function $\mathbb{R} \to \mathbb{R}$. Hence there exists $k \in \mathbb{N}$ with $\Phi_m(km) > 1$. Let p be a prime divisor of $\Phi_m(km)$. Since $\Phi_m(km) \mid ((km)^m - 1)$, we have $p \nmid m$ and the order r of $mk + p\mathbb{Z} \in \mathbb{F}_p^\times$ divides $\gcd(m, p-1)$. Suppose that $r < m$ and set $a := (mk)^r \equiv 1 \pmod{p}$. Then we have

$$\frac{m}{r} \equiv 1 + a + \ldots + a^{\frac{m}{r}-1} = \frac{a^{\frac{m}{r}} - 1}{a - 1} = \frac{(mk)^m - 1}{(mk)^r - 1} = \prod_{\substack{d \mid m \\ d \nmid r}} \Phi_d(mk) \equiv 0 \pmod{p}$$

in contradiction to $p \nmid m$. Hence $r = m$ and $m \mid p - 1$, i. e., $p \equiv 1 \pmod{m}$. Since $n \mid m$, we also have $p \equiv 1 \pmod{n}$. On the other hand, $m = np_1 \ldots p_s$ and $p \notin \{p_1, \ldots, p_s\}$. So we have found a new prime in the residue class $1 + n\mathbb{Z}$. □

Theorem 11.53 (Miller). *For $a, b, c \in \mathbb{N} \setminus \{1\}$ there exists a finite group $G = \langle x, y \rangle$ with $|\langle x \rangle| = a$, $|\langle y \rangle| = b$ and $|\langle xy \rangle| = c$.*

Proof (Holt). Instead of constructing G as a quotient of a von Dyck group, we give an independent proof. Theorem 11.52 guarantees the existence of a prime $p \equiv 1 \pmod{2abc}$. Let $\zeta_a, \zeta_b, \zeta_c \in \mathbb{F}_p^\times$ be elements of order $2a, 2b$ and $2c$, respectively. First let $\lambda \in \mathbb{F}_p$ be arbitrary. Then

$$x := \begin{pmatrix} \zeta_a & 1 \\ 0 & \zeta_a^{-1} \end{pmatrix} \in \mathrm{SL}(2, p), \qquad y := \begin{pmatrix} \zeta_b & 0 \\ \lambda & \zeta_b^{-1} \end{pmatrix} \in \mathrm{SL}(2, p)$$

have the eigenvalues $\zeta_a \neq \zeta_a^{-1}$ and $\zeta_b \neq \zeta_b^{-1}$, respectively. In particular, x, y are diagonalizable and it follows that $|\langle x \rangle| = 2a$ and $|\langle y \rangle| = 2b$. Since $\mathrm{tr}(xy) = \zeta_a\zeta_b + \zeta_a^{-1}\zeta_b^{-1} + \lambda$, we can choose λ such that $\mathrm{tr}(xy) = \zeta_c + \zeta_c^{-1}$ holds. For the eigenvalues e, f of xy (in a splitting field) we have $e + f = \mathrm{tr}(xy)$ and $ef = \det(xy) = \det(x)\det(y) = 1$. This shows $\{e, f\} = \{\zeta_c, \zeta_c^{-1}\}$ and xy is also diagonalizable with order $2c$. Since x^a, x^b and $(xy)^c$ have the double eigenvalue -1, we have $x^a = y^b = (xy)^c = -1_2 \in Z := \mathrm{Z}(\mathrm{SL}(2, p))$. With $\overline{x} := xZ$ and $\overline{y} := yZ$, the group $G = \langle \overline{x}, \overline{y} \rangle \leq \mathrm{PSL}(2, p)$ satisfies the claim. □

12 Free Products and Amalgamated Free Products

Example 12.1. The direct product $G \times H$ of two groups G and H is the largest group D with the following properties:

- $G, H \leq D$ and $D = \langle G, H \rangle$,
- $xy = yx$ for all $x \in G$ and $y \in H$.

We construct the largest group satisfying only the first property.

Definition 12.2. Let $G_I := \{G_i : i \in I\}$ be a nonempty family of groups. A *free product* of G_I is a group G with homomorphisms $\lambda_i \colon G_i \to G$ $(i \in I)$ with the following universal property: For every group H and homomorphisms $\rho_i \colon G_i \to H$ there exists exactly one homomorphism $\varphi \colon G \to H$ with $\rho_i = \varphi\lambda_i$ for all $i \in I$.

Lemma 12.3.

(i) The homomorphisms λ_i are injective and $G = \langle \lambda_i(G_i) : i \in I \rangle$.

(ii) Up to isomorphism there is at most one free product of G_I.

Proof.

(i) Let $H := G_i$, $\rho_i := \mathrm{id}_H$ and $\rho_j := 1$ for $j \neq i$. Then there exists a $\varphi\colon G \to H$ with $\mathrm{id}_H = \rho_i = \varphi\lambda_i$. In particular, λ_i is injective. Now let $H := \langle \lambda_i(G_i) : i \in I\rangle \leq G$ and $\rho_i := \lambda_i\colon G_i \to H$. Then there exists exactly one $\varphi\colon G \to H$ with $\lambda_i = \varphi\lambda_i$ for $i \in I$. For the same reason there exists exactly one homomorphism $\varphi'\colon G \to G$ with $\lambda_i = \varphi'\lambda_i$. Clearly, we must have $\varphi' = \mathrm{id}_G$. If we interpret φ as a map $G \to G$, we also have $\varphi = \varphi'$. This shows $G = H$.

(ii) Let H also be a free product of G_I with homomorphisms $\mu_i\colon G_i \to H$. Then there exist homomorphisms $\varphi\colon G \to H$ and $\psi\colon H \to G$ with $\mu_i = \varphi\lambda_i$ and $\lambda_i = \psi\mu_i$ for $i \in I$. It follows that $\psi\varphi\lambda_i = \psi\mu_i = \lambda_i$ and $\varphi\psi\mu_i = \varphi\lambda_i = \mu_i$. Now, as in (i), id_G and id_H are the only homomorphisms with $\mathrm{id}_G\,\lambda_i = \lambda_i$ and $\mathrm{id}_H\,\mu_i = \mu_i$. This shows $\varphi\psi = \mathrm{id}_H$ and $\psi\varphi = \mathrm{id}_G$. In particular, $\varphi\colon G \to H$ is an isomorphism. □

Remark 12.4. By Lemma 12.3, we can speak of *the* free product of G_I without mentioning the homomorphisms λ_i. One writes $G = \mathrm{Fr}_{i\in I} G_i$ or $G_1 * \ldots * G_n$ if $I = \{1, \ldots, n\}$.[32] Since the λ_i are injective, we may assume $G_i \leq G$ (this corresponds to the formal difference between the direct product and the direct sum).

Theorem 12.5. *For every nonempty family of groups G_I, the free product $\mathrm{Fr}_{i\in I} G_i$ exists.*

Proof. We modify Definition 12.2. W. l. o. g. let $G_i \cap G_j = \varnothing$ for $i \neq j$. Let W be the set of all formal words of the form $w = g_1 \ldots g_n$ with $n \in \mathbb{N}_0$ and $g_1, \ldots, g_n \in \bigcup_{i\in I} G_i$. One calls w *reduced* if $g_1, \ldots, g_n \neq 1$ and $\{g_i, g_{i+1}\} \nsubseteq G_j$ for $i = 1, \ldots, n-1$ and all $j \in I$. Clearly, w can always be reduced. Words $v, w \in W$ are called *equivalent* if they can be reduced to the same word. Let $G = \{[w] : w \in W\}$ be the set of equivalence classes of this relation. Clearly, G is a group with respect to concatenation. We define $\lambda_i\colon G_i \to G$ by $\lambda_i(g) := [g]$ for $g \in G_i$. Now let H be another group and $\rho_i\colon G_i \to H$ homomorphisms for $i \in I$. The map $\varphi\colon G \to H$ with $[w] \mapsto \rho_{i_1}(g_1) \ldots \rho_{i_n}(g_n)$ for $g_j \in G_{i_j}$ is a well-defined homomorphism with $\rho_i = \varphi\lambda_i$ for $i \in I$. Since G is generated by construction by the elements $[g]$ with $g \in \bigcup G_i$, φ is uniquely determined by the ρ_i. □

Lemma 12.6. *Every element in $\mathrm{Fr}_{i\in I} G_i$ can be written uniquely as a reduced word.*

Proof. We proceed as in Lemma 1.7. Let R be the set of reduced words $r = g_1 \ldots g_n$ with $g_1, \ldots, g_n \in \bigcup G_i$. For $g \in G_j$ let

$$ {}^g r := \begin{cases} r & \text{if } g = 1, \\ g g_1 \ldots g_n & \text{if } g_1 \notin G_j, \\ g_2 \ldots g_n & \text{if } g = g_1^{-1}, \\ (g g_1) g_2 \ldots g_n & \text{otherwise.} \end{cases} $$

One easily sees that this describes an action $G_j \to \mathrm{Sym}(R)$. By the universal property, the action can be extended to $G \to \mathrm{Sym}(R)$. For $v, w \in R$ with $[v] = [w]$ we now have $v = {}^{[v]}1 = {}^{[w]}1 = w$. □

[32] Note: Do not confuse this with the central product.

Remark 12.7. In what follows, we will replace $[w]$ by w. Every element in $\operatorname{Fr} G_i$ can then be written uniquely in the reduced form $g_1 \dots g_n$. Conversely, if $H = \langle G_i : i \in I \rangle$ such that every element in H can be written uniquely in reduced form, then $H \cong \operatorname{Fr}_{i \in I} G_i$, since one can extend the inclusions $\sigma_i \colon G_i \hookrightarrow H$ to an isomorphism.

Example 12.8.

(a) A free product of free groups is free: Let $F_i = F_{X_i}$ with $X_i \cap X_j = \varnothing$ for $i \neq j$. Let $X := \bigcup_{i \in I} X_i$ and $F := F_X$. Let $\rho_i \colon F_i \to H$ be homomorphisms. Then there exists exactly one homomorphism $\varphi \colon F \to H$ with $\varphi(x) := \rho_i(x)$ for $x \in X_i$. Since $F_i = \langle X_i \rangle$, we have $\rho_i = \varphi_{|F_i}$. As a special case we obtain $F_n = \mathbb{Z} * \ldots * \mathbb{Z}$.

(b) Let $G = \langle x \rangle * \langle y \rangle \cong C_2 * C_2$ and $z := xy \in G$. Since $z^{-1} = yx$, every element in G can be written uniquely in the form $z^a x^b$ with $a \in \mathbb{Z}$ and $b \in \{0, 1\}$. Since $xzx^{-1} = z^{-1}$, we have $G = \langle z \rangle \rtimes \langle x \rangle \cong \mathbb{Z} \rtimes C_2 \cong D_\infty$.

Lemma 12.9. *We have* $\operatorname{PSL}(2, \mathbb{Z}) \cong C_2 * C_3$.

Proof. Let $A := \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $B := \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$ in $\operatorname{SL}(2, \mathbb{Z})$. We first show $\operatorname{SL}(2, \mathbb{Z}) = \langle A, B \rangle$. Suppose indirectly that $C = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \operatorname{SL}(2, \mathbb{Z}) \setminus \langle A, B \rangle$ with $|a| + |c|$ minimal. Let us assume $a \neq 0 \neq c$. Then

$$(AB)^s C = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + sc & b + sd \\ c & d \end{pmatrix} \notin \langle A, B \rangle,$$
$$(BA)^{-r} C = \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ ra + c & rb + d \end{pmatrix}.$$

In the case $|a| \geq |c|$ we can choose s such that $|a + sc| + |c| < |a| + |c|$. Otherwise we can choose r such that $|a| + |ra + c| < |a| + |c|$. Contradiction. Hence $a = 0$ or $c = 0$. In the first case we would have

$$C = \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & d \end{pmatrix} \in \{BA^2(AB)^{-d-1}, B(AB)^{d-1}\}$$

and in the second case we would have

$$C = \begin{pmatrix} \pm 1 & b \\ 0 & \pm 1 \end{pmatrix} \in \{(AB)^b, A^2(AB)^{-b}\}.$$

This shows $\operatorname{SL}(2, \mathbb{Z}) = \langle A, B \rangle$.

Since $A^2 = -1 = B^3$, the elements $\overline{A}, \overline{B} \in \operatorname{PSL}(2, \mathbb{Z})$ have order 2 and 3, respectively (note that $\operatorname{Z}(\operatorname{SL}(2, \mathbb{Z})) = \langle -1_2 \rangle$). Let $G := \langle x \rangle * \langle y \rangle \cong C_2 * C_3$. Then there exists an epimorphism $\varphi \colon G \to \operatorname{PSL}(2, \mathbb{Z})$ with $\varphi(x) = \overline{A}$ and $\varphi(y) = \overline{B}$. Suppose that there exists $w \in \operatorname{Ker}(\varphi) \setminus \{1\}$. Then w is an alternating product of x and $y^{\pm 1}$. After conjugation we may assume $w = xy^{\epsilon_1} \dots xy^{\epsilon_n}$ with $\epsilon_1, \dots, \epsilon_n \in \{\pm 1\}$. The matrices AB and $AB^{-1} = \left(\begin{smallmatrix} -1 & 0 \\ -1 & -1 \end{smallmatrix}\right)$ have nonnegative and nonpositive entries, respectively. The same property would also have to hold for

$$\overline{AB^{\epsilon_1} \dots AB^{\epsilon_{n-1}} AB^{-\epsilon_n}} = \varphi(wy^{\epsilon_n}) = \varphi(w)\varphi(y^{\epsilon_n}) = \overline{B^{\epsilon_n}}$$

However, B^{ϵ_n} has both positive and negative entries. Hence φ is an isomorphism. □

Lemma 12.10. *Let* $F := \operatorname{Fr}_{i \in I} G_i$ *with* $G_i \neq 1$ *for all* $i \in I$. *Then:*

(i) Every element $g \in F$ *of finite order lies in some* G_i *up to conjugation.*

(ii) For $g \in G_i \setminus \{1\}$ we have $\mathrm{C}_F(g) = \mathrm{C}_{G_i}(g)$.

(iii) For $|I| \geq 2$ we have $|F| = \infty$ and $\mathrm{Z}(F) = 1$.

Proof.

(i) Let $g = g_1 \dots g_n \in F$ be reduced with $n \geq 2$. After conjugation we may assume that g_1 and g_n lie in different factors G_i. Clearly, $g^k \neq 1$ for all $k \geq 1$.

(ii) Suppose that there exists a reduced $x = x_1 \dots x_n \in \mathrm{C}_F(g) \setminus G_i$ of minimal length. In the case $x_1 = g^{-1}$ or $x_n = g$ we would also have $gx \in \mathrm{C}_F(g) \setminus G_i$ or $xg^{-1} \in \mathrm{C}_F(g) \setminus G_i$, respectively, of shorter length. Otherwise gxg^{-1} would be reduced, but different from x. Contradiction.

(iii) For $g \in G_i \setminus \{1\}$ and $h \in G_j \setminus \{1\}$ with $i \neq j$, gh is reduced of infinite order. Moreover, $\mathrm{Z}(F) \leq \mathrm{C}_F(g) \cap \mathrm{C}_F(h) \leq G_i \cap G_j = 1$ by (ii). □

Theorem 12.11. *Let $G = \mathrm{Fr}_{i\in I} G_i$ and $H = \mathrm{Fr}_{i\in I} H_i$ be free products over the same index set I. Let $\sigma_i \colon G_i \to H_i$ be homomorphisms for $i \in I$. Then there is exactly one homomorphism $\varphi \colon G \to H$ extending all σ_i. Moreover, $\mathrm{Ker}(\varphi)$ is the normal closure of $\bigcup_{i\in I} \mathrm{Ker}(\sigma_i)$ in G.*

Proof. The existence and uniqueness of φ follow from the universal property. Let $N \trianglelefteq G$ be the normal closure of $\bigcup_{i\in I} \mathrm{Ker}(\sigma_i)$ in G. Certainly $N \subseteq \mathrm{Ker}(\varphi)$. Let us assume $\mathrm{Ker}(\varphi) \nsubseteq N$ and choose $g = g_1 \dots g_n \in \mathrm{Ker}(\varphi) \setminus N$ reduced with minimal n and $g_k \in G_{i_k}$ for $k = 1, \dots, n$. If $\sigma_{i_k}(g_k) \neq 1$ for $k = 1, \dots, n$, then $1 = \varphi(g) = \sigma_{i_1}(g_1) \dots \sigma_{i_n}(g_n)$ would be reduced in H. This contradiction shows $\sigma_{i_k}(g_k) = 1$ for some $1 \leq k \leq n$. But then $g_1 \dots g_{k-1} g_{k+1} \dots g_n \in \mathrm{Ker}(\varphi) \setminus N$, contradicting the choice of n. □

Corollary 12.12. *Let $G_i = \langle X_i \mid R_i \rangle$ be presentations for $i \in I$. Then*

$$\mathrm{Fr}_{i\in I} G_i \cong \langle \dot{\bigcup}_{i\in I} X_i \mid \dot{\bigcup}_{i\in I} R_i \rangle.$$

Proof. Let $F_i := F_{X_i}$ and let F be the free group on $\dot{\bigcup}_{i\in I} X_i$. By Example 12.8, $F = \mathrm{Fr}_{i\in I} F_i$. Let $\sigma_i \colon F_i \to G_i$ be the canonical epimorphism. By Theorem 12.11 there exists an epimorphism $\varphi \colon F \to \mathrm{Fr}_{i\in I} G_i$ with $\mathrm{Ker}(\varphi) = \langle \bigcup \mathrm{Ker}(\sigma_i) \rangle^F = \langle \bigcup R_i^{F_i} \rangle^F = \langle \bigcup R_i \rangle^F$. This proves the claim. □

Remark 12.13.

(i) Kurosh showed that every subgroup H of $F := \mathrm{Fr}_{i\in I} G_i$ has the form

$$F_0 * \mathrm{Fr}_{i\in I} \mathrm{Fr}_{HxG_i \in H\backslash F/G_i} (H \cap xG_i x^{-1})$$

where F_0 is free (here $H\backslash F/G_i$ is the set of *double cosets* with respect to H and G_i).

(ii) From direct products one can make central products by factoring out central subgroups. We construct a noncommutative variant.

Definition 12.14. Let $G_I = \{G_i : i \in I\}$ be a family of groups. Let H be a group and $\sigma_i \colon H \to G_i$ monomorphisms for $i \in I$. Let N be the normal closure of $\{\sigma_i(h)^{-1}\sigma_j(h) : h \in H, i, j \in I\}$ in $F := \mathrm{Fr}_{i\in I} G_i$. One calls F/N the *amalgamated free product of G_I over H*.

Remark 12.15. Since $\sigma_i(h) \equiv \sigma_j(h) \pmod N$, we have $\sigma_i(H)N/N = \sigma_j(H)N/N \leq F/N$. Thus one identifies the isomorphic copies of H in G_i. In the case $H = 1$ we have $N = 1$ and obtain the ordinary free product. In contrast to the free product, F/N also depends on the choice of the monomorphisms σ_i.

Example 12.16. We proved $\mathrm{PSL}(2,\mathbb{Z}) = \langle \overline{A}, \overline{B}\rangle \cong C_2 * C_3$ in Lemma 12.9. One easily sees that $\mathrm{SL}(2,\mathbb{Z}) = \langle A, B\rangle$ is an amalgamated free product of $\langle A\rangle = C_4$ and $\langle B\rangle \cong C_6$ over $\langle A^2\rangle = \langle B^3\rangle = \langle -1_2\rangle = \mathrm{Z}(\mathrm{SL}(2,\mathbb{Z}))$.

Lemma 12.17. *In the situation of Definition 12.14, let R_i be a set of coset representatives for the cosets of $\sigma_i(H)$ in G_i with $1 \in R_i$. Then every element in F/N can be written uniquely in the reduced form $r_1 \dots r_n hN$ with $n \in \mathbb{N}_0$, $r_k \in R_{i_k} \setminus \{1\}$, $i_k \neq i_{k+1}$ and $h \in \sigma_{i_n}(H)$.*

Proof. Let $g = g_1 \dots g_n N \in F/N$ with $g_k \in G_{i_k}$ arbitrary. Then there exist $r_1 = \overline{g_1} \in R_{i_1}$ and $h_1 \in H$ with $g_1 = r_1\sigma_{i_1}(h_1)$. Since $\sigma_{i_1}(h_1)N = \sigma_{i_2}(h_1)N$, there exist $r_2 = \overline{\sigma_{i_2}(h_1)g_2} \in R_{i_2}$ and $h_2 \in H$ with $g = r_1(\sigma_{i_2}(h_1)g_2)g_3 \dots g_n N = r_1 r_2 \sigma_{i_2}(h_2) g_3 \dots g_n N$ and so on. Elements $r_i = 1$ can be omitted. In this way g can be brought into reduced form.

Now let $R := \{r_1 \dots r_n \sigma_{i_n}(h) : r_i \in R_i, h \in H\} \subseteq F$ be the set of reduced words. We show as in Lemma 12.6 that F acts on R. It suffices to show that every G_j acts on R. Let $x = \overline{x}\sigma_j(h_1) \in G_j$ and $r = r_1 \dots r_n \sigma_{i_n}(h) \in R$. We first assume $j \neq i_1$ and define

$$ {}^x r := \overline{x}\,\overline{\sigma_{i_1}(h_1)r_1}\,\overline{\sigma_{i_2}(h_2)r_2} \dots \overline{\sigma_{i_n}(h_n)r_n}\sigma_{i_n}(h_{n+1}h) \in R $$

with $\sigma_{i_k}(h_k)r_k = \overline{\sigma_{i_k}(h_k)r_k}\sigma_{i_k}(h_{k+1})$ for $k = 1, \dots, n$ (in the case $x \in \sigma_j(H)$ one has to remove $\overline{x}$). For $y \in G_j$ it follows that

$$ {}^y({}^x r) = \overline{yx}\,\overline{\sigma_{i_1}(h_1'h_1)r_1} \dots \overline{\sigma_{i_n}(h_n'h_n)r_n}\sigma_{i_n}(h_{n+1}'h_{n+1}h) $$

with $y\overline{x} = \overline{yx}\sigma_j(h_1')$ and $\sigma_{i_k}(h_k')\overline{\sigma_{i_k}(h_k)r_k} = \overline{\sigma_{i_k}(h_k'h_k)r_k}\sigma_{i_k}(h_{k+1}')$ for $k = 1, \dots, n$. On the other hand,

$$ {}^{yx} r = \overline{yx}\,\overline{\sigma_{i_1}(h_1'')r_1} \dots \overline{\sigma_{i_n}(h_n'')r_n}\sigma_{i_n}(h_{n+1}''h) $$

with $yx = \overline{yx}\sigma_j(h_1'')$ and $\sigma_{i_k}(h_k'')r_k = \overline{\sigma_{i_k}(h_k'')r_k}\sigma_{i_k}(h_{k+1}'')$ for $k = 1, \dots, n$. It follows that

$$ \overline{yx}\sigma_j(h_1'') = yx = y\overline{x}\sigma_j(h_1) = \overline{yx}\sigma_j(h_1'h_1) $$

and $h_1'' = h_1'h_1$, since σ_j is injective. By induction on k we obtain

$$ \begin{aligned} \overline{\sigma_{i_k}(h_k'')r_k}\sigma_{i_k}(h_{k+1}'') &= \sigma_{i_k}(h_k'')r_k = \sigma_{i_k}(h_k'h_k)r_k = \sigma_{i_k}(h_k')\overline{\sigma_{i_k}(h_k)r_k}\sigma_{i_k}(h_{k+1}) \\ &= \overline{\sigma_{i_k}(h_k'h_k)r_k}\sigma_{i_k}(h_{k+1}'h_{k+1}) = \overline{\sigma_{i_k}(h_k'')r_k}\sigma_{i_k}(h_{k+1}'h_{k+1}), \end{aligned} $$

i. e., $h_{k+1}'h_{k+1} = h_{k+1}''$. The case $j = i_1$ works analogously. This proves ${}^y({}^x r) = {}^{yx} r$, and F acts on R with ${}^x[r] \equiv [xr] \pmod N$. For $h \in H$, $\sigma_i(h)$ acts on R in the same way as $\sigma_j(h)$ for $i, j \in I$. Hence N acts trivially and F/N acts on R. For $v, w \in R$ with $[v]N = [w]N$ we now have $v = {}^{[v]N}1 = {}^{[w]N}1 = w$. □

Lemma 12.18. *Let $G = F/N$ be an amalgamated free product of G_I over H and $\sigma_i \colon H \to G_i$. Then there exist subgroups $H \cong \overline{H} \leq G$ and $G_i \cong \overline{G_i} \leq G$ with $G = \langle \overline{G_i} : i \in I\rangle$ and $\overline{H} = \overline{G_i} \cap \langle \overline{G_j} : j \neq i\rangle$ for all $i \in I$.*

Proof. Let $\overline{G_i} := G_iN/N$ and $\overline{H} := \sigma_i(H)N/N \leq G$ (does not depend on i). From Lemma 12.17 it follows that $\overline{G_i} \cong G_i/G_i \cap N \cong G_i$. Since $\overline{H} \leq \overline{G_i}$, we also have $\overline{H} \cong H$. From $F = \langle G_i : i \in I \rangle$ it follows that $G = \langle \overline{G_i} : i \in I \rangle$. Finally, $\overline{H} = \overline{G_i} \cap \langle \overline{G_j} : j \neq i \rangle$ also follows from Lemma 12.17. □

Remark 12.19. In what follows, we will regard the groups G_i and H as subgroups of $G = F/N$. Let R_i be a set of coset representatives for G_i/H with $1 \in R_i$ for $i \in I$. Every element in G can then be written uniquely in the reduced form $r_1 \dots r_n h$ with $r_k \in R_{i_k} \setminus \{1\}$, $i_k \neq i_{k+1}$ and $h \in H$.

Lemma 12.20. *Let G be an amalgamated free product of G_I over H. Then:*

(i) Every element of finite order of G lies in some G_i up to conjugation.

(ii) If there exist distinct $i, j \in I$ with $G_i \neq H \neq G_j$, then $|G| = \infty$.

Proof.

(i) Let $g = r_1 \dots r_n h$ be in reduced form with finite order. After conjugation we may assume $i_1 \neq i_n$ (note: possibly all r_i are changed). Then the elements $g, g^2, \dots$ are all distinct. This contradiction shows $n \leq 1$ and $g \in G_{i_1}$.

(ii) For $g_1 \in G_i \setminus H$ and $g_2 \in G_j \setminus H$, $g_1 g_2$ has infinite order. □

Theorem 12.21. *Every finite group G is a subgroup of a finite group $\hat{G}$ with the following property: If $H, K \leq G$ are isomorphic subgroups and $\varphi\colon H \to K$ is an isomorphism, then there exists an $x \in \hat{G}$ with $\varphi(h) = xhx^{-1}$ for all $h \in H$. In particular, H and K are conjugate in $\hat{G}$.*

Proof. We set $\hat{G} := \mathrm{Sym}(G)$ and embed G via the regular representation $\sigma\colon G \to \hat{G}$, $x \mapsto \sigma_x$ with $\sigma_x(y) = xy$. Let $\{r_1, \dots, r_n\}$ and $\{s_1, \dots, s_n\}$ be sets of representatives for the right cosets of H and K in G, respectively.[33] Let $\hat{\varphi} \in \hat{G}$ with $\hat{\varphi}(hr_i) := \varphi(h)s_i$ for all $h \in H$ and $1 \leq i \leq n$. For $h \in H$, $k \in K$ and $1 \leq i \leq n$ we have

$$(\hat{\varphi}\sigma_h\hat{\varphi}^{-1})(ks_i) = (\hat{\varphi}\sigma_h)(\varphi^{-1}(k)r_i) = \hat{\varphi}(h\varphi^{-1}(k)r_i) = \varphi(h)ks_i = \sigma_{\varphi(h)}(ks_i)$$

and $\hat{\varphi}\sigma_h\hat{\varphi}^{-1} = \sigma_{\varphi(h)}$. □

Remark 12.22. The proof of Theorem 12.21 no longer works for infinite groups. For example, the canonical isomorphism $\varphi\colon \mathbb{Z} \to 2\mathbb{Z}$ cannot be extended to a bijection $\mathbb{Z} \to \mathbb{Z}$.

Theorem 12.23 (Higman-Neumann-Neumann)**.** *Let G be a group, $H, K \leq G$ and $\varphi\colon H \to K$ an isomorphism. Then G is a subgroup of a group $\hat{G}$ such that there exists an $x \in \hat{G}$ with $\varphi(h) = xhx^{-1}$ for all $h \in H$. In particular, H and K are conjugate in $\hat{G}$.*

Proof. Let $\langle a_i \rangle \cong \mathbb{Z}$ and $G_i := G * \langle a_i \rangle$ for $i = 1, 2$. By the universal property there exist homomorphisms $\sigma_1\colon G * H \to G_1$ and $\sigma_2\colon G * H \to G_2$ with $(\sigma_i)_{|G} = \mathrm{id}$ and $\sigma_1(h) := a_1 h a_1^{-1}$ as well as $\sigma_2(h) := a_2\varphi(h)a_2^{-1}$ for $h \in H$. By Lemma 12.6, σ_1 and σ_2 are injective. Let $\hat{G}$ be the amalgamated free product of G_1 and G_2 over $G * H$. By Lemma 12.18 we can identify G with $\sigma_1(G) = \sigma_2(G)$ in $\hat{G}$. For $h \in H$ we have $a_1 h a_1^{-1} = \sigma_1(h) = \sigma_2(h) = a_2\varphi(h)a_2^{-1}$ and the claim follows with $x := a_2^{-1}a_1 \in \hat{G}$. □

[33] Using Hall's marriage theorem one can show that a *common* set of representatives for these right cosets exists. This is not needed here.

Remark 12.24. In the situation of Theorem 12.23, one calls $\langle G, x\rangle \leq \hat{G}$ an *HNN extension* of G with respect to $\varphi\colon H \to K$. If G is torsion-free, then so is every HNN extension by Lemma 12.20.

Theorem 12.25. *Every countable group is a subgroup of a group with two generators.*

Proof. Let $G = \{1 = g_0, g_1, \ldots\}$ be a countable group. If G is finite, the claim follows from $G \leq \mathrm{Sym}(G)$. So let $|G| = \infty$. Let $F_2 = \langle x, y\rangle$, $X := \{y^k x y^{-k} : k \geq 0\}$ and

$$Y := \{x^k y x^{-k} g_k : k \geq 0\} \subseteq G * F_2.$$

From Lemma 12.6 it easily follows that $\langle X\rangle \cong F_X \cong F_Y \cong \langle Y\rangle$. Let $\varphi\colon \langle X\rangle \to \langle Y\rangle$, $y^k x y^{-k} \mapsto x^k y x^{-k} g_k$ for $k \geq 0$. Let $\hat{G} := \langle G * F_2, s\rangle$ be an HNN extension of $G * F_2$ with respect to φ. Then $s y^k x y^{-k} s^{-1} = x^k y x^{-k} g_k$ for $k \geq 0$. In particular, $y = sxs^{-1} \in \langle x, s\rangle$ and $g_k = (x^k y x^{-k})^{-1} s y^k x y^{-k} s^{-1} \in \langle x, s\rangle$ for $k \geq 1$. This shows $G \leq \hat{G} = \langle x, s\rangle$. □

Theorem 12.26. *There exists an infinite group with exactly two conjugacy classes.*

Proof. First let $G = G_1 = \{1 = g_0, g_1, g_2, \ldots\}$ be an arbitrary countable torsion-free group (for example $G = \mathbb{Z}$). Inductively, there exists a torsion-free countable HNN extension G_{n+1} of G_n such that g_1 and g_{n+1} are conjugate in G_{n+1}. In the torsion-free countable group $G^* := \bigcup_{n\in\mathbb{N}} G_n$ all g_i are then conjugate. We now set $H_1 := G$ and $H_{n+1} := H_n^*$ for $n \geq 1$. Finally, let $H := \bigcup_{n\geq 1} H_n$. For $x, y \in H \setminus \{1\}$ there exists an $n \geq 1$ with $x, y \in H_n$. By construction, x and y are conjugate in $H_{n+1} \leq H$. Hence H has only two conjugacy classes. □

Remark 12.27. Let G be a torsion group with class number 2. Then every nontrivial element in G has the same order p. Certainly p is a prime. In the case $p > 2$, $x \in G \setminus \{1\}$ cannot be conjugate to x^{-1}. Hence $p = 2$ and G is abelian. It follows that $G \cong C_2$.

13 The Burnside Problem

Remark 13.1. Burnside asked in 1902:

(I) Is every finitely generated torsion group finite?

(II) Is every finitely generated periodic group finite?[34]

(III) For $d, e \in \mathbb{N}$, are there at most finitely many finite groups with d generators and exponent e?

The first question was answered in the negative by Golod in 1964 (Theorem 13.3). Novikov and Adian answered (II) in the negative as well in 1968. Zelmanov, on the other hand, proved in 1990 that (III) is true (he received the Fields Medal for it).[35]

Theorem 13.2. *Every finitely generated solvable torsion group is finite.*

Proof. Let G be a finitely generated solvable torsion group. By the fundamental theorem of finitely generated abelian groups, G/G' is finite. By Theorem 2.11, G' is finitely generated. By induction on the derived length, it follows that G is finite. □

[34] Reminder: Periodic means that the orders of all elements are globally bounded (i.e., $\exp(G) < \infty$).
[35] The proof uses Schreier's conjecture and is therefore based on the CFSG.

Theorem 13.3 (Golod). *For every prime $p > 2$ there exists an infinite p-group with two generators.*[36]

Proof (Gupta).
Step 1: Construction of G.
Let $\langle a\rangle \cong \langle t\rangle \cong C_p$ and $H = \langle a\rangle * \langle t\rangle$. Let $a_i := t^i a t^{-i}$ and

$$A := \langle a\rangle^H = \langle a_0, a_1, \ldots, a_{p-1}\rangle.$$

Since every element of A can be written uniquely in the form $a_{i_1}^{n_1} \ldots a_{i_s}^{n_s}$ with $1 \le n_1, \ldots, n_s < p$ and $i_k \ne i_{k+1}$, we have $A \cong \langle a_0\rangle * \ldots * \langle a_{p-1}\rangle$ (Remark 12.7). Moreover, $H = A \rtimes \langle t\rangle$. For $0 \le k < p$ let $\theta_k \colon A \to H$ be the homomorphism with $\theta_k(a_k) := a$ and $\theta_k(a_i) := t^{i-k}$ for $i \ne k$. Set $N_0 = 1$ and

$$N_{k+1} := \{x \in A : \forall i : \theta_i(x) \in N_k\}.$$

Clearly $1 = N_0 \le N_1 \le \ldots$. We show $N_k \trianglelefteq H$. This is clear for $k = 0$. Let $N_k \trianglelefteq H$ by induction. Then first of all $N_{k+1} \trianglelefteq A$. Let $w = w(a_0, \ldots, a_{p-1}) \in N_{k+1}$. Since $\theta_i(a_j) = \theta_{i-1}(a_{j-1})$, we have

$$\theta_i(twt^{-1}) = \theta_i(w(a_1, \ldots, a_{p-1}, a_0)) = \theta_{i-1}(w(a_0, \ldots, a_{p-1})) \in N_k$$

and $twt^{-1} \in N_{k+1}$. This shows $N_{k+1} \trianglelefteq H$. Hence also $N := \bigcup_{k\ge 0} N_k \trianglelefteq H$. Finally, let

$$G := H/N.$$

Step 2: G is a finitely generated p-group.
Certainly $G = \langle aN, tN\rangle$. For $h = wt^i \in H$ with $w = a_{j_1}^{w_1} \ldots a_{j_s}^{w_s} \in A$ reduced and $0 \le i < p$ let

$$l(h) := \begin{cases} 1 + s & \text{if } i \ne 0, \\ s & \text{if } i = 0. \end{cases}$$

We show $h^{p^{l(h)}} \in N_{l(h)}$ for all $h \in H$. Then G is a p-group. For $l(h) \le 1$ we have $h \in \langle t\rangle \cup \langle a_0\rangle \cup \ldots \cup \langle a_{p-1}\rangle$ and $h^p = 1 \in N_0$. Let $l(h) = n + 1$ and suppose the claim has already been proved for n. First assume $i \ne 0$. Then $l(w) = n$ and

$$h^p = (wt^i)^p = wt^i wt^{-i} t^{2i} wt^{-2i} \ldots t^{(p-1)i} wt^{-(p-1)i}.$$

For $w = w(a_0, \ldots, a_{p-1})$ we have $t^{ji} w t^{-ji} = w(a_{ji}, a_{ji+1}, \ldots, a_{ji+p-1})$, where the indices are to be read modulo p. Let d_k be the number of powers of a_k in w. The number of powers of a_k in h^p is then at most

$$\sum_{j=0}^{p-1} d_{k-ji} = \sum_{j=0}^{p-1} d_j \le n$$

because $p \nmid i$. Hence the number of powers of a in $\theta_k(h^p)$ is also $\le n$. We count how many t occur in $\theta_k(h^p)$. Let s_r be the sum of the exponents of all a_r in w. The contribution of these powers to the exponent sum of t in $\theta_k(h^p)$ is then

$$\sum_{j=0}^{p-1} s_r(r + ij - k) \equiv s_r\Big(p(r-k) + i\binom{p}{2}\Big) \equiv 0 \pmod p,$$

since $p > 2$ (note that nothing is counted for $ij + r = k$). Since $A \trianglelefteq H$, all t in $\theta_k(h^p)$ can be moved to the right without changing the exponent sum. This shows $\theta_k(h^p) \in A$. By moving the t, some of

[36] This also holds for $p = 2$, but with a different proof.

the a are converted into a_j. Their number is, however, still $\leq n$, i. e., $l(\theta_k(h^p)) \leq n$. By induction, $\theta_k(h^{p^{n+1}}) = \theta_k(h^p)^{p^n} \in N_n$ for all k, i. e., $h^{p^{n+1}} \in N_{n+1}$.

Now let $i = 0$, so $h = w \in A$. After conjugation we may assume that w does not begin and end with the same a_j. If $l(\theta_k(w)) \leq n$ already holds for all k, then $\theta_k(w^{p^n}) \in N_n$ by induction, and it follows that $w^{p^n} \in N_{n+1}$ and $w^{p^{n+1}} \in N_{n+1}$. So suppose $l(\theta_k(w)) = n + 1$ for some k (note that the length cannot become larger). If w contains two powers of a_j with $j \neq k$, then these are transformed into powers of t under θ_k. After moving all t in $\theta_k(w)$ to the right, one obtains the contradiction $l(\theta_k(w)) \leq n$. Hence w can contain at most one power of each a_j with $j \neq k$. On the other hand, w begins and ends with different a_j. This shows $n = 1$, and after conjugation w has the form $w = a_k^r a_j^s$ with $j \neq k$. It follows that $\theta_k(w) = a^r t^{j-k}$. As in the case $i \neq 0$, we obtain $\theta_m(\theta_k(w)^p) \in A$ with $l(\theta_m(\theta_k(w)^p)) \leq 1$ for $m = 0, \ldots, p-1$ (at this point we had not yet used induction). Hence $\theta_m(\theta_k(w^{p^2})) = 1$ and $\theta_k(w^{p^2}) \in N_1$. Interchanging the roles of k and j, we obtain $\theta_j(w^{p^2}) \in N_1$. For $m \notin \{k, j\}$ we even have $\theta_m(w^p) = 1$. Hence $w^{p^2} \in N_2$ as claimed.

Step 3: $|G| = \infty$.
Suppose G is finite. By Reidemeister-Schreier, N is then finitely generated. In particular, $N = N_n$ for some $n \in \mathbb{N}$. We define $v_0 := [a_1, a] \in A$ and $v_{k+1} = [v_k, a] \in A$ for $k \geq 0$. Then $\theta_0(v_0) = [t, a] = a_1 a^{-1}$ and $\theta_0(v_1) = [a_1 a^{-1}, a] = v_0$. Inductively, $\theta_0(v_{k+1}) = v_k$ for $k \geq 0$. By Lemma 12.10, all v_i have infinite order, since their reduced form begins with a_1 and ends with a^{-1}. Let $r \geq 0$ with $v_0^r \in N_2$. Then $\theta_0(v_0^r) = (a_1 a^{-1})^r \in N_1$ and one obtains $(ta^{-1})^r = \theta_0(a_1 a^{-1})^r = 1$. This shows $p \mid r$. It follows that

$$1 = \big((ta^{-1})^p\big)^{r/p} = \big((at^{-1})^p\big)^{-r/p} = (a_0 a_{p-1} \ldots a_1)^{-r/p}$$

and $r = 0$. We have thus proved $\langle v_0 \rangle \cap N_2 = 1$. Now let inductively $\langle v_{k-1} \rangle \cap N_{k+1} = 1$ and $v_k^r \in N_{k+2}$. Then $v_{k-1}^r = \theta_0(v_k)^r \in N_{k+1}$ and $r = 0$. Thus $\langle v_k \rangle \cap N_{k+2} = 1$ for all $k \geq 0$. On the other hand, $v_n^{p^{l(v_n)}} \in N = N_n = N_{n+2}$ by Step 2. Contradiction. □

Definition 13.4. For $d, e \in \mathbb{N}$ let $B(d, e) := F_d / \langle g^e : g \in F_d \rangle$ be the *Burnside group* with d generators and exponent e. The Burnside problem (II) is equivalent to $|B(d, e)| < \infty$ for all $d, e \in \mathbb{N}$.

Theorem 13.5. *For $d, e \in \mathbb{N}$ we have*

(i) $B(1, e) \cong C_e$.

(ii) $B(d, 2) \cong C_2^d$.

(iii) $|B(d, 3)| \leq 3^{3^{d-1}}$.

Proof.

(i) Trivial.

(ii) Every group of exponent 2 is abelian. By Burnside's basis theorem, C_2^d is the largest elementary abelian group with d generators.

(iii) Induction on d: The case $d = 1$ follows from (i). Let $d \geq 2$ and $G := B(d, 3) = \langle x_1, \ldots, x_d \rangle$. By induction, $H := \langle x_1, \ldots, x_{d-1} \rangle$ has order $\leq 3^{3^{d-2}}$. Every $g \in G$ has the form $g = h_1 x_d^{\epsilon_1} h_2 \ldots x_d^{\epsilon_k} h_{k+1} \ldots$ with $\epsilon_i \in \{\pm 1\}$ and $h_i \in H$. Since $\exp(G) = 3$, we have

$$x_d^{\pm 1} h x_d^{\pm 1} = h^{-1} x_d^{\mp 1} h^{-1} \qquad\qquad x_d^{\pm 1} h x_d^{\mp 1} = h^{-1} x_d^{\mp 1} h^{-1} x_d^{\pm 1}$$

for all $h \in H$. Using the first equation, two adjacent occurrences of x_d in g can be eliminated if they have the same exponent. Using the second equation, two adjacent x_d can be interchanged. In this way one obtains a representation with $k \leq 2$. It follows that

$$G = Hx_dH \cup Hx_d^{-1}H \cup Hx_dHx_d^{-1}H,$$
$$|G| \leq 2|H|^2 + |H|^3 < 3|H|^3 \leq 3^{3^{d-1}+1}$$

(note that $H \subseteq Hx_dHx_d^{-1}H$). Since G is a 3-group, we even have $|G| \leq 3^{3^{d-1}}$. □

Definition 13.6 (Review from GT)**.** For $x_1, \ldots, x_n \in G$ let $[x_1, x_2] := x_1x_2x_1^{-1}x_2^{-1}$ and

$$[x_1, \ldots, x_n] := [x_1, [x_2, \ldots, x_n]]$$

for $n \geq 3$. We set $G^{[1]} := G$ and $G^{[k]} := [G, G^{[k-1]}]$ for $k \geq 2$.

Lemma 13.7. *For $G = \langle X \rangle$ and $k \geq 1$ we have $G^{[k]} = \langle [x_1, \ldots, x_k] : x_1, \ldots, x_k \in X \rangle G^{[k+1]}$.*

Proof. The claim holds for $k = 1$ if one interprets $[x] = x$. Now suppose the claim has already been proved for some $k \geq 1$. Let $N := \langle [x_1, \ldots, x_{k+1}] : x_1, \ldots, x_{k+1} \in X \rangle G^{[k+2]} \leq G^{[k+1]}$. For $g \in G$ and $x_1, \ldots, x_{k+1} \in X$ we have

$$g[x_1, \ldots, x_{k+1}]g^{-1} = [g, x_1, \ldots, x_{k+1}][x_1, \ldots, x_{k+1}] \in G^{[k+2]}N = N.$$

This implies $N \trianglelefteq G$. Modulo N, each of the generators $x \in X$ of G commutes with the generators $[x_1, \ldots, x_k]$ and $g \in G^{[k+1]}$ of $G^{[k]}$. This shows $G^{[k+1]} = [G, G^{[k]}] \leq N$. □

Lemma 13.8 (Levi)**.** *Let G be a group of exponent 3. Then*

(i) $[x, x, y] = 1$ *for* $x, y \in G$.

(ii) $[x, y, z] = [y, z, x] = [x, z, y]^{-1}$ *for* $x, y, z \in G$.

(iii) G *is nilpotent of class at most* 3.

Proof.

(i) Follows from

$$x \cdot yxy^{-1} = (xy)^2y^{-2} = (xy)^{-1}y = y^{-1}x^{-1}y = y^2(y^{-1}x)^{-1} = y^2(y^{-1}x)^2 = yxy^{-1} \cdot x.$$

(ii) By (i), any two conjugates of x commute. Hence $\langle x \rangle^G$ is abelian. It follows that

$${}^x[y, z][x, z] = [xy, z] = {}^{xy}[xy, z] = {}^{xy}(\underbrace{{}^x[y, z]}_{\in \langle y \rangle^G}[x, z]) = {}^x({}^x[y, z] \cdot {}^y[x, z]) = {}^{x^{-1}}[y, z] \cdot {}^y[x, z].$$

Conjugation with x^{-1} yields

$${}^x[y, z] \cdot {}^y[x, z] = \underbrace{[y, z]}_{\in \langle z \rangle^G}\underbrace{[x, z]}_{\in \langle z \rangle^G} = [x, z][y, z].$$

Now we have

$$[x, y, z] = {}^x[y, z][z, y] = {}^y[z, x][x, z][y, z][z, y] = {}^y[z, x][x, z] = [y, z, x].$$

Since

$$[x, y^{-1}] = [x, yy] = [x, y] \cdot {}^y[x, y] = [x, y]^2 = [x, y]^{-1}$$

we have $[x, y, z]^{-1} = [x, [y, z]^{-1}] = [x, z, y]$.

(iii) From (ii) it follows that $[x,y,z,w] = [x,[y,z,w]] = [x,z,w,y]$ and

$$[x,y,z,w] = [x,y,[z,w]] = [x,[z,w],y]^{-1} = [y,x,[z,w]]^{-1} = [y,x,z,w]^{-1}.$$

Hence for $\pi \in \langle (1,2),(2,3,4)\rangle = S_4$ and $x_1,\dots,x_4 \in G$ we have

$$[x_{\pi(1)},x_{\pi(2)},x_{\pi(3)},x_{\pi(4)}] = [x_1,x_2,x_3,x_4]^{\operatorname{sgn}(\pi)}.$$

On the other hand,

$$\begin{aligned}[x,y,z,w] &= [x,y,[z,w]] = [[z,w],x,y] = [[z,w],[x,y]] = [[x,y],[z,w]]^{-1}\\ &= [[x,y],z,w]^{-1} = [z,w,[x,y]]^{-1} \underset{\pi=(1,3)(2,4)}{=} [x,y,z,w]^{-1}.\end{aligned}$$

This shows $[x,y,z,w] = 1$ for all $x,y,z,w \in G$, and the claim follows. □

Remark 13.9.

(i) As is well known, every group of exponent 2 is abelian, hence nilpotent of class at most 1. For p-groups with exponent $p \geq 5$, on the other hand, there is no longer an absolute bound for the nilpotency class (provided the group is nilpotent).

(ii) For $p > 10^{75}$, Ol'shanskii proved the existence of infinite p-groups such that every nontrivial proper subgroup has order p. One calls these groups *Tarski monsters* (like the sporadic Monster, they are simple groups).

(iii) The next theorem improves Theorem 13.5.

Theorem 13.10 (Levi, Van der Waerden)**.** *For $d \geq 1$ we have* $\boxed{|B(d,3)| = 3^{d(d^2+5)/6}.}$

Proof. Let $G := B(d,3) = \langle x_1,\dots,x_d\rangle$. By von Dyck's theorem, there exists an epimorphism $G \to C_3^d$. Since G/G' is generated by d elements, it follows that $G/G' \cong C_3^d$. By Lemma 13.7, the elementary abelian group $G'/G^{[3]}$ is generated by the elements $[x_i,x_j]G^{[3]}$ with $i<j$. This shows $|G'/G^{[3]}| \leq 3^{\binom{d}{2}}$. By Lemma 13.8, $G^{[3]}$ is also elementary abelian with $[x,y,z] = [y,z,x]$ as well as $[x,y,z] = [x,z,y]^{-1}$. By Lemma 13.7, $G^{[3]}$ is therefore generated by the elements $[x_i,x_j,x_k]$ with $i<j<k$. Thus $|G^{[3]}| \leq 3^{\binom{d}{3}}$ and $|G| \leq 3^k$ with $k = d + \binom{d}{2} + \binom{d}{3} = d(d^2+5)/6$.

Every element in G has the form

$$x_1^{a_1}\dots x_d^{a_d}[x_1,x_2]^{b_{12}}\dots[x_{d-1},x_d]^{b_{d-1,d}}[x_1,x_2,x_3]^{c_{123}}\dots[x_{d-2},x_{d-1},x_d]^{c_{d-2,d-1,d}}$$

with $a_i, b_{ij}, c_{ijk} \in \{0,1,2\}$. Suppose that an element g has two different factorizations of this kind. These factorizations must be equal in G/G', since $|G/G'| = 3^d$. So we may assume $g \in G'$. Since $G'' \subseteq G^{[4]} = 1$ (GT, Lemma 3.11), G' is abelian. By rearranging, one obtains a nontrivial relation of the form

$$[x_1,x_2]^{b_{12}}\dots[x_{d-1},x_d]^{b_{d-1,d}}[x_1,x_2,x_3]^{c_{123}}\dots[x_{d-2},x_{d-1},x_d]^{c_{d-2,d-1,d}} = 1.$$

Let $i<j<k$ with $b_{ij} \neq 0$ or $c_{i,j,k} \neq 0$. By setting all other $x_l = 1$, we obtain a nontrivial relation in $B(3,3)$. If we can show $|B(3,3)| = 3^{3(3^2+5)/6} = 3^7$, we would have the desired contradiction. In GAP one can show $B(3,3) \cong \texttt{SmallGroup}(3^7, 4487)$.[37] Nevertheless, we give a theoretical argument.

[37] For example with $\texttt{OneGroup}(3^7, \texttt{RankPGroup}, 3, \texttt{Exponent}, 3)$;

Let $E_{ij} \in \mathbb{F}_3^{4\times 4}$ with $E_{ij}E_{kl} = \delta_{jk}E_{il}$ as usual. For the lower triangular matrices

$$Y := 1 + E_{31} - E_{42}, \qquad Z := 1 - E_{21} - E_{43} \in \mathrm{GL}(4,3)$$

we have

$$\begin{aligned} (Y-1)^2 = 0 = (Z-1)^2, &\qquad (Y-1)(Z-1) = E_{41} = -(Z-1)(Y-1),\\ E_{41}(Y-1) = 0 = (Y-1)E_{41}, &\qquad E_{41}(Z-1) = 0 = (Z-1)E_{41}. \end{aligned}$$

This shows

$$H := \langle Y, Z\rangle = \{1 + \alpha(Y-1) + \beta(Z-1) + \gamma E_{41} : \alpha,\beta,\gamma \in \mathbb{F}_3\}$$

and $(h-1)^2 = 0$ for all $h \in H$. In particular, $|H| = 27$ and $h^3 = 1 + (h-1)^3 = 1$, i. e., $\exp(H) = 3$. Let $N := \mathbb{F}_3^4$ with standard basis $e_1, \dots, e_4$ and $G := N \rtimes H$, so $|G| = 3^7$. Since $1 + h + h^2 = (h-1)^2 = 0$, we have for $v \in N$ and $h \in H$

$$(vh)^3 = \big((1 + h + h^2)v\big)h^3 = h^3 = 1,$$

i. e., $\exp(G) = 3$. From $[Y, e_1] = e_3$, $[Z, e_1] = -e_2$ and $[Y, e_2] = -e_4$ it follows that $G = \langle e_1, Y, Z\rangle$. Hence $G \cong B(3,3)$. □

Theorem 13.11 (Sanov)**.** *For $d \geq 1$ we have $|B(d,4)| < \infty$.*

Proof. Let $G := B(d,4)$. W. l. o. g. we may assume $d \geq 2$. By induction, $H := \langle x_1, \dots, x_{d-1}\rangle$ is finite. Let $y := x_d^2$. We first show that $K := \langle y, H\rangle$ is finite. The claim then follows by a similar argument for K and x_d in place of H and y. Every $g \in K$ has the form $g = h_1 y h_2 \dots y h_n$ with $h_1, \dots, h_n \in H$. Let n be as small as possible. Suppose $n \geq 2|H| + 4$. Then among the elements $h_2 h_3^{-1}$, $h_2 h_4 h_6 (h_3 h_5)^{-1}$, $h_2 h_4 (h_3 h_5 h_7)^{-1}, \dots$ there are two equal ones, say

$$h_2 h_4 \dots h_{2r} (h_3 h_5 \dots h_{2r+1})^{-1} = h_2 \dots h_{2s} (h_3 \dots h_{2s+1})^{-1}$$

with $r < s \leq (n-2)/2$. Hence

$$h_{2r+2} h_{2r+4} \dots h_{2s} (h_{2r+3} h_{2r+5} \dots h_{2s+1})^{-1} = 1.$$

Since $\exp(G) = 4$ and $y^2 = 1$, we have

$$yhy = h^{-1}y^{-1}h^{-1}y^{-1}h^{-1} = h^{-1}yh^{-1}yh^{-1}$$

for $h \in H$. We apply this rule to $h = h_{2s+1}$ in g:

$$g = h_1 y h_2 \dots h_{2s} h_{2s+1}^{-1} y h'_{2s+1} y \dots y h_n.$$

By applying it again with $h = h_{2s} h_{2s+1}^{-1}$, h_{2s-1} is replaced by $h_{2s-1}(h_{2s}h_{2s+1}^{-1})^{-1} = h_{2s-1}h_{2s+1}h_{2s}^{-1}$. Then h_{2s-2} is replaced by $h_{2s-2}h_{2s}(h_{2s-1}h_{2s+1})^{-1}$, and so on. Finally, h_{2r+2} is replaced by 1. This contradicts the minimality of n. Hence $n \leq 2|H| + 2$ and one obtains $|K| < \infty$. □

Example 13.12. Let $F := F_2$ and $N := \langle x^2 : x \in F\rangle$. Then $|F/N| = 4$ and Schreier's formula shows $N \cong F_5$. Hence $M := \langle x^2 : x \in N\rangle$ has index 32 in N. Since M is characteristic in N, we have $M \trianglelefteq F$. Now $G = F/M$ has order 2^7, exponent 4 and two generators. This shows $|B(2,4)| \geq 2^7$. The same argument yields $|B(2,8)| \geq 2^{7+2^7+1} = 2^{136}$ and $|B(3,4)| \geq 2^{3+17} = 2^{20}$.

Remark 13.13. We have $|B(2,4)| = 2^{12}$, $|B(3,4)| = 2^{69}$, $|B(4,4)| = 2^{422}$ and $|B(5,4)| = 2^{2728}$. Hall proved

$$\boxed{|B(d,6)| = 2^{1+3^a(d-1)}3^{b+\binom{b}{2}+\binom{b}{3}}}$$

with $a = d + \binom{d}{2} + \binom{d}{3}$ and $b = 1 + 2^d(d-1)$. It is known that $|B(d,e)| = \infty$ for all $d \geq 2$ and $e \geq 8000$ (and all odd $e \geq 557$).[38][39]

Definition 13.14.

- Let M be the intersection of all normal subgroups of finite index of $B(d,e)$ and $B_0(d,e) := B(d,e)/M$. The Burnside problem (III) is equivalent to $|B_0(d,e)| < \infty$ for all $d, e \in \mathbb{N}$.
- Let G be a finite group and p a prime. We define $N_0 := 1$ and

$$N_{2k-1}/N_{2k-2} := \mathrm{O}_{p'}(G/N_{2k-2}), \qquad N_{2k}/N_{2k-1} := \mathrm{O}_p(G/N_{2k-1})$$

  for $k \geq 1$. If there exists $k \geq 0$ with $N_k = G$, then G is called *p-solvable*. In this case, the smallest number $l := l_p(G) \geq 0$ with $N_{2l+1} = G$ is called the *p-length* of G.

Remark 13.15.

(i) If G is p-solvable, then the normal series $N_0 < \ldots < N_k$ can be refined to a chief series and a composition series. Hence all chief and composition factors are p-groups or p'-groups. In particular, every minimal normal subgroup of G is a p-group or a p'-group.

(ii) Clearly every solvable group is p-solvable for every prime p.

Lemma 13.16 (Hall-Higman)**.** *Let G be a p-solvable group with $\mathrm{O}_{p'}(G) = 1$. Then*

$$\mathrm{C}_G(\mathrm{O}_p(G)) \leq \mathrm{O}_p(G).$$

Proof. W. l. o. g. let $N := \mathrm{O}_p(G)$. Then $\mathrm{C}_G(N)N/N \trianglelefteq G/N$. In the case $\mathrm{C}_G(N) \not\subseteq N$ there exists a minimal normal subgroup $M/N \trianglelefteq G/N$ with $M \leq \mathrm{C}_G(N)N$. Since $\mathrm{O}_p(G/N) = 1$, M/N is a p'-group (Remark 13.15). By Schur-Zassenhaus, $M = N \rtimes H$. Since $\mathrm{C}_G(N)N/\mathrm{C}_G(N) \cong N/\mathrm{Z}(N)$ is a p-group, we have $H \leq \mathrm{C}_G(N)$ and $M = N \times H$. But then we would have $H \leq \mathrm{O}_{p'}(M) \leq \mathrm{O}_{p'}(G) = 1$. □

Theorem 13.17. *Let G be p-solvable and c the nilpotency class of a Sylow p-subgroup of G. Then $l_p(G) \leq c$.*

Proof. Induction on c: In the case $c = 0$ we have $G = \mathrm{O}_{p'}(G)$ and $l_p(G) = 0$. Now let $c > 0$ and $P \in \mathrm{Syl}_p(G)$. W. l. o. g. let $\mathrm{O}_{p'}(G) = 1$ and $N := \mathrm{O}_p(G) > 1$. By Hall-Higman, $\mathrm{Z}(P) \leq \mathrm{C}_G(P) \leq \mathrm{C}_G(N) \leq N$. The Sylow p-subgroup

$$PN/N \cong P/P \cap N \cong (P/\mathrm{Z}(P))/((N \cap P)/\mathrm{Z}(P))$$

of G/N thus has nilpotency class $\leq c - 1$. By induction, $l_p(G) = l_p(G/N) + 1 \leq c$. □

Theorem 13.18. *For $d \geq 1$ we have $|B_0(d,6)| \leq |B_0(d,12)| < \infty$.*

[38] See `https://arxiv.org/abs/2303.15997v5`

[39] Magnus commented on the first paper of this kind as "This paper is possibly the most difficult paper to read that has ever been written on mathematics."

Proof. Since $B_0(d,6)$ is a quotient group of $B_0(d,12)$, it suffices to show $|B_0(d,12)| < \infty$. Let G be a finite group with d generators and exponent 12. By Burnside's p^aq^b-theorem, G is solvable. A Sylow 3-subgroup P of G has exponent ≤ 3 and nilpotency class ≤ 3 by Lemma 13.8. By Theorem 13.17, $l_3(G) \leq 3$. So let $1 = N_0 \leq \ldots \leq N_7 = G$ as in Definition 13.14. Then G/N_6 is a 2-group with d generators and exponent ≤ 4. By Theorem 13.11, $|G/N_6|$ is bounded by a function of d. By Reidemeister-Schreier, the number of generators of N_6 is bounded by a function of d. Moreover, N_6/N_5 has exponent ≤ 3. By Theorem 13.10, $|N_6/N_5|$ is also bounded by a function of d. Continuing in this way, one obtains an upper bound for $|G|$. □

Remark 13.19. Hall and Higman showed that the p-length of G can be bounded by a function of the exponent of a Sylow p-subgroup of G. With the help of the CFSG, one can now reduce the question $|B_0(d,n)| < \infty$ to the case where $n = p^k$ is a prime power. Finally, Zelmanov proved $|B_0(d,p^k)| < \infty$.

Definition 13.20. Let $G = \langle x_1, \ldots, x_n \rangle$. We define a set of *basic commutators* $c_1, c_2, \ldots$ recursively as follows:

- For $i = 1, \ldots, n$ let $c_i := x_i$ be the basic commutators of *weight* $\omega(c_i) := 1$.
- Suppose that the basic commutators $c_1, \ldots, c_k$ have already been defined. For $1 \leq i < j \leq k$, $c := [c_i, c_j]$ is a basic commutator if either $\omega(c_j) = 1$ or $c_j = [c_s, c_t]$ with $i \geq s$. In this case, let $\omega(c) := \omega(c_i) + \omega(c_j)$.
- The basic commutators are numbered in ascending order of weight, where the order of the commutators of the same weight does not matter.

For $k \in \mathbb{N}$ let $\delta_n(k)$ be the (finite) number of basic commutators of weight k. Here, basic commutators count as distinct even if they represent the same group elements (e. g., if G is abelian). We show in Remark 13.33 that in the free group $G = F_n$ basic commutators are indeed pairwise distinct.

Example 13.21. Clearly, $\delta_n(1) = n$ and $\delta_n(2) = \binom{n}{2}$. For $n = 3$ and $(x_1, x_2, x_3) = (x, y, z)$,

$$x, y, z, [x,y], [x,z], [y,z], [x,x,y], [x,x,z], [y,x,y], [y,x,z], [y,y,z], [z,x,y], [z,x,z], [z,y,z]$$

are the basic commutators of weight ≤ 3. Hence $\delta_3(3) = 8$.

Theorem 13.22. *Let $G = \langle x_1, \ldots, x_n \rangle$ and $k \in \mathbb{N}$. Then $G^{[k]}/G^{[k+1]}$ is generated by the cosets of the basic commutators of weight k.*

Proof. For $k = 1$ the claim is clear since $G^{[1]} = G$. Suppose that the statement has already been proved for all smaller values of k. By definition, $G^{[k]}$ is generated by the commutators $[g,h]$ with $g \in G$ and $h \in G^{[k-1]}$. Since $G^{[k]}/G^{[k+1]} \leq \mathrm{Z}(G/G^{[k+1]})$, we have

$$\begin{aligned} [xy,h] &= {}^x[y,h][x,h] \equiv [x,h][y,h] \pmod{G^{[k+1]}}, \\ [x^{-1},h] &\equiv {}^x[x^{-1},h] = [h,x] = [x,h]^{-1} \pmod{G^{[k+1]}} \end{aligned}$$

for all $x, y \in G$. Hence we may assume $g = x_i = c_i$ with $1 \leq i \leq n$. By induction there exist basic commutators $c_{j_1}, \ldots, c_{j_l}$ of weight $k-1$ and $h' \in G^{[k]}$ with $h = c_{j_1}^{\pm 1} \ldots c_{j_l}^{\pm 1} h'$. Now similarly

$$[g,h] \equiv [g,c_{j_1}]^{\pm 1} \ldots [g,c_{j_l}]^{\pm 1} [g,h'] \equiv [g,c_{j_1}]^{\pm 1} \ldots [g,c_{j_l}]^{\pm 1} \pmod{G^{[k+1]}}.$$

Thus we may assume $h = c_j$. Since $[c_j, c_i] = [c_i, c_j]^{-1}$, it suffices to show the following additional claim:

> If c_i, c_j are basic commutators with $i < j$ and $\omega(c_i) + \omega(c_j) = k$, then $[c_i, c_j]$ is a product of basic commutators of weight k and their inverses modulo $G^{[k+1]}$.

For this we use induction on $(\omega(c_j), -i)$ with respect to the lexicographic order (note that i is bounded above by a function of $\omega(c_j)$). In the case $\omega(c_j) = 1$, $[c_i, c_j]$ is a basic commutator. So let $c_j = [c_s, c_t]$ with $s < t$. In the case $i \geq s$, $[c_i, c_j]$ is again a basic commutator. So let $i < s < t$. The Hall-Witt identity (GT, Lemma 3.6) simplifies to

$$1 = {}^{c_s}[c_i, c_s^{-1}, c_t] \cdot {}^{c_t}[c_s, c_t^{-1}, c_i] \cdot {}^{c_i}[c_t, c_i^{-1}, c_s] \equiv [c_i, c_s^{-1}, c_t][c_s, c_t^{-1}, c_i][c_t, c_i^{-1}, c_s] \pmod{G^{[k+1]}}.$$

There exists $z \in G^{[k]}$ with $[c_s^{-1}, c_t] = [c_s, c_t]^{-1} z$ and

$$[c_i, c_s^{-1}, c_t] \equiv [c_i, [c_s, c_t]^{-1}] \equiv [c_i, c_j]^{-1} \pmod{G^{[k+1]}}.$$

Since $[G^{[a]}, G^{[b]}] \leq G^{[a+b]}$ (GT, Lemma 3.11), we similarly obtain

$$[c_s, c_t^{-1}, c_i] \equiv [c_s, [c_t, c_i]^{-1}] \equiv [c_s, c_i, c_t] \pmod{G^{[k+1]}}$$

and $[c_t, c_i^{-1}, c_s] \equiv [c_t, c_i, c_s]^{-1} \pmod{G^{[k+1]}}$. Altogether,

$$[c_i, c_j] \equiv [c_s, c_i, c_t][c_t, c_i, c_s]^{-1} \pmod{G^{[k+1]}}.$$

Since $\omega(c_i) + \omega(c_t) < k$ and $\omega(c_i) + \omega(c_s) < k$, $[c_i, c_t]$ and $[c_i, c_s]$ are products of basic commutators modulo $G^{[\omega(c_i)+\omega(c_t)+1]}$ and $G^{[\omega(c_i)+\omega(c_s)+1]}$, respectively (induction on k). As above, it therefore suffices to treat $[c_s, c_u]$ and $[c_t, c_v]$ for basic commutators c_u, c_v with $\omega(c_u) = \omega(c_i) + \omega(c_t)$ and $\omega(c_v) = \omega(c_i) + \omega(c_s)$. Since $\omega(c_u) > \omega(c_t) \geq \omega(c_s)$, we have $s < u$. From $\omega(c_i) \leq \omega(c_s)$ we obtain $\omega(c_u) \leq \omega(c_j)$, and because of $i < s$ we get $(\omega(c_u), -s) < (\omega(c_j), -i)$. Hence we can apply the second induction to $[c_s, c_u]$.

Similarly, $(\omega(c_v), -t) < (\omega(c_j), -i)$ because of $\omega(c_i) \leq \omega(c_t)$ and $i < t$. In the case $t < v$ we are done. In the case $t = v$ we have $[c_t, c_v] = 1$, and in the case $v < t$ we have $[c_t, c_v] = [c_v, c_t]^{-1}$ with $(\omega(c_t), -v) < (\omega(c_j), -i)$ because of $\omega(c_t) < \omega(c_j)$. In all cases, $[c_s, c_u]$ and $[c_t, c_v]$ are products of basic commutators of weight k and their inverses modulo $G^{[k+1]}$. □

Remark 13.23. One can show that every element of $G^{[k]}$ is a product of basic commutators of weight $\geq k$ (this does not follow directly from Theorem 13.22).

Corollary 13.24. *Every nilpotent group that is generated by finitely many elements of finite order is finite.*

Proof. Let $G = \langle x_1, \ldots, x_n \rangle$ be nilpotent with elements $x_1, \ldots, x_n$ of finite order. Then $|G/G'| < \infty$. Inductively, we may assume $e := |G/G^{[k]}| < \infty$. W. l. o. g. we may assume $G^{[k+1]} = 1$. By Theorem 13.22, $G^{[k]}$ is generated by finitely many commutators. From GT, Exercise 18 we obtain

$$[x, y]^e = [x^e, y] \in G^{[k+1]} = 1$$

for all $[x, y] \in G^{[k]} \leq \mathrm{Z}(G)$. Hence $G^{[k]}$ is also finite. □

Remark 13.25. The proof shows: If a nilpotent group G is generated by finitely many p-elements, then G is a p-group. For solvable groups the statement is false by Exercise 14 (cf. Theorem 13.2).

Lemma 13.26. *Let $c_1, c_2, \ldots$ be the basic commutators of $G = \langle x_1, \ldots, x_n \rangle$. For all $k \geq 1$ there are exactly n^k sequences $(a_1, a_2, \ldots) \in \mathbb{N}_0^\infty$ with $a_1 \geq a_2 \geq \ldots$ and $\sum_{a_i \geq 1} \omega(c_{a_i}) = k$.*

Proof. For $s \in \mathbb{N}$ let F_s be the set of all sequences $(a_1, a_2, \ldots) \in \mathbb{N}_0^\infty$ such that:

(i) There exists $l \in \mathbb{N}_0$ with $a_1, \ldots, a_{l-1} \geq s$, $a_l > s$, $s \geq a_{l+1} \geq a_{l+2} \geq \ldots$.

(ii) If $c_{a_m} = [c_i, c_j]$ for some $m \leq l$, then $i < s$.

(iii) $\sum_{a_i \geq 1} \omega(c_{a_i}) = k$.

If s is large enough, then $\omega(c_s) > k$. In this case, $l = 0$ in (i) and (ii) gives no restrictions. Hence F_s consists exactly of the sequences we want to count. For $s = 1$ we have $\omega(c_{a_i}) = 1$ for all $i \in \mathbb{N}$ with $a_i \geq 1$. By (iii), we then have $a_k > 0 = a_{k+1} = a_{k+2} = \ldots$. Otherwise, there are no further restrictions on the sequence. Since $a_i \in \{1, \ldots, n\}$, we obtain $|F_1| = n^k$.

For a given $s \geq 1$ it suffices to construct a bijection $\varphi\colon F_s \to F_{s+1}$. Let $a := (a_i) \in F_s$ and l as in (i). We traverse the sequence from right to left, starting at a_l. If we encounter a pair $(a_i, a_{i+1}) = (s, t)$ with $t > s$, then $c_j := [c_s, c_t]$ is a basic commutator by (ii). In this case, we replace (a_i, a_{i+1}) by j. This does not change (iii). Then we consider the pair (a_{i-1}, j) and iterate. We denote the sequence constructed in this way by $\varphi(a)$. By construction, $\varphi(a)_i \geq s+1$ for $i = 1, \ldots, l$. In the case $\varphi(a)_1 = \ldots = \varphi(a)_l = s+1$, (i) holds with $l = 0$ with respect to $s + 1$. Otherwise, (i) holds with $l := \max\{i : \varphi(a)_i > s + 1\}$. By construction, $\varphi(a)$ also satisfies (ii). Hence $\varphi(a) \in F_{s+1}$.

Conversely, let $(b_i)_i \in F_{s+1}$ be given with l as in (i). We traverse b from left to right, starting at b_1. If $c_{b_i} = [c_s, c_t]$ for some $i \leq l$, then we replace b_i by the pair (s, t). Then we consider t or c_t, respectively. Because of $\sum_{b_j \geq 1} \omega(c_{b_j}) = k$, this process must terminate after finitely many steps. We denote the final result by $\psi(b)$. Clearly, $\psi(b) \in F_s$, and the maps φ and ψ are mutually inverse bijections. □

Lemma 13.27. *For $k \in \mathbb{N}$ we have*

$$\sum_{d|k} \delta_n(d) d = n^k.$$

Proof. Let S_k be the set of sequences $(a_1, a_2, \ldots) \in \mathbb{N}_0^\infty$ from Lemma 13.26. If a basic commutator c_j occurs exactly t times in $(a_i)_i \in S_k$, then it contributes $t\omega(c_j)$ to the sum $\sum_{a_i \geq 1} \omega(c_{a_i}) = k$. Let K_w be the set of basic commutators of weight w. From Lemma 13.26 we obtain an identity of formal power series:

$$\begin{aligned}\prod_{w=1}^{\infty} (1 - X^w)^{-\delta_n(w)} &= \prod_{w=1}^{\infty} (1 + X^w + X^{2w} + \ldots)^{\delta_n(w)} = \prod_{w=1}^{\infty} \prod_{c \in K_w} (1 + X^{\omega(c)} + X^{2\omega(c)} + \ldots) \\ &= \prod_{i=1}^{\infty} (1 + X^{\omega(c_i)} + X^{2\omega(c_i)} + \ldots) = 1 + \sum_{k=1}^{\infty} \sum_{(a_i) \in S_k} X^k = \sum_{k=0}^{\infty} n^k X^k = (1 - nX)^{-1}.\end{aligned}$$

We apply the formal logarithm $\log(1 - X) := -\sum_{k=1}^{\infty} \frac{1}{k} X^k$ to both sides:[40]

$$\sum_{k=1}^{\infty} \frac{n^k}{k} X^k = -\log(1 - nX) = -\sum_{w=1}^{\infty} \delta_n(w) \log(1 - X^w) = \sum_{w=1}^{\infty} \sum_{k=1}^{\infty} \frac{\delta_n(w)}{k} X^{wk} = \sum_{k=1}^{\infty} \sum_{d|k} \frac{d\delta_n(d)}{k} X^k.$$

Comparing coefficients shows the claim. □

[40] See the Combinatorics lecture notes

Theorem 13.28 (Witt's formula)**.** *Let $G = \langle x_1, \dots, x_n \rangle$ and $k \in \mathbb{N}$. Then $G^{[k]}/G^{[k+1]}$ is generated by*

$$\boxed{\delta_n(k) = \frac{1}{k} \sum_{d|k} \mu(d) n^{k/d}}$$

elements. Here, μ is the Möbius function from number theory.

Proof. By Theorem 13.22, $G^{[k]}/G^{[k+1]}$ is generated by the $\delta_n(k)$ basic commutators of weight k. The formula for $\delta_n(k)$ follows from Lemma 13.27 by Möbius inversion. □

Remark 13.29. Next, we show that the estimate for the number of generators of $G^{[k]}/G^{[k+1]}$ in Theorem 13.28 is optimal for the free group.

Definition 13.30.

- Let $M := \langle X_1, \dots, X_n \rangle$ be the monoid of all monomials $X_{i_1} \dots X_{i_k}$ with $k \geq 0$ and $1 \leq i_1, \dots, i_k \leq n$ in the noncommuting variables $X_1, \dots, X_n$ under concatenation (so the free group without inverses).
- Let R be a commutative ring and $A := RM$ the free R-module with basis M. Extending the concatenation distributively turns A into a ring. For example,

  $$(1 + X_2 X_1^2)(X_1 + X_3 X_2) = X_1 + X_2 X_1^3 + X_3 X_2 + X_2 X_1^2 X_3 X_2.$$

  One calls A the *free algebra* over $X_1, \dots, X_n$.
- Let $\deg(X_{i_1} \dots X_{i_k}) := k$ be the *degree* of a monomial. Let A_k be the free R-submodule of A generated by the monomials of degree k, i. e., A_k consists of all *homogeneous polynomials* of degree k (and 0). Clearly, $\operatorname{rk}(A_k) = n^k$ and $A = \bigoplus_{k \geq 0} A_k$.
- The *commutator* of $a, b \in A$ is $[a, b] := ab - ba$. We define the basic commutators in A analogously to Definition 13.20. Instead of c_i we write C_i. For $n = 2$ and $(X_1, X_2) = (X, Y)$, for example,

  $$C_4 = [C_1, C_3] = X[X, Y] - [X, Y]X = X(XY - YX) - (XY - YX)X = X^2 Y - 2XYX + YX^2$$

  (cf. Example 13.21).

Lemma 13.31. *For all $k \in \mathbb{N}$,*

$$B_k := \left\{ C_{a_1} \dots C_{a_m} : m \in \mathbb{N},\ a_1 \geq \dots \geq a_m,\ \sum_{i=1}^{m} \omega(c_{a_i}) = k \right\}$$

is a basis of A_k. In particular, the basic commutators of weight k are linearly independent over $\mathbb{Z}$.

Proof. One easily sees inductively that a basic commutator of weight k lies in A_k. Hence $B_k \subseteq A_k$. We first regard B_k as a multiset (the second part of the proof shows that the elements of B_k are pairwise distinct). By Lemma 13.26, $|B_k| = n^k = \operatorname{rk}(A_k)$. We first show that B_k is a generating set of A_k. For this we consider the sets F_s from the proof of Lemma 13.26. If s is large enough, then B_k consists exactly of the products $C_{a_1} \dots C_{a_m}$ with $a = (a_i)_i \in F_s$ (zeros are removed from the sequence). On the other hand, F_1 corresponds exactly to the monomials $X_{i_1} \dots X_{i_k}$. Hence it suffices to show that $C := C_{a_1} \dots C_{a_m}$ with $a \in F_s$ can be expressed as a linear combination of $C_{b_1} \dots C_{b_{m'}}$ with $b \in F_{s+1}$.

As in the proof of Lemma 13.26, we traverse the sequence a from right to left. If $(a_i, a_{i+1}) = (s, t)$ with $t > s$, we replace C_sC_t in C by the equivalent expression $[C_s, C_t] + C_tC_s$. With $C_j = [C_s, C_t]$ we have

$$C = C_{a_1} \dots C_{a_{i-1}} C_j C_{a_{i+2}} \dots C_{a_m} + C_{a_1} \dots C_{a_{i+1}} C_{a_i} \dots C_{a_m}.$$

We proceed with the first summand as before. In the second summand, $C_{a_i} = C_s$ has moved to the right. The corresponding sequence still lies in F_s and can be treated like the initial sequence. After finitely many steps we reach a linear combination of $C_{b_1} \dots C_{b_{m'}}$ with $b \in F_{s+1}$.

Hence B_k is a generating set of A_k.[41] Therefore there exists a matrix $Q \in R^{n^k \times n^k}$ such that the map $\varphi\colon R^{n^k} \to A_k \cong R^{n^k}$, $v \mapsto Qv$ is surjective. For the i-th standard basis vector $e_i \in R^{n^k}$ there exists $x_i \in R^{n^k}$ with $Qx_i = e_i$. Let $L \in R^{n^k \times n^k}$ be the matrix with columns $x_1, \dots, x_{n^k}$. Then $QL = 1_{n^k}$ and $\det(Q) \in R^\times$. For the adjugate matrix Q^* of Q we obtain $\det(Q)^{-1}Q^*Q = 1_{n^k}$. This shows that φ is injective. Thus B_k is linearly independent. □

Definition 13.32. As before, let $A := R\langle X_1, \dots, X_n\rangle$. For $k \in \mathbb{N}_0$ let $A_{>k} := \sum_{l>k} A_l$ be the ideal of A generated by A_{k+1}. For fixed k we use the abbreviations $I := A_{>k}$, $\overline{A} := A/I$ and $\mathcal{A} := \overline{1 + A_{>0}} \subseteq \overline{A}$.

Remark 13.33. For $a \in A_{>l}$ we have

$$(1 + a) \sum_{i=0}^{k} (-a)^i = 1 + (-1)^k a^{k+1} \in 1 + I.$$

This shows $\overline{1 + A_{>l}} \leq \overline{A}^\times$. In particular, $\mathcal{A} \leq \overline{A}^\times$.

Theorem 13.34. *Let F be the free group of rank n and $k \in \mathbb{N}$. Then $F^{[k]}/F^{[k+1]}$ is a free abelian group of rank $\delta_n(k)$.*

Proof. Let F be free with respect to $x_1, \dots, x_n$. We consider $A = \mathbb{Z}\langle X_1, \dots, X_n\rangle$ with $R = \mathbb{Z}$. By the universal property of F there exists a homomorphism $\varphi\colon F \to \mathcal{A}$ with $\varphi(x_i) = 1 + X_i + I$. For $l \in \mathbb{N}_0$ let $D_l := \varphi^{-1}(\overline{1 + A_{>l}}) \leq F$ (Remark 13.33). Let $r, s \in \mathbb{N}_0$, $a = 1 + u \in \overline{1 + A_{>r}}$ and $b = 1 + v \in \overline{1 + A_{>s}}$. Since $[a, b] = ab - ba = uv - vu \in \overline{A_{>r+s+1}}$, we have

$$aba^{-1}b^{-1} = 1 + (ab - ba)a^{-1}b^{-1} = 1 + [a, b]a^{-1}b^{-1} \in \overline{1 + A_{>r+s+1}}$$

and hence $[D_r, D_s] \leq D_{r+s+1}$. Since $\varphi(F) \subseteq \mathcal{A} = \overline{1 + A_{>0}}$, we have $F = D_0$. Inductively, it follows that

$$F^{[l+1]} = [F, F^{[l]}] \leq [D_0, D_{l-1}] \leq D_l$$

for all $l \geq 1$. In particular, $F^{[k+1]} \leq D_k = \mathrm{Ker}(\varphi)$.

Next we show

$$\varphi(c_i) \equiv 1 + C_i \pmod{\overline{A_{>\omega(c_i)}}}$$

for every basic commutator c_i. This is clear for $\omega(c_i) = 1$. So let $[c_i, c_j]$ be a basic commutator with $w_i := \omega(c_i)$ and $w_j := \omega(c_j)$. By induction there exist $\epsilon_i \in \overline{A_{>w_i}}$ and $\epsilon_j \in \overline{A_{>w_j}}$ with $a := \varphi(c_i) = 1 + C_i + \epsilon_i$ and $b := \varphi(c_j) = 1 + C_j + \epsilon_j$. As above,

$$\varphi([c_i, c_j]) = 1 + [a, b]a^{-1}b^{-1} = 1 + [C_i + \epsilon_i, C_j + \epsilon_j]a^{-1}b^{-1} \equiv 1 + [C_i, C_j] \pmod{\overline{A_{>w_i+w_j}}}$$

as claimed.

[41] Since R is not a field in general, linear independence does not follow automatically.

Let $c_{i_1}, \ldots, c_{i_m}$ be the basic commutators of F of weight k. By Theorem 13.22, $F^{[k]}/F^{[k+1]}$ is generated by $c_{i_1}F^{[k+1]}, \ldots, c_{i_m}F^{[k+1]}$. Let $a_1, \ldots, a_m \in \mathbb{Z}$ with $g := c_{i_1}^{a_1} \ldots c_{i_m}^{a_m} \in F^{[k+1]}$. Then

$$1 = \varphi(g) = (1 + C_{i_1})^{a_1} \ldots (1 + C_{i_m})^{a_m} = 1 + a_1 C_{i_1} + \ldots + a_m C_{i_m}.$$

Hence $a_1 C_{i_1} + \ldots + a_m C_{i_m} \in A_k \cap I = 0$. From Lemma 13.31 we obtain $a_1 = \ldots = a_m = 0$. Thus $c_{i_1}F^{[k+1]}, \ldots, c_{i_m}F^{[k+1]}$ is a basis of $F^{[k]}/F^{[k+1]}$. □

Remark 13.35. By Remark 2.10, F_2' is free of (countably) infinite rank. Hence F_2'/F_2'' is not finitely generated.

Corollary 13.36. *The free nilpotent group $F_n/F_n^{[k+1]}$ of rank n and nilpotency class k is torsion-free.*

Theorem 13.37. *For every prime power $q = p^e$ and $n \geq 2$ we have:*

(i) There exists a unique group P with n generators, exponent q, nilpotency class $p-1$ and $P^{[k]}/P^{[k+1]} \cong C_q^{\delta_n(k)}$ for $k = 1, \ldots, p-1$.

(ii) There exists a group P with n generators, exponent q and nilpotency class $q-1$ such that $P^{[k]}/P^{[k+1]}$ has rank $\delta_n(k)$ for $k = 1, \ldots, q-1$.

Proof. Let F be the free group with generators $x_1, \ldots, x_n$. Let $N := \langle g^q : g \in F \rangle \trianglelefteq F$ and $\overline{F} := F/N$.

(i) We consider $A := R\langle X_1, \ldots, X_n \rangle$ with $R = \mathbb{Z}/q\mathbb{Z}$. In the definition of $\mathcal{A}$ let $k < p$. For $1 + a \in \mathcal{A}$ we then have

$$(1 + a)^q = 1 + qa + \binom{q}{2}a^2 + \ldots + \binom{q}{p}a^p + \ldots = 1.$$

Hence for the map $\varphi\colon F \to \mathcal{A}$, $x_i \mapsto 1 + X_i + I$ from the proof of Theorem 13.34 we have $\varphi(N) = 1$. The cosets of the basic commutators $c_{i_1}, \ldots, c_{i_m} \in F$ of weight k form a generating set of $\overline{F}^{[k]}/\overline{F}^{[k+1]}$ as before. Let $a_1, \ldots, a_m \in R$ with $g := c_{i_1}^{a_1} \ldots c_{i_m}^{a_m} \in NF^{[k+1]}$. Then, as before, $\varphi(g) = 1$ and $a_1 = \ldots = a_m = 0$. This shows that $\overline{F}^{[k]}/\overline{F}^{[k+1]} \cong F^{[k]}N/F^{[k+1]}N$ is a free R-module of rank $\delta_n(k)$. Since $n \geq 2$, we have $\delta_n(k) > 0$. Hence $P := \overline{F}/\overline{F}^{[p]} \cong F/NF^{[p]}$ has the desired properties.

Conversely, let Q be a group with the given properties. Then there exists an epimorphism $\sigma\colon F \to Q$ with $NF^{[p]} \leq \mathrm{Ker}(\sigma)$. In particular, $|P| \geq |Q|$. From

$$|Q| = q^{\delta_n(1)+\ldots+\delta_n(p-1)} = |P|$$

we obtain $Q \cong P$.

(ii) This time we consider $A = \mathbb{F}_p\langle X_1, \ldots, X_n \rangle$ and $k < q$ in the definition of $\mathcal{A}$. For $1 + a \in \mathcal{A}$ we still have

$$(1 + a)^q = \sum_{k=0}^{q} \binom{q}{k} a^k = 1$$

because of $\binom{q}{k} \equiv 0 \pmod{p}$ for $k = 1, \ldots, q-1$. So we again obtain a homomorphism $\varphi\colon F \to \mathcal{A}$ with $\varphi(N) = 1$. Since the basic commutators of weight k are linearly independent over $\mathbb{F}_p$, $\overline{F}^{[k]}/\overline{F}^{[k+1]}$ has rank $\delta_n(k)$ (but not necessarily exponent q). To show that $P := \overline{F}/\overline{F}^{[q]} \cong F/NF^{[q]}$ has exponent q, we consider an epimorphism $\psi\colon F \to C_q^n$. Since $\psi(NF') = 1$, we have $P/P' \cong F/F'N \cong C_q^n$. □

Remark 13.38.

(i) Theorem 13.37 shows $\lim_{p\to\infty} \log_p |B(n,p)| = \infty$ for $n \geq 2$. For the group P from Theorem 13.37(i) we have more precisely

$$\log_q |P| = \sum_{k=1}^{p-1} \delta_n(k) = \sum_{k=1}^{p-1} \frac{1}{k} \sum_{d|k} \mu(d) n^{k/d} = \sum_{d=1}^{q-1} \frac{\mu(d)}{d} \sum_{s=1}^{\lfloor (p-1)/d \rfloor} \frac{n^s}{s}.$$

For $q \in \{3,5\}$ we obtain $|B(n,3)| \geq 3^{n+\binom{n}{2}} = 3^{\binom{n+1}{2}}$ (cf. Theorem 13.10) and $|B(2,5)| \geq 5^{2+1+2+3} = 5^8$. In fact, $|B_0(2,5)| = 5^{34}$ and $|B_0(3,5)| = 5^{2282}$ (without proof).

(ii) As is well known, every group of exponent 2 is abelian, i. e., the nilpotency class is ≤ 1. By Lemma 13.8, groups of exponent 3 can have nilpotency class at most 3 (by Theorem 13.10 this case also occurs). The groups $B_0(2,5)$ and $B_0(3,5)$ have nilpotency class 12 and 17, respectively. For $p \geq 5$ it is not known whether the nilpotency class of a group of exponent p is bounded.

(iii) Let $p = q > 2$ and P the group from Theorem 13.37(i). Then $P/P^{[3]}$ is a maximal Schur extension of C_p^n (Example 5.21).

(iv) In Theorem 13.37(ii), $P^{[k]}/P^{[k+1]}$ does not necessarily have exponent q. Indeed, there is no group with two generators, order 2^6, exponent 4 and nilpotency class 2:

```
OneGroup(2^6,RankPGroup,2,Exponent,4,NilpotencyClassOfGroup,2);
```

(v) Uniquely determined groups with certain properties as in Theorem 13.37 will be constructed more generally in Theorem 14.29.

14 Group Classes and Varieties

Definition 14.1. A class $\mathcal{X}$ of groups is called a *class of groups* if $1 \in \mathcal{X}$ and $G \cong H \in \mathcal{X} \Rightarrow G \in \mathcal{X}$. The elements of $\mathcal{X}$ are called $\mathcal{X}$-groups or groups with property $\mathcal{X}$. A group G is called a *residually $\mathcal{X}$ group* if for every $g \in G \setminus \{1\}$ there exists $N \trianglelefteq G$ with $g \notin N$ and $G/N \in \mathcal{X}$. The residually $\mathcal{X}$ groups form the class of groups $\mathcal{X}^r$.

Remark 14.2. Clearly, $G \in \mathcal{X}^r$ if and only if $\bigcap_{\substack{N\trianglelefteq G \\ G/N\in\mathcal{X}}} N = 1$. In this case, G is isomorphic to a subgroup of $\times_{\substack{N\trianglelefteq G \\ G/N\in\mathcal{X}}} G/N$. If $\{G_i : i \in I\}$ is an arbitrary family of groups, then one calls $H \leq \times_{i\in I} G_i$ a *subdirect product* if the projection of H onto every G_i is surjective.

Theorem 14.3. *Let $\mathcal{X}$ be a class of groups. Then $\mathcal{X}^r$ consists precisely of the subdirect products of $\mathcal{X}$-groups.*

Proof. By Remark 14.2, every residually $\mathcal{X}$ group is a subdirect product of $\mathcal{X}$-groups. Conversely, let $H \leq \times_{i\in I} G_i$ be a subdirect product of $\mathcal{X}$-groups G_i. Let $\pi_i \colon H \to G_i$ be the projection and $K_i := \operatorname{Ker}(\pi_i) \trianglelefteq H$. By assumption, $H/K_i \cong G_i \in \mathcal{X}$ and $\bigcap_{i\in I} K_i = 1$. □

Example 14.4. If $\mathcal{X}$ is the class of abelian groups, then $\mathcal{X}^r = \mathcal{X}$.

Theorem 14.5 (Iwasawa)**.** *Every free group is a residually finite p-group for every prime p.*

Proof. Let $a = x_1^{a_1} \dots x_n^{a_n} \in F_X \setminus \{1\}$ with $x_i \neq x_{i+1}$ and $a_1, \dots, a_n \in \mathbb{Z} \setminus \{0\}$ as in Corollary 1.9. Let q be a power of p that does not divide $a_1 \dots a_n$. Let $E_{st} \in \mathbb{Z}/q\mathbb{Z}^{(n+1)\times(n+1)}$ be the matrix with a 1 at position (s,t) and zeros elsewhere. For $x \in X$ let

$$\sigma(x) := \prod_{\substack{1 \leq i \leq n \\ x_i = x}} (1_{n+1} + E_{i,i+1}) \in \mathrm{GL}(n+1, \mathbb{Z}/q\mathbb{Z}).$$

Clearly, $G := \langle \sigma(x) : x \in X \rangle$ consists of upper triangular matrices with ones on the main diagonal. Hence G is a finite p-group. By the universal property, σ extends to a homomorphism $\widehat{\sigma} \colon F \to G$. Because of $E_{st}E_{uv} = \delta_{tu}E_{sv}$ and $x_i \neq x_{i+1}$, the factors of $\sigma(x)$ commute pairwise. This shows

$$\sigma(x_i)^{a_i} = \prod_{x_j = x_i} (1_{n+1} + E_{j,j+1})^{a_i} = \prod_{x_j = x_i} (1_{n+1} + a_i E_{j,j+1}) = 1_{n+1} + a_i \sum_{x_j = x_i} E_{j,j+1}.$$

It follows that

$$\widehat{\sigma}(a) = \sigma(x_1)^{a_1} \dots \sigma(x_n)^{a_n} = 1_{n+1} + a_1 \dots a_n E_{1,n+1} + \dots \neq 1_{n+1}. \qquad \square$$

Corollary 14.6 (Magnus)**.** *Every free group is residually nilpotent.*

Remark 14.7. For every nonabelian free group F we have $F^{[\infty]} = \bigcap_{k=1}^{\infty} F^{[k]} = 1$ and $\mathrm{Z}(F) = 1$. Thus the ascending and descending central series behave completely differently.

Lemma 14.8. *Let G be a finitely generated group. Then the following holds:*

(i) For every $n \in \mathbb{N}$, G has only finitely many subgroups of index n.

(ii) (Hall) *If $H \leq G$ with $|G : H| < \infty$, then there exists a characteristic subgroup $K \leq G$ with $K \leq H$ and $|G : K| < \infty$.*

Proof.

(i) Every subgroup $H \leq G$ of index n induces a homomorphism $G \to S_n$ with kernel H_G. Since G is finitely generated, there are only finitely many such homomorphisms. Because of $H_G \leq H$, there are also only finitely many possibilities for H given H_G.

(ii) By (i), $\{\alpha(H) : \alpha \in \mathrm{Aut}(G)\}$ is finite. Then the characteristic subgroup $K := \bigcap_{\alpha \in \mathrm{Aut}(G)} \alpha(H)$ has finite index. $\square$

Theorem 14.9 (Baumslag)**.** *If G is finitely generated and residually finite, then $\mathrm{Aut}(G)$ is residually finite as well.*

Proof. Let $\alpha \in \mathrm{Aut}(G) \setminus \{1\}$ and $g \in G$ with $\alpha(g) \neq g$. Then there exists $N \trianglelefteq G$ with $|G : N| < \infty$ and $\alpha(g)g^{-1} \notin N$. By Lemma 14.8, the characteristic subgroup $M := \bigcap_{\beta \in \mathrm{Aut}(G)} \beta(N)$ has finite index in G. For $A := \mathrm{C}_{\mathrm{Aut}(G)}(G/M) \trianglelefteq \mathrm{Aut}(G)$ we therefore have $|\mathrm{Aut}(G)/A| \leq |\mathrm{Aut}(G/M)| < \infty$. Clearly, $\alpha \notin A$. $\square$

Definition 14.10. A group G is called *Hopfian* if G is not isomorphic to any proper quotient group of G. This means that every epimorphism $G \to G$ is an automorphism.

Example 14.11. Clearly, all finite groups and all simple groups are Hopfian. For dimension reasons, all finite-dimensional vector spaces are Hopfian as well. On the other hand, $C_2^{\mathbb{N}}$ is not Hopfian.

Theorem 14.12 (Mal'cev)**.** *Every finitely generated residually finite group is Hopfian.*

Proof. Let $\varphi\colon G \to G$ be a noninjective epimorphism. Let $g \in \operatorname{Ker}(\varphi) \setminus \{1\}$. Since G is residually finite, there exists $N \trianglelefteq G$ with $g \notin N$ and $|G : N| < \infty$. Since G is finitely generated, there are only finitely many homomorphisms $\gamma_1, \ldots, \gamma_n\colon G \to G/N$. Here let γ_1 be the canonical epimorphism. From the surjectivity of φ it follows that $\{\gamma_1, \ldots, \gamma_n\} = \{\gamma_1\varphi, \ldots, \gamma_n\varphi\}$. So let $\gamma_1 = \gamma_k\varphi$. Then we obtain the contradiction $1 \neq gN = \gamma_1(g) = \gamma_k(\varphi(g)) = \gamma_k(1) = 1$. □

Corollary 14.13. *Every free group of finite rank is Hopfian.*

Proof. Follows from Theorem 14.5. □

Corollary 14.14. *Let F be free of rank $n < \infty$ and let $F = \langle Y \rangle$ with $|Y| = n$. Then F is free with respect to Y.*

Proof. Let $F = F_X$. An arbitrary bijection $X \to Y \subseteq F$ extends to an epimorphism $\sigma\colon F \to F$. By Mal'cev, $\sigma \in \operatorname{Aut}(F)$. Let G be a group and $\tau\colon Y \to G$. Then the map $\tau\sigma_{|X}$ extends to a homomorphism $\gamma\colon F \to G$. Finally, $\gamma\sigma^{-1}\colon F \to G$ is an extension of τ. Because of $F = \langle Y \rangle$, this is the only extension. Hence F satisfies the universal property with respect to Y. □

Theorem 14.15. *Every finitely generated residually nilpotent group is Hopfian.*

Proof. Let G be finitely generated and residually nilpotent. By assumption, $\bigcap_{k=1}^{\infty} G^{[k]} = 1$, and by Lemma 13.7 the factors $G^{[k]}/G^{[k+1]}$ are finitely generated abelian groups. Suppose that $G \cong G/N$ for some $1 \neq N \trianglelefteq G$. Let $k \in \mathbb{N}$ with $N \subseteq G^{[k]}$ and $N \nsubseteq G^{[k+1]}$. Then

$$G^{[k]}/G^{[k+1]} \cong (G/N)^{[k]}/(G/N)^{[k+1]} \cong G^{[k]}/G^{[k+1]}N \cong (G^{[k]}/G^{[k+1]})/(G^{[k+1]}N/G^{[k+1]}).$$

By the fundamental theorem of finitely generated abelian groups, however, $G^{[k]}/G^{[k+1]}$ is Hopfian. Contradiction. □

Theorem 14.16 (Hall)**.** *There exists a finitely generated solvable group that is not Hopfian.*

Proof. Let R be the ring of all numbers of the form $a2^b$ with $a, b \in \mathbb{Z}$. Let $E_{ij} \in R^{3\times 3}$ be the matrix with a 1 at position (i, j) and zeros elsewhere. For $a, b, c \in R$ we set

$$n(a, b, c) := 1_3 + aE_{12} + bE_{23} + cE_{13} = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}.$$

From $E_{ij}E_{kl} = \delta_{jk}E_{il}$ (cf. GT, Chapter 10) it follows that

$$n(a_1, b_1, c_1)\, n(a_2, b_2, c_2) = n(a_1 + a_2,\ b_1 + b_2,\ c_1 + c_2 + a_1b_2). \tag{14.1}$$

In particular, $N := \{n(a,b,c) : a,b,c \in R\} \leq \mathrm{GL}(3,R)$ is the group of upper triangular matrices with ones on the main diagonal, and $n(a,b,c)^{-1} = n(-a,-b,ab-c)$. From (14.1) we also obtain

$$\begin{aligned}[n(a_1,b_1,c_1), n(a_2,b_2,c_2)] &= n(a_1+a_2,\, b_1+b_2,\, c_1+c_2+a_1b_2) \\ &\quad \cdot n(-a_1-a_2,\, -b_1-b_2,\, a_1b_1-c_1+a_2b_2-c_2+a_1b_2) \\ &= n(0,0,a_1b_1+2a_1b_2+a_2b_2-(a_1+a_2)(b_1+b_2)) = n(0,0,a_1b_2-a_2b_1).\end{aligned} \tag{14.2}$$

We set $u := n(1,0,0)$, $v := n(0,1,0)$, $w := n(0,0,1)$, $t := \mathrm{diag}(1,2,1) \in \mathrm{GL}(3,R)$ and $H := \langle t,u,v\rangle$.

Conjugation by an invertible diagonal matrix $\mathrm{diag}(\delta_1,\delta_2,\delta_3)$ maps E_{ij} to $\delta_i\delta_j^{-1}E_{ij}$. In particular,

$$t^d n(a,b,c) t^{-d} = n(a2^{-d},\, b2^d,\, c) \tag{14.3}$$

for $d \in \mathbb{Z}$. By (14.1) and (14.3), $w \in \mathrm{Z}(H)$ and $\langle w\rangle \trianglelefteq H$. Because of $w^k = n(0,0,k)$ for $k \in \mathbb{Z}$, we have $\langle w^2\rangle < \langle w\rangle$.

Now let $s := \mathrm{diag}(2,2,1) \in \mathrm{GL}(3,R)$. Conjugation by s maps E_{12}, E_{23}, E_{13} to E_{12}, $2E_{23}$, $2E_{13}$. For the monomorphism $f\colon H \to \mathrm{GL}(3,R)$, $x \mapsto sxs^{-1}$ we therefore have $f(t) = t$, $f(u) = u$, $f(v) = v^2$ and $f(w) = w^2$. In particular, $f(H) \subseteq H$. Because of $t^{-1}vt = n(0,2^{-1},0)$, we also have $v = f(t^{-1}vt) \in f(H)$. Hence f is an automorphism of H with $f(\langle w\rangle) = \langle w^2\rangle$, i. e., f induces an isomorphism $H/\langle w\rangle \cong H/\langle w^2\rangle$. For $G := H/\langle w^2\rangle$ and $K := \langle w\rangle/\langle w^2\rangle \neq 1$ it follows that

$$G/K \cong H/\langle w\rangle \cong H/\langle w^2\rangle = G.$$

Hence G is not Hopfian. Together with H, G is also finitely generated. By (14.2), $N' = \langle w\rangle$ and $H/N = \langle tN\rangle$ are abelian. Therefore H and G are solvable. □

Definition 14.17. Let $w = w(x_1,\ldots) \in W \subseteq F_X$. For a group G and $g_1,\ldots \in G$ let $w(g_1,\ldots) \in G$ be the element obtained by replacing x_i in w by g_i (only finitely many x_i occur in w). One calls

$$W(G) := \langle w(g_1,\ldots) : w \in W,\ g_1,\ldots \in G\rangle$$

the *verbal subgroup* with respect to W.

Example 14.18. For $W = \{[x,y]\}$ we obtain $W(G) = G'$.

Remark 14.19. For every homomorphism $f\colon G \to H$ we have $f(W(G)) \leq W(H)$, i. e., verbal subgroups are fully invariant.

Theorem 14.20 (Neumann)**.** *Every fully invariant subgroup of a free group is verbal.*

Proof. Let $F = F_X$ and let $H \leq F$ be fully invariant. Let $w = w(x_1,\ldots) \in H$ and let $y_1,\ldots \in F$ be arbitrary. Then there exists an endomorphism $\alpha\colon F \to F$ with $\alpha(x_i) = y_i$. We have $w(y_1,\ldots) = \alpha(w) \in H$. For $W := H$ we therefore have $H = W(F)$. □

Definition 14.21. Let again $W \subseteq F = F_X$. A normal subgroup N of a group G is called *W-marginal* if

$$w(g_1,\ldots,g_ia,g_{i+1},\ldots) = w(g_1,\ldots)$$

holds for all $w \in W$, $g_1,\ldots \in G$, $i \in \mathbb{N}$ and $a \in N$. This is equivalent to

$$w(g_1,\ldots) = w(h_1,\ldots)$$

for all $g_1,h_1,\ldots \in G$ with $g_i \equiv h_i \pmod N$ for $i \in \mathbb{N}$. The subgroup generated by all W-marginal normal subgroups is called the *W-marginal subgroup* $W^*(G)$.

Remark 14.22. Clearly, $W^*(G)$ itself is W-marginal. For $W = \{[x, y]\}$ we have $W^*(G) = \mathrm{Z}(G)$, since

$$g \in W^*(G) \Rightarrow \forall h \in G : [g, h] = [1g, h] = [1, h] = 1 \Rightarrow g \in \mathrm{Z}(G) \Rightarrow g \in W^*(G).$$

Lemma 14.23. *We have $W(G) = 1$ if and only if $W^*(G) = G$.*

Proof. From $W(G) = 1$ it certainly follows that $W^*(G) = G$. Now let $W^*(G) = G$ and $g_1, \ldots \in G$. From $g_i \equiv 1 \pmod{G}$ it follows that $w(g_1, \ldots) = w(1, \ldots) = 1$. This shows $W(G) = 1$. □

Definition 14.24. A class of groups $\mathcal{X}$ is called a *variety*, if there exists $W \subseteq F_X$ with $\mathcal{X} = \{G : W(G) = 1\}$. In this case, one writes $\mathcal{X} = \mathcal{X}(W)$.

Example 14.25. The abelian groups form the variety $\mathcal{X}([x, y])$. The elementary abelian p-groups form the variety $\mathcal{X}([x, y], x^p)$. The groups whose exponent divides n form the variety $\mathcal{X}(x^n)$, etc. The next theorem implies that the solvable groups of derived length $\leq n$ form a variety.

Theorem 14.26 (Birkhoff)**.** *A class of groups $\mathcal{X}$ is a variety if and only if $\mathcal{X}$ is closed under forming quotient groups and subdirect products.*

Proof. Every variety $\mathcal{X}(W)$ is closed with respect to subgroups, quotient groups and (sub)direct products, because for $H \leq G$ we have $W(H) \leq W(G)$, for $N \trianglelefteq G$ we have $W(G/N) = W(G)N/N$, and for a family $\{G_i : i \in I\}$ we have $W(\times G_i) \leq \times W(G_i)$. Now let $\mathcal{X}$ be a class of groups that is closed with respect to quotient groups and subdirect products. Let $X = \{x_1, x_2, \ldots\}$ be countable and

$$W := \{w \in F_X : \forall G \in \mathcal{X} : w(G) = 1\}.$$

Then $\mathcal{X} \subseteq \mathcal{X}(W)$. Conversely, let $G \in \mathcal{X}(W)$. For every $w \in F_X \setminus W$ we choose $H(w) \in \mathcal{X}$ with $w(h_1, \ldots) \neq 1$ for certain $h_1, \ldots \in H(w)$. We set

$$Y := X \cup G \cup \bigcup_{w \in F_X \setminus W} H(w).$$

Then Y is infinite and at least as large as G and all $H(w)$. Let $\varphi\colon F_Y \to G$ be an epimorphism with $N := \mathrm{Ker}(\varphi)$. Let $v \in F_Y \setminus N$. Only finitely many letters $y_1, \ldots, y_n \in Y$ occur in v, i. e., $v = w(y_1, \ldots, y_n)$ for some $w = w(x_1, \ldots, x_n) \in F_X$. Because of $w(\varphi(y_1), \ldots, \varphi(y_n)) = \varphi(v) \neq 1$ and $G \in \mathcal{X}(W)$, we have $w \notin W$. Because of $|Y| \geq |H(w)|$, there exists an epimorphism $\psi_v\colon F_Y \to H(w)$ with $\psi_v(y_i) = h_i$ for $i = 1, \ldots, n$. For $K(v) := \mathrm{Ker}(\psi_v)$ we have $\psi_v(v) = w(h_1, \ldots, h_n) \neq 1$, so $v \notin K(v)$. We set

$$K := \bigcap_{v \in F_Y \setminus N} K(v) \trianglelefteq F_Y.$$

Then F_Y/K is a subdirect product of the $\mathcal{X}$-groups $F_Y/K(v) \cong H(w)$. By assumption, $F_Y/K \in \mathcal{X}$. For $v \in F_Y \setminus N$ we have $v \notin K(v)$, in particular $v \notin K$. Hence $K \subseteq N$, and $G \cong F_Y/N \cong (F_Y/K)/(N/K)$ is a quotient group of a $\mathcal{X}$-group, and thus $G \in \mathcal{X}$. This shows that $\mathcal{X} = \mathcal{X}(W)$ is a variety. □

Corollary 14.27. *If a class of groups is closed under forming quotient groups and subdirect products, then it is also closed under forming subgroups.*

Definition 14.28. Let $\mathcal{X}$ be a variety. A group $F \in \mathcal{X}$ is called *$\mathcal{X}$-free with respect to $X \subseteq F$* if for every $\mathcal{X}$-group G and every map $\sigma\colon X \to G$ there exists exactly one homomorphism $\widehat{\sigma}\colon F \to G$ with $\widehat{\sigma}(x) = \sigma(x)$ for all $x \in X$.

Theorem 14.29. *Let $F = F_X$ be a free group and $\mathcal{X} = \mathcal{X}(W)$ a variety with $W \subseteq F$. Then $\overline{F} := F/W(F)$ is $\mathcal{X}$-free with respect to $\overline{X} := \{xW(F) : x \in X\}$, and every $\mathcal{X}$-free group with respect to $\overline{X}$ is isomorphic to $\overline{F}$.*

Proof. Because of $W(\overline{F}) = 1$ we have $\overline{F} \in \mathcal{X}$. Let G be a $\mathcal{X}$-group and $\sigma\colon \overline{X} \to G$. We define $\tau\colon X \to G$ by $\tau(x) := \sigma(xW(F))$ for $x \in X$. Then there exists exactly one homomorphism $\widehat{\tau}\colon F \to G$ with $\widehat{\tau}(x) = \tau(x)$ for $x \in X$. Because of $G \in \mathcal{X}$ we have $\widehat{\tau}(w) = 1$ for all $w \in W(F)$. We obtain a homomorphism $\widehat{\sigma}\colon \overline{F} \to G$ with $\widehat{\sigma}(xW(F)) = \widehat{\tau}(x) = \tau(x) = \sigma(xW(F))$ for $x \in X$. Because of $\overline{F} = \langle \overline{X} \rangle$, $\widehat{\sigma}$ is uniquely determined by the images of $\overline{X}$.

Let H also be $\mathcal{X}$-free with respect to $\overline{X}$. Then $\mathrm{id}_{\overline{X}}$ can be extended to homomorphisms $\varphi\colon \overline{F} \to H$ and $\psi\colon H \to \overline{F}$. As usual, it follows that φ is an isomorphism. □

Remark 14.30. In contrast to free groups, there is not a $\mathcal{X}$-free group with respect to X for every set X. For $\mathcal{X} = \mathcal{X}(x) = \{1\}$, for instance, every $\mathcal{X}$-free group is trivial.

Theorem 14.31. *Let $\mathcal{X}$ be a variety. Then every $\mathcal{X}$-group is isomorphic to a quotient group of a $\mathcal{X}$-free group.*

Proof. Let $G \in \mathcal{X} = \mathcal{X}(W)$ with generating set X. Let $F := F_X$ and $\overline{X} := \{xW(F) : x \in X\}$ as in Theorem 14.29. Let $x, y \in X$ with $x \equiv y \pmod{W(F)}$. Then $y^{-1}x \in W(F)$. From $W(G) = 1$ it follows that $x = y$. Hence $|\overline{X}| = |X|$, and there exists a bijection $\sigma\colon \overline{X} \to X \subseteq G$ that extends to an epimorphism $\overline{F} \to G$. The claim follows from Theorem 14.29. □

Theorem 14.32 (Engel)**.** *Let $x_1, \ldots, x_k \in X$ with $x_{k-1} \neq x_k$. Then every finite group $G \in \mathcal{X}([x_1, \ldots, x_k])$ is nilpotent.*

Proof. Let G be a minimal counterexample. Then every proper subgroup/quotient group of G is nilpotent. By Schmidt (GT, Theorem 5.45), $G = P \rtimes Q$ with $P \in \mathrm{Syl}_p(G)$ and $Q \in \mathrm{Syl}_q(G)$. Suppose that $\Phi(P) \neq 1$. Then $G/\Phi(P)$ is nilpotent as well. Because of $\Phi(P) \leq \Phi(G)$, $G/\Phi(G)$ and G would then be nilpotent. Hence $\Phi(P) = 1$, i. e., P is elementary abelian. Let $y \in Q \setminus \{1\}$. Since P is abelian, there exists $\alpha \in \mathrm{End}(P)$ with $\alpha(g) = [y, g]$ for $g \in P$. For $x, g \in P$ we have $[y, xy] = [y, x] = \alpha(x)$ and

$$[xy, g] = {}^{x}[y, g][x, g] = [y, g] = \alpha(g).$$

In $[x_1, \ldots, x_k]$ we replace every occurrence of the variable x_k by xy and all other variables by y. Because of $x_{k-1} \neq x_k$, the innermost bracket is then $[y, xy] = \alpha(x)$, while the outer positions are filled with y or xy and therefore each apply α. Inductively, the assumption yields $\alpha^{k-1}(x) = 1$. If $\beta \in \mathrm{Aut}(P)$ is the conjugation by y, then $\alpha = \beta - 1 \in \mathrm{End}(P)$. Let $m \in \mathbb{N}$ with $p^m \geq k$. Since P is elementary abelian, it follows that

$$0 = \alpha^{p^m} = \sum_{i=0}^{p^m} \binom{p^m}{i} (-1)^i \beta^{p^m - i} = \beta^{p^m} - 1$$

(for $p = 2$ we have $-1 = 1$). Hence $y^{p^m} \in \mathrm{C}_G(P)$, and $y \in \mathrm{C}_G(P)$ because of $p \neq q$. But then G would be nilpotent. □

Corollary 14.33 (Zorn)**.** *Let G be a finite group and $k \in \mathbb{N}$ with $[\underbrace{g, \ldots, g}_{k}, h] = 1$ for all $g, h \in G$. Then G is nilpotent.*

Remark 14.34. If even $[g_1, \ldots, g_k] = 1$ holds for all $g_1, \ldots, g_k \in G$, then G is nilpotent of nilpotency class $\leq k-1$ by GT, Lemma 3.3 (cf. Lemma 13.8).

15 p-Groups

Definition 15.1. Let $X = \{x_1, \ldots\}$ and $\mathcal{X} = \mathcal{X}(x^{p^2}, [x,y]^p, [x,y,z] : x,y,z \in F_X)$. Let G_r be the $\mathcal{X}$-free group of rank $r \geq 1$ from Theorem 14.29. As usual, we identify x_i with the corresponding coset in G_r.

Lemma 15.2. *A finite p-group G lies in $\mathcal{X}$ if and only if $\Phi(G) \leq \mathrm{Z}(G)$ and $\Phi(\Phi(G)) = 1$.*

Proof. Let G be a p-group with $\Phi(G) \leq \mathrm{Z}(G)$ and $\Phi(\Phi(G)) = 1$. For $x,y,z \in G$ we have $x^p \in \Phi(G)$ and $x^{p^2} = 1$. Moreover, $[x,y] \in G' \leq \Phi(G)$ and $[x,y]^p = 1$ follows. Finally, $[y,z] \in \Phi(G) \leq \mathrm{Z}(G)$ and therefore $[x,y,z] = [x,[y,z]] = 1$.

Conversely, let $G \in \mathcal{X}$ and $N := \langle x^p, [x,y] : x,y \in G\rangle$. Then G/N is elementary abelian and $\Phi(G) \leq N$. Since $[x^p, y] = [x,y]^p = 1$ and $[x,y,z] = 1$, we have $N \leq \mathrm{Z}(G)$ and N is elementary abelian, i. e., $\Phi(\Phi(G)) \leq \Phi(N) = 1$. □

Lemma 15.3. *We have:*

(i) $\Phi(G_r)$ is elementary abelian of rank $r(r+1)/2$.

(ii) $G_r/\Phi(G_r)$ is elementary abelian of rank r.

(iii) If $\alpha \in \mathrm{Aut}(G_r)$ acts trivially on $G_r/\Phi(G_r)$, then it also acts trivially on $\Phi(G_r)$.

Proof.

(i) Since $[x_i, x_j x_k] = [x_i, x_j] \cdot {}^{x_j}[x_i, x_k] = [x_i, x_j][x_i, x_k]$ for $1 \leq i,j,k \leq r$, the group $G_r' = \langle [x_i, x_j] : 1 \leq i < j \leq r\rangle$ is elementary abelian of rank at most $r(r-1)/2$. Since $(x_i x_j)^p \equiv x_i^p x_j^p \pmod{G_r'}$, the group $G_r^p G_r' = \langle x_i^p, [x_i, x_j]\rangle$ is elementary abelian of rank at most $r(r-1)/2 + r = r(r+1)/2$. Since $G_r/G_r^p G_r'$ is also elementary abelian of rank at most r, G_r is a finite p-group and $\Phi(G_r) = G_r^p G_r' \leq \mathrm{Z}(G_r)$.

Suppose that there is a relation of the form

$$g := \prod_{i=1}^{r} x_i^{pa_i} \prod_{1 \leq i < j \leq r} [x_i, x_j]^{b_{ij}} = 1$$

with $0 \leq a_i, b_{ij} \leq p-1$. For $\langle h\rangle \cong C_{p^2}$ and $1 \leq i \leq r$ there exists a homomorphism $\varphi\colon G_r \to \langle h\rangle$ with $\varphi(x_i) = h$ and $\varphi(x_j) = 1$ for $j \neq i$ (definition of the $\mathcal{X}$-free group). It follows that $1 = \varphi(g) = h^{pa_i}$ and $a_i = 0$ for $i = 1, \ldots, r$. For $1 \leq i < j \leq r$ let $\langle h_i, h_j\rangle \cong p_+^{1+2}$ analogously. Again there exists a homomorphism $\varphi\colon G_r \to \langle h_i, h_j\rangle$ with $\varphi(x_i) = h_i$, $\varphi(x_j) = h_j$ and $\varphi(x_k) = 1$ for $i \neq k \neq j$. From $1 = \varphi(g) = [h_i, h_j]^{b_{ij}}$ we obtain $b_{ij} = 0$. This shows that $\Phi(G_r)$ has rank $r(r+1)/2$.

(ii) Let $0 \leq a_1, \ldots, a_r \leq p-1$ with $g = x_1^{a_1} \ldots x_r^{a_r} \in \Phi(G_r)$. As in (i) there exists a homomorphism $\varphi\colon G_r \to \langle h\rangle \cong C_{p^2}$ with $h^{a_i} = \varphi(g) \in \Phi(\langle h\rangle) = \langle h^p\rangle$. This shows $g = 1$ and the claim follows.

(iii) Let $\alpha \in \mathrm{Aut}(G_r)$ act trivially on $G_r/\Phi(G_r)$. Then there exist $h_1, \dots, h_r \in \Phi(G_r)$ with $\alpha(x_i) = x_i h_i$ for $i = 1, \dots, r$. By (i) we have $\alpha(x_i^p) = \alpha(x_i)^p = (x_i h_i)^p = x_i^p$ and $\alpha([x_i, x_j]) = [x_i h_i, x_j h_j] = [x_i, x_j]$ for $1 \le i < j \le r$. □

Lemma 15.4. *Let $N, M \le \Phi(G_r)$. Then $G_r/N \cong G_r/M$ if and only if there exists an $\alpha \in \mathrm{Aut}(G_r)$ with $\alpha(N) = M$.*

Proof. Every $\alpha \in \mathrm{Aut}(G_r)$ with $\alpha(N) = M$ induces an isomorphism $G_r/N \cong G_r/M$. Conversely, let an isomorphism $\alpha' \colon G_r/N \to G_r/M$ be given. Choose $y_1, \dots, y_r \in G_r$ with $\alpha'(x_i N) = y_i M$ for $i = 1, \dots, r$. There exists a homomorphism $\alpha \colon G_r \to G_r$ with $\alpha(x_i) = y_i$ for $i = 1, \dots, r$. Since G_r is a finite p-group by Lemma 15.3, we have

$$G_r = \langle y_1, \dots, y_r \rangle M = \langle y_1, \dots, y_r \rangle \Phi(G_r) = \langle y_1, \dots, y_r \rangle.$$

Hence α is surjective and an isomorphism. Since $\alpha(x_i)M = y_i M = \alpha'(x_i N)$, we have $\alpha(g)M = \alpha'(gN)$ for all $g \in G_r$. Here we have

$$g \in N \iff \alpha'(gN) = 1 \iff \alpha(g)M = 1 \iff \alpha(g) \in M.$$

This shows $\alpha(N) = M$. □

Lemma 15.5. *The number of d-dimensional subspaces of $\mathbb{F}_p^n$ is*

$$\binom{n}{d}_p = \frac{(p^n - 1)(p^n - p) \dots (p^n - p^{d-1})}{(p^d - 1)(p^d - p) \dots (p^d - p^{d-1})}$$

and we have $p^{d(n-d)} \le \binom{n}{d}_p \le p^{d(n-d+1)}$.

Proof. The formula is known from combinatorics. The lower estimate follows from $p^{n-d} \le \frac{p^n - p^i}{p^d - p^i}$ for $0 \le i \le d-1$. On the other hand, $p^n + p^{n-d+i+1} \le 2p^n \le p^{n+1} + p^i$ and this yields $\frac{p^n - p^i}{p^d - p^i} \le p^{n-d+1}$. □

Lemma 15.6. *Let $r, s \in \mathbb{N}$ with $1 \le s \le r(r+1)/2$. Then there exist at least $p^{rs(r+1)/2 - r^2 - s^2}$ nonisomorphic groups of order p^{r+s}.*

Proof. Let $\mathcal{N}$ be the set of all $N \le \Phi(G_r)$ with $|\Phi(G_r) : N| = p^s$. By Lemma 15.2, $N \le \mathrm{Z}(G_r)$ and $N \trianglelefteq G_r$. By Lemma 15.3 and Lemma 15.5 we have $|\mathcal{N}| \ge p^{rs(r+1)/2 - s^2}$. For $N \in \mathcal{N}$ we have $|G_r/N| = p^{r+s}$. By Lemma 15.4 it suffices to estimate the number of orbits of $\mathrm{Aut}(G_r)$ on $\mathcal{N}$. Let $\Gamma \colon \mathrm{Aut}(G_r) \to \mathrm{Aut}(G_r/\Phi(G_r))$ be the canonical homomorphism. By Lemma 15.3, the automorphisms in $\mathrm{Ker}(\Gamma)$ act trivially on $\mathcal{N}$. On the other hand,

$$|\mathrm{Aut}(G_r/\Phi(G_r))| = |\mathrm{GL}(r, p)| = (p^r - 1)(p^{r-1} - p) \dots (p^r - p^{r-1}) \le p^{r^2}.$$

Hence every orbit of $\mathrm{Aut}(G_r)$ on $\mathcal{N}$ has length at most p^{r^2}. Thus the number of orbits is at least $p^{rs(r+1)/2 - r^2 - s^2}$. □

Theorem 15.7 (Higman)**.** *Let $f(p^n)$ be the number of nonisomorphic groups of order p^n for a prime p. Then*

$$\boxed{p^{\frac{2}{27} n^2 (n-6)} \le f(p^n) \le p^{\frac{1}{6}(n^3 - n)}.}$$

Proof. For the lower estimate we may assume $n \geq 6$. Let

$$s := \begin{cases} n/3 & \text{if } n \equiv 0 \pmod 3, \\ (n+2)/3 & \text{if } n \equiv 1 \pmod 3, \\ (n+1)/3 & \text{if } n \equiv 2 \pmod 3 \end{cases}$$

and $r = n - s$. Then $s \leq r \leq r(r+1)/2$ and

$$\begin{aligned} rs(r+1)/2 - r^2 - s^2 &= \begin{cases} \frac{1}{27}n^2(2n+3) - \frac{4}{9}n^2 - \frac{1}{9}n^2 & \text{if } n \equiv 0 \pmod 3 \\ \frac{1}{27}(n+2)(n-1)(2n+1) - \frac{4}{9}(n-1)^2 - \frac{1}{9}(n+2)^2 & \text{if } n \equiv 1 \pmod 3 \\ \frac{1}{27}(n+1)(2n-1)(n+1) - \frac{1}{9}(2n-1)^2 - \frac{1}{9}(n+1)^2 & \text{if } n \equiv 2 \pmod 3 \end{cases} \\ &= \begin{cases} \frac{2}{27}n^2(n-6) & \text{if } n \equiv 0 \pmod 3 \\ \frac{2}{27}n^2(n-6) + \frac{1}{3}n - \frac{26}{27} & \text{if } n \equiv 1 \pmod 3 \\ \frac{2}{27}n^2(n-6) + \frac{2}{9}n - \frac{7}{27} & \text{if } n \equiv 2 \pmod 3 \end{cases} \\ &\geq \frac{2}{27}n^2(n-6). \end{aligned}$$

By Lemma 15.6 we therefore have $f(p^n) \geq p^{\frac{2}{27}n^2(n-6)}$.

For the upper estimate, let G be a group of order p^n with chief series $G = H_0 > H_1 > \ldots > H_n = 1$. We choose $g_i \in H_{i-1} \setminus H_i$ for $i = 1, \ldots, n$. Every element of G can then be written uniquely in the form $g = g_1^{a_1} \ldots g_n^{a_n}$ with $0 \leq a_1, \ldots, a_n \leq p-1$. Here $g \in H_i$ if and only if $a_1 = \ldots = a_i = 0$. Let $0 \leq b_{ij} \leq p-1$ with

$$g_i^p = g_{i+1}^{b_{i,i+1}} \ldots g_n^{b_{i,n}}. \tag{15.1}$$

Since $H_{j-1}/H_j \leq \mathrm{Z}(G/H_j)$, we have $[g_i, g_j] \in H_j$ for $1 \leq i < j \leq n$. Hence there exist $0 \leq c_{ijk} \leq p-1$ with

$$[g_i, g_j] = g_{j+1}^{c_{i,j,j+1}} \ldots g_n^{c_{i,j,n}} \qquad (i < j). \tag{15.2}$$

We now show that G is uniquely determined by the relations (15.1) and (15.2). For this, one has to bring a product $g_1^{a_1} \ldots g_n^{a_n} g_1^{a_1'} \ldots g_n^{a_n'}$ into the form $g_1^{a_1''} \ldots g_n^{a_n''}$. This is clear for $n = 1$. Let $n \geq 2$. Using (15.2) one can move $g_1^{a_1'}$ to the left by inserting terms of the form $g_j^{c_{1,j,j}}$ with $j \geq 2$. Afterwards one obtains $g_1^{a_1+a_1'}h$ with $h \in H_1$. Using (15.1) one can replace $a_1 + a_1'$ by a_1'' by inserting further terms of the form $g_i^{b_{1i}}$ with $i \geq 2$. The claim now follows by induction. Hence the number of nonisomorphic groups of order p^n is bounded by the choice of the parameters b_{ij} and c_{ijk}. There are $p^{n(n-1)/2}$ possibilities for the choice of the b_{ij}. There are p^α possibilities for the choice of the c_{ijk}, where

$$\alpha = \sum_{i=1}^{n-2} \binom{n-i}{2} = \binom{n}{3} = \frac{n^3 - 3n^2 + 2n}{6}.$$

(Count the number of 3-element subsets of $\{1, \ldots, n\}$ with given minimum.) Altogether we obtain at most $p^{(n^3-n)/6}$ groups of order p^n. □

Remark 15.8. Sims-Newman-Seeley proved the stronger estimate $f(p^n) \leq p^{\frac{2}{27}n^3 + O(n^{5/2})}$. Since we only used quotient groups of G_r in the proof of the lower estimate, it follows: Almost all p-groups G have a normal subgroup $N \leq \mathrm{Z}(G)$ such that N and G/N are elementary abelian. In particular, almost all p-groups have nilpotency class 2.

Definition 15.9. For finite sets $A, B \subseteq \mathbb{N}$ let $A \prec B$ if $|A| < |B|$ or if $|A| = |B|$ and A is lexicographically smaller than B (i. e., $\min((A \cup B) \setminus (A \cap B)) \in A$).

Remark 15.10. Clearly, $\prec$ is a total order on the set of finite subsets of $\mathbb{N}$. From $A \subsetneq B$ it follows that $A \prec B$.

Lemma 15.11 (HALL's collection process)**.** *Let $X = \{x_1, \dots, x_n\} = X_1 \dot\cup \dots \dot\cup X_m$ and $F := F_X$. For $I \subseteq \{1, \dots, m\}$ let*

$$C_I := F^{[|I|]} \cap \bigcap_{i \in I} \langle X_i \rangle^F \cap \langle \bigcup_{i \in I} X_i \rangle \leq F.$$

Then there exist $c_I \in C_I$ with

$$x_1 \dots x_n = \prod_{I \subseteq \{1,\dots,m\}} c_I,$$

where the subsets I are traversed according to $\prec$.

Proof. For $J \subseteq \{1, \dots, m\}$ we show

$$x_1 \dots x_n = \prod_{I \prec J} c_I \cdot y_1 \dots y_t$$

with $t \in \mathbb{N}_0$, where every y_i lies in some C_K with $J \preceq K$. The claim then follows with $J = \{1, \dots, m\}$ by setting $c_J = y_1 \dots y_t$. For $J = \varnothing$ the product over $I \prec J$ is empty. For $x_i \in X_j \leq \langle X_j \rangle = C_{\{j\}}$ we can choose $t = n$ and $y_i = x_i$. Now suppose the claim has already been proved for J and let J' be the successor of J with respect to $\prec$. We may assume that at least one y_i lies in C_J, since otherwise $y_i \in C_K$ with $J' \preceq K$ for all i. Let i be minimal. In the case $i > 1$ we write $y_{i-1}y_i = y_i y_{i-1} z$ with $z := [y_{i-1}^{-1}, y_i^{-1}]$ and $y_{i-1} \in C_K \subseteq F^{[|K|]}$, where $J \prec K$. Then

- $z \in [F^{[|K|]}, F^{[|J|]}] \leq F^{[|K|+|J|]} \leq F^{[|J \cup K|]}$.
- $z \in \langle X_j \rangle^F$ for all $j \in J \cup K$.
- $z \in \langle \bigcup_{j \in J \cup K} X_j \rangle$.

This shows $z \in C_{J \cup K}$ with $J \prec J \cup K$. In finitely many steps we can move all $y_i \in C_J$ to the left in this way. We call the product of these y_i c_J. For the remaining factors we now have $y_j \in C_K$ with $J' \preceq K$ as desired. □

Theorem 15.12 (HALL-PETRESCU formula)**.** *For every group G and $x, y \in G$ there exist uniquely determined elements $c_i \in G^{[i]}$ such that for all $n \in \mathbb{N}$ we have*

$$\boxed{x^n y^n = (xy)^n c_2^{\binom{n}{2}} c_3^{\binom{n}{3}} \dots c_n^{\binom{n}{n}}.}$$

Proof. W. l. o. g. let $G := \langle x, y \rangle$. The uniqueness of the c_i follows inductively:

$$c_2 = (xy)^{-2} x^2 y^2, \qquad c_n = c_{n-1}^{-\binom{n}{n-1}} \dots c_2^{-\binom{n}{2}} (xy)^{-n} x^n y^n.$$

Let $N \in \mathbb{N}$ be arbitrary. We show that the formula holds for all $n \leq N$. By uniqueness it then holds for all $n \in \mathbb{N}$. Let $X_i := \{x_i, x_{i+N}\}$ and $X = X_1 \dot\cup \dots \dot\cup X_N = \{x_1, \dots, x_{2N}\}$. Let $F := F_X$. For $I \subseteq \{1, \dots, N\}$ let $\mu_I \colon F \to F$ be the endomorphism with

$$\mu_I(x_i) = \begin{cases} x_i & \text{if } i \in I \vee i - N \in I, \\ 1 & \text{otherwise.} \end{cases}$$

Let $x_1 \dots x_{2N} = \prod c_I$ as in Lemma 15.11. For $I \subseteq J$ we have $\mu_J(c_I) = c_I$, since $C_I \subseteq \langle \bigcup_{i\in I} X_i \rangle$. For $I \not\subseteq J$, on the other hand, we have $\mu_J(c_I) = 1$, since for $i \in I \setminus J$ we have $C_I \subseteq \langle X_i \rangle^F = \langle x_i, x_{i+N} \rangle^F \subseteq \mathrm{Ker}(\mu_J)$. By Lemma 15.11 we therefore have

$$\prod_{j\in J} x_j \prod_{j\in J} x_{j+N} = \mu_J(x_1 \dots x_{2N}) = \prod_{I\subseteq J} c_I,$$

where the sets I are ordered with respect to $\prec$.

Let $\varphi\colon F \to G$ be a homomorphism with $\varphi(x_i) = x$ and $\varphi(x_{i+N}) = y$ for $i = 1, \dots, N$. For $|J| = n$ we obtain

$$x^n y^n = \varphi(\mu_J(x_1 \dots x_{2N})) = \prod_{I\subseteq J} \varphi(c_I).$$

We claim that $\varphi(c_I)$ depends only on $|I|$. This is clear for $n = 0$ with $c_\varnothing = 1$. For $n = 1$ we obtain $\varphi(c_I) = xy$ for all one-element sets I. Now let $n := |J| = |J'|$ and

$$\Big(\prod_{I\subsetneq J} \varphi(c_I)\Big)\varphi(c_J) = \prod_{I\subseteq J} \varphi(c_I) = x^n y^n = \prod_{I'\subseteq J'} \varphi(c_{I'}) = \Big(\prod_{I'\subsetneq J'} \varphi(c_{I'})\Big)\varphi(c_{J'}).$$

On both sides the I are sorted in ascending order by size. For $l := |I| = |I'| < n$ we already have $\varphi(c_I) = \varphi(c_{I'})$ by induction. The number of these factors is $\binom{n}{l}$ on both sides. Hence $\varphi(c_J) = \varphi(c_{J'})$.

We set $c_l := \varphi(c_{\{1,\dots,l\}})$ for $l = 1, \dots, N$. Then $c_l \in \varphi(F^{[l]}) \subseteq G^{[l]}$ and

$$x^n y^n = \prod_{l=1}^{n} c_l^{\binom{n}{l}} = c_1^n \prod_{l=2}^{n} c_l^{\binom{n}{l}} = (xy)^n \prod_{l=2}^{n} c_l^{\binom{n}{l}}. \qquad \square$$

Example 15.13. We have $c_2 = (xy)^{-2}x^2y^2 = y^{-1}x^{-1}y^{-1}xy^2 = [y^{-1}x^{-1}, y^{-1}]$. In the case $\langle x, y\rangle^{[3]} = 1$ one obtains the formula $(xy)^n = x^n y^n [y, x]^{\binom{n}{2}}$ known from GT, Exercise 18.

Lemma 15.14. *Let p be a prime and G nilpotent of class $k < p$. For all $x, y \in G$ and $q = p^n$ we have:*

(i) *There exists a $z \in G'$ with $x^q y^q = (xy)^q z^q$.*

(ii) $x^q = y^q \iff (xy^{-1})^q = 1$.

(iii) $\Omega_q(G) := \{x \in G : x^q = 1\} \trianglelefteq G$.

(iv) $\mho_q(G) := \{x^q : x \in G\} \trianglelefteq G$.

(v) $|G| = |\Omega_q(G)||\mho_q(G)|$.

(vi) $\mho_q(G)' \le \mho_{q^2}(G')$.

Proof. Induction on k: In the case $k \le 1$, G is abelian and all statements follow easily. Let $k \ge 2$ and w. l. o. g. $G = \langle x, y\rangle$ (this is not used in (v)).

(i) By assumption we have $G^{[p]} \le G^{[k+1]} = 1$. By the Hall-Petrescu formula there exist $c_i \in G^{[i]} \le G'$ with

$$x^q y^q = (xy)^q c_2^{\binom{q}{2}} \dots c_{p-1}^{\binom{q}{p-1}}.$$

One easily sees that $\binom{q}{i}$ is divisible by q for $0 < i < p$. By induction, (iv) holds for G'. This shows $c_2^{\binom{q}{2}} \dots c_{p-1}^{\binom{q}{p-1}} \in \mho_q(G')$.

(ii) Let $x^q = y^q$. Then

$$x^q = yx^q y^{-1} = (yxy^{-1})^q.$$

Let $H := \langle x, yxy^{-1}\rangle \leq G$. Since

$$[x, yxy^{-1}] = [x, yxy^{-1}x^{-1}] = [x, y, x] \in G^{[3]}$$

we have $H' = \langle [x, yxy^{-1}]\rangle^H \leq G^{[3]}$. Hence H has nilpotency class $< k$. By induction, $[x, y]^q = (x(yxy^{-1})^{-1})^q = 1$. Since G' also has nilpotency class $< k$, we have

$$G' = \langle [x, y]\rangle^G = \langle g[x, y]g^{-1} : g \in G\rangle \leq \Omega_q(G') = G',$$

i. e., $\exp(G') \leq q$. From (i) it follows that $(xy^{-1})^q = x^q y^{-q} = 1$.

Conversely, let $(xy^{-1})^q = 1$. Then also $(y^{-1}x)^q = x^{-1}(xy^{-1})^q x = 1$. By the first part of the equivalence we obtain $[x, y^{-1}]^q = (xy^{-1}x^{-1}y)^q = 1$. As before, we get $\exp(G') \leq q$. By (i), $x^q y^{-q} = (xy^{-1})^q = 1$ and $x^q = y^q$.

(iii) For $x, y \in \Omega_q(G)$ we have $x^q = 1 = y^q$ and $(xy^{-1})^q = 1$ by (ii). This shows $xy^{-1} \in \Omega_q(G)$ and $\Omega_q(G) \leq G$. Clearly, $\Omega_q(G)$ is a normal subgroup of G.

(iv) For $x^q, y^q \in \mho_q(G)$ there exists a $z \in G'$ with $x^q y^{-q} = (xy^{-1})^q z^q$ by (i). For $H := \langle xy^{-1}, z\rangle$ we have $[xy^{-1}, z] \in G^{[3]}$ and $H' = \langle [xy^{-1}, z]\rangle^H \leq G^{[3]}$. In particular, H has nilpotency class $< k$. By induction we have $(xy^{-1})^q z^q \in \mho_q(H) \leq \mho_q(G)$ and $\mho_q(G) \leq G$. Clearly, $\mho_q(G)$ is normal in G.

(v) The map $f\colon G \to \mho_q(G)$, $x \mapsto x^q$ is surjective (but in general not a homomorphism). By (ii) we have $f(x) = f(y)$ if and only if $xy^{-1} \in \Omega_q(G)$. Hence $f^{-1}(x^q) = x\Omega_q(G)$ and $|f^{-1}(x^q)| = |\Omega_q(G)|$ for all $x \in G$. If $\Omega_q(G)$ is infinite, then so is G. Otherwise, $|G : \Omega_q(G)| = |\mho_q(G)|$.

(vi) Let $x^q, y^q \in \mho_q(G)$. We apply (ii) to $\overline{G} := G/\mho_{q^2}(G')$:

$$[x^q, y^q] = x^q(y^q x y^{-q})^{-q} \in \mho_{q^2}(G') \iff (xy^q x^{-1} y^{-q})^q \in \mho_{q^2}(G') \iff \overline{[x, y^q]} \in \Omega_q(\overline{G}).$$

Now we apply (ii) to $\overline{G}/\Omega_q(\overline{G})$:

$$\overline{[x, y^q]} = (\overline{xyx^{-1}})^q \overline{y}^{-q} \in \Omega_q(\overline{G}) \iff \overline{[x, y]}^q \in \Omega_q(\overline{G}) \iff [x, y]^{q^2} \in \mho_{q^2}(G').$$

The statement on the right-hand side is clearly true. □

Example 15.15. The statements in Lemma 15.14 do not hold in general for groups with nilpotency class $\geq p$. For example, since $|\Omega_2(D_8)| = 6$, $\Omega_2(D_8)$ cannot be a subgroup of D_8. For

$$G := \langle x, y \mid x^4 = y^4 = 1,\ yxy^{-1} = x^{-1}\rangle \cong C_4 \rtimes C_4$$

the set $\mho_2(G) = \{1, x^2, y^2\}$ is not a subgroup of G.

Remark 15.16. The next theorem shows that p-groups with "small" nilpotency class behave approximately like abelian groups.

Theorem 15.17 (Groves)**.** *Let P be a p-group with nilpotency class $< p$. Then there exists an operation $P \times P \to P$, $(x, y) \mapsto x + y$ with the following properties:*

(i) $P_+ := (P, +)$ is an abelian group.

(ii) The orders of x in P and P_+ are equal.

(iii) *For all $x, y \in P$ we have $x + y \in \langle x, y\rangle$.*

(iv) *Every endomorphism of P is also an endomorphism of P_+.*

(v) *Every subgroup of P is also a subgroup of P_+.*

Proof. Let P be generated by n elements and $F := F_n/F_n^{[p]}$ the free nilpotent group with rank n and nilpotency class $p-1$. Then there exists an epimorphism $f\colon F \to P$, $x \mapsto \overline{x}$. Let $q = \exp(P')$. For all $x, y \in F$ there exists, by Lemma 15.14 (applied to $\langle x, y\rangle$), an $s = s(x, y) \in \langle x, y\rangle$ with $x^q y^q = s^q$. Let also $t \in F$ with $s^q = t^q$. Then $(st^{-1})^q = 1$ follows from Lemma 15.14. Since F is torsion-free (Corollary 13.36), we obtain $s = t$, i. e., s is uniquely determined by x, y. We regard s as a word in x, y and define $a + b := s(a, b)$ for $a, b \in P$.

(i) For $x, y, z \in F$ we have

$$s(s(x, y), z)^q = s(x, y)^q z^q = (x^q y^q) z^q = x^q (y^q z^q) = x^q s(y, z)^q = s(x, s(y, z))^q.$$

As above, it follows that $s(s(x, y), z) = s(x, s(y, z))$. This shows that $+$ is associative on P. Since $s(x, 1) = x = s(1, x)$, 1 is neutral with respect to $+$. Since $s(x, x^{-1}) = 1$, x^{-1} is inverse to x with respect to $+$.

By Lemma 15.14 we have $s(x, y)^q s(y, x)^{-q} = [x^q, y^q] \in \mathfrak{V}_{q^2}(F')$ and $(s(x, y)s(y, x)^{-1})^q \in \mathfrak{V}_{q^2}(F')$. Let $z \in F'$ with $(s(x, y)s(y, x)^{-1})^q = z^{q^2} = (z^q)^q$. From Lemma 15.14 it follows that $(s(x, y)s(y, x)^{-1}z^{-q})^q = 1$. Since F is torsion-free, we obtain $s(x, y)s(y, x)^{-1} = z^q \in \mathfrak{V}_q(F')$. Since $\mathfrak{V}_q(P') = 1$, it follows that

$$a + b = s(a, b) = s(b, a) = b + a$$

for all $a, b \in P$, i. e., P_+ is an abelian group.

(ii) Clearly, $a + a = s(a, a) = a^2$ and inductively $k \cdot a = a^k$ for all $a \in P$ and $k \in \mathbb{N}$.

(iii) Follows from $s(x, y) \in \langle x, y\rangle$.

(iv) Let $\alpha \in \mathrm{End}(P)$. Since $a + b$ is a word in a and b, $\alpha(a + b)$ is the corresponding word in $\alpha(a)$ and $\alpha(b)$. This shows $\alpha(a + b) = \alpha(a) + \alpha(b)$.

(v) Let $Q \leq P$. For $a, b \in Q$, $a - b = a + b^{-1}$ is a word in a and b. This shows $a - b \in Q$ and $Q \leq P_+$. □

Remark 15.18.

(i) The isomorphism type of P_+ can easily be determined from $|\Omega_q(P)| = |\Omega_q(P_+)|$ for $q = p, p^2, \ldots$.

(ii) Let $p > 2$ and P a p-group with nilpotency class 2. Let $q := \exp(P')$. In the free group with nilpotency class 2 we have

$$x^q y^q = (xy)^q [y, x]^{-\binom{q}{2}} = (xy[y, x]^{\frac{1-q}{2}})^q$$

by Example 15.13. Since the map $P \to P$, $x \mapsto x^2$ is bijective, every $x \in P$ has exactly one "root" $\sqrt{x} \in P$ with $\sqrt{x}^2 = x$. As in the proof of Theorem 15.17,

$$x + y := xy[y, x]^{\frac{1-q}{2}} = xy\sqrt{[y, x]}$$

yields an abelian group structure on P. This special case was first constructed by Baer.

(iii) For p-groups with nilpotency class $3 < p$ the formula reads

$$x + y := xy\sqrt{[y^{-1}, x^{-1}]}\ \sqrt[12]{[x, y, x]}\ \sqrt[12]{[y, x, y]}$$

(without proof). In general, these terms are obtained from the *Baker-Campbell-Hausdorff formula*.

(iv) Theorem 15.17 already holds if every subgroup of P generated by three elements has nilpotency class $< p$ (three generators are necessary to prove the associative law).

(v) The *Lazard correspondence* yields a Lie ring structure on p-groups with nilpotency class $< p$. Here P_+ is the additive group of this Lie ring.

Example 15.19.

(i) All p-groups P with $|P| \leq p^p$ have nilpotency class $< p$ and therefore satisfy the hypothesis of Theorem 15.17. For $P = \texttt{SmallGroup}(5^5, 21)$ we have $|\Omega_5(P)| = 5^3$ and $|\Omega_{25}(P)| = 5^4$. Hence $P_+ \cong C_{5^3} \times C_5^2$.

(ii) In general, not every endomorphism of P_+ is an endomorphism of P. For example, the center of $P := p_+^{1+2}$ is characteristic (or has order p), while $P_+ \cong C_p^3$ is characteristically simple. Moreover, P has only $p + 1$ maximal subgroups, while P_+ has, by Lemma 15.5,

$$\binom{3}{2}_p = \frac{(p^3 - 1)(p^3 - p)}{(p^2 - 1)(p^2 - p)} = p^2 + p + 1$$

maximal subgroups.

16 Decision Problems

Remark 16.1 (Dehn's problems)**.**

(i) Let G be a group with generating set X. In "practice" one would like to solve the following tasks algorithmically:

- (Word problem) When does a word in X represent the identity element in G?
- (Conjugacy problem) When are two words in X conjugate as elements of G?
- (Isomorphism problem) When is G isomorphic to another given group H?

(ii) A solution of the word problem means (for countable X) that every element in G can be brought into a "normal form". By Lemma 1.7, the word problem is therefore solvable for free groups (provided one already knows that G is free).

(iii) If the conjugacy problem is solvable for G, then so is the word problem, since $g = 1$ if and only if g and 1 are conjugate.

(iv) Novikov and Boone constructed a finitely presented group for which the word problem (and hence also the conjugacy problem) is unsolvable. Adian and Rabin showed the unsolvability of the isomorphism problem for a certain finitely presented group. That is, there is no algorithm that solves one of the three problems for these groups in finite time. Even the question whether a given group is finite or trivial cannot be decided!

(v) In general, the solvability of the problems depends on the chosen presentation. For finitely generated groups the situation is better.

Theorem 16.2. *Let $G = \langle X \mid R\rangle$ be finitely generated. If the word problem (respectively, the conjugacy problem) is solvable with respect to this presentation of G, then the word problem (respectively, the conjugacy problem) is also solvable for every other finitely generated presentation of G.*

Proof. Let $F := F_X$ and let $\varphi\colon F \to G$ be the canonical epimorphism. Let $F_1 := F_{X_1}$ and let $\varphi_1\colon F_1 \to G$ be another presentation with $|X_1| < \infty$. Then there exists a function $\psi\colon X_1 \to F$ with $(\varphi\psi)(x) = \varphi_1(x)$ for all $x \in X_1$. By the universal property, ψ extends to F_1. This extension can be computed explicitly from the finitely many values $\psi(x)$ with $x \in X_1$. Now let $w \in F_1$. If the word problem for $\langle X \mid R\rangle$ is solvable, then $\varphi(\psi(w)) = 1$ can be decided. Hence $\varphi_1(w) = 1$ can also be decided, i. e., the word problem for $\langle X_1 \mid R_1\rangle$ with $\mathrm{Ker}(\varphi_1) = \langle R_1\rangle^{F_1}$ is solvable. Similarly for the conjugacy problem. □

Theorem 16.3. *The conjugacy problem is solvable for all free groups.*

Proof. Clearly every $g \in F_X$ is conjugate to a cyclically reduced word $\hat{g}$. If $\hat{g}$ and $\hat{h}$ differ only by shifts, then g and h are conjugate in F_X. Conversely, suppose that g and h are conjugate. Then $\hat{g}$ and $\hat{h}$ are also conjugate. Let $a \in F_X$ be reduced with $a\hat{g}a^{-1} = \hat{h}$. Since $\hat{h}$ is cyclically reduced, a or a^{-1} must cancel completely with $\hat{g}$. Hence $\hat{g}$ is a shift of $\hat{h}$. In this way one can decide whether g and h are conjugate. □

Theorem 16.4. *Let G be finitely presented and residually finite. Then the word problem is solvable for G.*

Proof. Let $G = \langle X \mid R\rangle$ be a finite presentation and w a word in $X \subseteq F_X$. We run the following algorithms in parallel (or alternately):

- One constructs all (countably many) products of elements of the form $gr^{\pm 1}g^{-1}$ with $r \in R$ and $g \in F_X$ (for example sorted by length and lexicographically). If $w = 1$ in G, then w must appear at some point.
- One constructs all (countably many) finite groups H. Since X is finite, there are only finitely many maps $X \to H$. Since $|R| < \infty$, one can check which of them extend to homomorphisms $\varphi\colon G \to H$. One checks $\varphi(w) \neq 1$. If $w \neq 1$, then there exist H and φ with $\varphi(w) \neq 1$, since G is residually finite. □

Theorem 16.5. *The word problem is solvable for Coxeter groups.*

Proof. Let $G = \langle x_1, \ldots, x_n\rangle$ be a Coxeter group of rank n and $w = x_{i_1} \ldots x_{i_l} \in G$. Let $\sigma\colon G \to \mathrm{GL}(n, \mathbb{R})$ be the monomorphism from Theorem 11.14. Then $\sigma(w) = \sigma_{i_1} \ldots \sigma_{i_l}$ can be computed from the numbers m_{ij} alone. (On a computer, one could realize the matrix entries $\cos(\pi/m_{ij})$ discretely in a cyclotomic field instead of working with floating-point numbers, which are prone to rounding errors.) We have $w = 1$ if and only if $\sigma(w) = 1$. □

Remark 16.6.

(i) Tits gave an algorithm for Theorem 16.5: Let $G = \langle x_1, \ldots, x_n\rangle$ be a Coxeter group and $\pi\colon F_X \to G$ the canonical epimorphism with $X = \{x_1, \ldots, x_n\}$. For $w = x_{i_1} \ldots x_{i_k} \in F_X$ let $R(w) \subseteq F_X$ be the set of all words that can be obtained from w by the following operations:

- Remove a subword of the form $x_i x_i$ in w.

- Replace a subword of the form $x_i x_j x_i \ldots$ ($m_{ij} < \infty$ letters) by $x_j x_i x_j \ldots$ (m_{ij} letters).

Since the length of the words in $R(w)$ is bounded by k, $R(w)$ is finite. Clearly $\pi(r) = \pi(w)$ for all $r \in R(w)$. One can show that $\pi(w) = 1$ if and only if $1 \in R(w)$ (without proof).

(ii) The isomorphism problem for Coxeter groups has not yet been solved completely.

(iii) The difficulty of Dehn's problems is exploited in cryptography. As an example, we describe the Anshel-Anshel-Goldfeld *protocol* for determining a shared secret key between persons A and B. Let a group G with efficiently computable normal form be given. The public key of A (respectively B) consists of random elements $a_1, \ldots, a_n$ (respectively $b_1, \ldots, b_n$) in G. The private keys consist of words w_A (respectively w_B) in $a_1, \ldots, a_n$ (respectively $b_1, \ldots, b_n$). The key exchange is based on the following principle:

(1) A sends the n-tuple $(w_A b_1 w_A^{-1}, \ldots, w_A b_n w_A^{-1})$ (in normal form) to B.

(2) B sends $(w_B a_1 w_B^{-1}, \ldots, w_B a_n w_B^{-1})$ to A.

(3) Both can now compute

$$w_A w_A(w_B a_1 w_B^{-1}, \ldots, w_B a_n w_B^{-1})^{-1} = [w_A, w_B] = w_B(w_A b_1 w_A^{-1}, \ldots, w_A b_n w_A^{-1}) w_B^{-1}$$

and use it as a shared key. In general (without special group structure), an attacker cannot draw any conclusions about w_A, w_B or $[w_A, w_B]$ from a_i, b_i, $w_A b_i w_A^{-1}$ and $w_B a_i w_B^{-1}$.

Such cryptographic methods become interesting if existing methods such as RSA become unusable with powerful quantum computers.

Exercises

Exercise 1. Let F be a free group of rank > 1. Show $\mathrm{Z}(F) = 1$.

Exercise 2. Show:

(a) $Q_{2^n} \cong \langle x, y \mid x^{2^{n-2}} = y^2,\ yxy^{-1} = x^{-1} \rangle$ for $n \geq 3$ (see GT, Theorem 8.15)

(b) $A_4 \cong \langle x, y \mid x^2 = y^3 = (xy)^3 = 1 \rangle$.

Exercise 3. A group G is called *metacyclic* if there exists a cyclic normal subgroup N with cyclic quotient group G/N (example: GT, Theorem 7.24). Let P be a finite metacyclic p-group. Show that $a, b, c, k \geq 0$ with $k^{p^b} - 1 \equiv 0 \equiv p^c(k-1) \pmod{p^a}$ exist such that

$$P \cong \langle x, y \mid x^{p^a} = 1,\ y^{p^b} = x^{p^c},\ yxy^{-1} = x^k \rangle$$

holds.
Remark: Different parameters can give rise to isomorphic groups. An exact classification was given by Liedahl.

Exercise 4. Let p be a prime, $a \geq 2$ and $b \geq 1$. Let

$$P(a,b) := P = \langle x, y \mid x^{p^a} = y^{p^b} = 1,\ yxy^{-1} = x^{1+p^{a-1}} \rangle.$$

Show:

(a) $|P| = p^{a+b}$.

(b) $P' = \langle x^{p^{a-1}} \rangle \cong C_p$.

(c) $\Phi(P) = \mathrm{Z}(P) = \langle x^p, y^p \rangle \cong C_{p^{a-1}} \times C_{p^{b-1}}$.

Exercise 5. A nonabelian group G is called *minimal nonabelian* if every proper subgroup of G is abelian. Show that for a finite p-group P the following statements are equivalent:

(a) P is minimal nonabelian.

(b) $|P : \Phi(P)| = |P : \mathrm{Z}(P)| = p^2$.

(c) $|P : \Phi(P)| = p^2$ and $|P'| = p$.

Hint: GT, Chapter 4.

Exercise 6. Let p be a prime and $a, b \in \mathbb{N}$. Let

$$Q(a,b) := Q = \langle x, y \mid x^{p^a} = y^{p^b} = [x,y]^p = [x,x,y] = [y,x,y] = 1 \rangle$$

(recall: $[x,y,z] := [x,[y,z]]$). Show:

(a) $|Q| = p^{a+b+1}$.

(b) $Q' = \langle [x,y] \rangle \cong C_p$.

(c) $\Phi(Q) = \mathrm{Z}(Q) = \langle x^p, y^p, [x,y] \rangle \cong C_{p^{a-1}} \times C_{p^{b-1}} \times C_p$.

Remark: Rédei showed that every minimal nonabelian p-group is isomorphic to $P(a,b)$, $Q(a,b)$ or Q_8.

Exercise 7. Decide whether the Petersen graph is the Cayley graph of a group.

Exercise 8. Let F be the free group over the alphabet $\{x, y\}$ and $G \leq F$ the subgroup of all words of even length. Show that G is freely generated by x^2, y^2 and xy.

Exercise 9. Let $G = \langle X \mid R \rangle$ be a simple group, where every relator in R has even length. Show: $G \cong C_2$.

Exercise 10. Let F be a free group and $w \in F \setminus \{1\}$. Show $\mathrm{C}_F(w) \cong C_\infty$. Deduce that F is indecomposable.

Exercise 11. Show that a group H is free if and only if every extension of H splits.

Exercise 12. The *coproduct*

$$A := \coprod_{n\in\mathbb{Z}} C_2 \leq \prod_{n\in\mathbb{Z}} C_2$$

consists of all sequences $(a_n)_{n\in\mathbb{Z}}$ with $|\{n \in \mathbb{Z} : a_n \neq 1\}| < \infty$. Clearly there exists $\gamma \in \mathrm{Aut}(A)$ with $\gamma((a_n)_n) := (a_{n+1})_n$ for all $(a_n)_n \in A$. Let $G := A \rtimes \langle\gamma\rangle$. Show that G is finitely generated and metabelian, but the subgroup $A \leq G$ is not finitely generated.

Exercise 13. Determine all extensions of C_2 by C_n up to equivalence ($n \in \mathbb{N}$).

Exercise 14. Let

$$D_\infty := \langle x, y \mid x^2 = y^2 = 1\rangle \cong C_\infty \rtimes C_2$$

(GT, Exercise 61(b)). Show that $G := \langle x, y \mid x^2 = y^2\rangle$ is an extension of D_∞ by $\mathrm{Z}(G) \cong C_\infty$. Deduce that G is supersolvable and G' is cyclic.

Exercise 15. Let $A := \langle a\rangle \cong C_4$, $S := \mathrm{SL}(2,3)$ and $G := (S \times A)/\langle(-1_2, a^2)\rangle \cong S * A$ (central product). Show:

(a) $S \cong Q_8 \rtimes C_3$.

(b) $Q_8 \trianglelefteq G$ has a complement in $H := Q_8 * A \in \mathrm{Syl}_2(G)$.

(c) Q_8 has no complement in G.

Remark: Hence Theorem 4.23 of Gaschütz is false for nonabelian groups N.

Exercise 16. Let $N := D_8$ and $H \cong C_2$.

(a) Show $\mathrm{Aut}(N) \cong D_8$.

(b) Determine all extensions of H by N up to equivalence. Which of them are isomorphic?

Exercise 17. Let $N = \langle a, b \mid a^8 = b^2 = 1,\ bab^{-1} = a^{-1}\rangle \cong D_{16}$ and $H = \langle x\rangle \cong C_2$. Show:

(a) There exists an automorphism $\beta \in \mathrm{Aut}(N)$ with $\beta(a) = a^3$ and $\beta(b) = ab$.

(b) There exists a homomorphism $\omega\colon H \to \mathrm{Out}(N)$ with $\omega(x) = \beta\mathrm{Inn}(N)$.

(c) There is no parameter system of H by N with coupling ω.

Hint: Theorem 4.16.

Exercise 18. We consider the simple group $N := \mathrm{PSL}(2,9)$ of order 360. We identify the elements in N with their preimages in $\mathrm{SL}(2,9)$ (note: $\mathrm{Z}(\mathrm{SL}(2,9)) = \langle -1_2\rangle$). Let $\mathbb{F}_9^\times = \langle\zeta\rangle$ and $d := \mathrm{diag}(\zeta, 1) \in \mathrm{GL}(2,9)$. Show:

(a) The map $\sigma\colon N \to N$, $x \mapsto dxd^{-1}$ is an outer automorphism with $\sigma^2 \in \mathrm{Inn}(N)$.

(b) The map $\tau\colon N \to N$, $(x_{ij}) \mapsto (x_{ij}^3)$ is an outer automorphism of order 2.

(c) Let $H = \langle x\rangle \cong C_2$. By Corollary 4.28, for each $\alpha_x \in \{1, \sigma, \tau, \sigma\tau\}$ there is a parameter system (α, κ) of H by N. Investigate which of the extensions split and which are isomorphic.

Remark: We have $N \cong A_6$.

Exercise 19. Let N be abelian and $\alpha\colon H \to \mathrm{Aut}(N)$ a group homomorphism. A map $\delta\colon H \to N$ with

$$\boxed{\delta(xy) = \delta(x)\alpha_x(\delta(y))}$$

for all $x, y \in H$ is called a *crossed homomorphism* with respect to α. Show:

(a) The crossed homomorphisms form a group $\mathrm{Hom}_\alpha(H,N) \leq C^1(H,N)$.

(b) The map $\Gamma\colon N \to \mathrm{Hom}_\alpha(H,N)$, $a \mapsto \delta_a$ with $\delta_a(x) := \alpha_x(a)a^{-1}$ is a homomorphism. One sets $H^1_\alpha(H,N) := \mathrm{Hom}_\alpha(H,N)/\Gamma(N)$.

(c) Let K be a complement of N in $G := N \rtimes_\alpha H$. For $x \in H$ there exists exactly one $y \in K$ with $\delta_K(x) := xy^{-1} \in N$. We have $\delta_K \in \mathrm{Hom}_\alpha(H,N)$.

(d) Every $\delta \in \mathrm{Hom}_\alpha(H,N)$ defines a complement $K_\delta := \{\delta(x)^{-1}x : x \in H\}$ of N in G with $\delta_{K_\delta} = \delta$.

(e) Two complements K_1, K_2 of N in G are conjugate if and only if $\delta_{K_1}\delta_{K_2}^{-1} \in \Gamma(N)$ holds.

(f) The map $K \to \delta_K$ induces a bijection between the conjugacy classes of complements of N in G and $H^1_\alpha(H,N)$.

Exercise 20 (Gaschütz)**.** Let N be an abelian normal subgroup of a finite group G and $N \leq H \leq G$ with $\gcd(|N|, |G:H|) = 1$. Show: If all complements of N in H are conjugate, then all complements of N in G are conjugate.

Exercise 21.

(a) Show $\mathrm{Aut}(S_n) \cong \mathrm{Aut}(A_n)$ for $n \geq 4$.

(b) Show that $\varphi \in \mathrm{Aut}(S_6)$ with

$$\begin{aligned}
&\varphi((1,2)) = (1,5)(2,3)(4,6), && \varphi((1,3)) = (1,4)(2,6)(3,5),\\
&\varphi((1,4)) = (1,3)(2,4)(5,6), && \varphi((1,5)) = (1,2)(3,6)(4,5),\\
&\varphi((1,6)) = (1,6)(2,5)(3,4)
\end{aligned}$$

is an outer automorphism of order 2.

Exercise 22. Let $P = P(a,1) = M_{p^{a+1}}$ be the minimal nonabelian group from Exercise 4 (or GT, Theorem 8.15) with $a \geq 2$. In the case $p = 2$ let $a \geq 3$. Show $M(P) = 1$.
Hint: Theorem 5.25.

Exercise 23. Let $G_i \approx H_i$ be isoclinic groups for $i = 1, 2$. Show $G_1 \times G_2 \approx H_1 \times H_2$.

Exercise 24. Show that the simple group $\mathrm{PSp}(4,3)$ is isomorphic to no alternating group and no projective special linear group.
Remark: By Example 8.13 we have $\mathrm{PSp}(4,3) \cong \mathrm{PSU}(4,2)$. With the help of *Zsigmondy primes* one can show that $\mathrm{PSp}(2n,q)$ for $n \geq 2$ is isomorphic to no alternating group and no projective special linear group.

Exercise 25. Let V be a unitary space. Show that V has a basis consisting of elements in V_0.

Exercise 26. Let V be a unitary space of dimension n and $U \leq V$ with $U \subseteq U^\perp$. Show $\dim U \leq \lfloor n/2 \rfloor$. Give an example in which equality holds.
Remark: In general (i. e., for arbitrary "scalar products" on V) one calls

$$\max\{\dim U : U \leq V,\ U \subseteq U^\perp\}$$

the *Witt index* of V.

Exercise 27. Show $\mathrm{Sp}(2n, q) \leq \mathrm{SU}(2n, q)$ for all $n \in \mathbb{N}$ and prime powers $q \neq 1$.

Exercise 28. Prove $\mathrm{PSU}(3, 2) \cong M_9$.
Hint: GT, Theorem 6.21.

Exercise 29. We consider the orthogonal group $\mathrm{GO}^\epsilon(2k, q)$ with q odd. Show: The Witt index is k for $\epsilon = 1$ and $k - 1$ for $\epsilon = -1$ (see Exercise 26).

Exercise 30. Show that an action of G on Ω is 2-transitive if and only if $G = G_\omega \cup G_\omega x G_\omega$ for $\omega \in \Omega$ and some $x \in G$.

Exercise 31. Show that one must fill out at least 82 lottery tickets (6 out of 49) in order to have two "correct numbers" (on one ticket).

Exercise 32. Let $S = (\Omega, \mathcal{B})$ be a $(2, k, v)$-Steiner system. Show that the following statements are equivalent:

(1) $v = k^2$.

(2) $|\mathcal{B}| = k^2 + k$.

(3) For $B \in \mathcal{B}$ and $\omega \in \Omega \setminus B$ there exists exactly one block $B' \in \mathcal{B}$ with $\omega \in B'$ and $B \cap B' = \varnothing$.

Remark: In the affine plane $\mathbb{F}_q^2$, (3) is the *parallel postulate*: For every point x and every line g there exists exactly one parallel of g passing through x.

Exercise 33. Show that the Higman-Sims graph is 22-regular, i. e., every vertex is connected to exactly 22 other vertices.

Exercise 34.

(a) Show that $\mathrm{SU}(3, 3)$ has exactly 63 involutions.

(b) Show with GAP that the graph Γ defined for J_2 in Remark 10.20(vii) is 36-regular.

(c) Show $\mathrm{Aut}(\Gamma) \cong J_2 \rtimes C_2$ using the packages `grape` and `atlasrep`.

Exercise 35. A *lattice* is a free abelian subgroup of $\mathbb{R}^n$ of rank n. Hence every lattice has the form $L = \mathbb{Z}b_1 + \ldots + \mathbb{Z}b_n$, where $b_1, \ldots, b_n \in \mathbb{R}^n$ is a basis of $\mathbb{R}^n$. Show that

$$\mathrm{Aut}(L) = \{f \in \mathrm{O}(\mathbb{R}^n) : f(L) = L\}$$

is a finite group. Determine $\mathrm{Aut}(\mathbb{Z}^n)$.

Exercise 36. Let $p > 2$ be a prime and $G := \langle x, y \mid x^p = y^p = (xy)^p = 1\rangle$. Show $|G| = \infty$.
Hint: Realize G as a subgroup of $\mathrm{Sym}(\mathbb{Z})$.

Exercise 37. Show:

(a) Every Hurwitz group G is perfect and 84 divides $|G|$.

(b) $\mathrm{GL}(3, 2)$ is a Hurwitz group.

Exercise 38. Let $G = \mathrm{Fr}_{i\in I} G_i$ and $H = \bigoplus_{i\in I} G_i$.

(a) Construct a natural epimorphism $G \to H$ and describe its kernel.

(b) Show $G/G' \cong H/H'$.

(c) Let $N_i \trianglelefteq G_i$ for $i \in I$. Let N be the normal closure of $\bigcup_{i\in I} N_i$ in G. Show $G/N \cong \mathrm{Fr}_{i\in I} G_i/N_i$.

Exercise 39. Let G be an amalgamated free product of G_I over H with $H < G_i$ for all $i \in I$. Show: $\mathrm{Z}(G) = \bigcap_{i\in I} \mathrm{Z}(G_i)$.

Exercise 40. Show that every finitely generated periodic solvable group is finite.

Exercise 41. Write the element c_3 in the Hall-Petrescu formula in a form that shows that it lies in $G^{[3]}$.

Exercise 42.

(a) Let P be a p-group such that $P^{[p-1]}$ is cyclic. Show that P satisfies the statements of Lemma 15.14.

(b) Construct p-groups with $p > 2$ as in (a) of arbitrarily large nilpotency class.

Index

Symbols

A^*, 34
$B^2(H,N)$, 32
Co_1, 80
Co_2, 80
Co_3, 80
$\delta_n(k)$, 113
D_∞, 136
E_{ij}, 54
$\mathrm{Fr}_{i\in I} G_i$, 101
$G * H$, 101
$\mathrm{G}_2(q)$, 70
$G \approx H$, 44
$\mathrm{GO}(V,\rho)$, 67
$\mathrm{GU}(V)$, 59
$\mathrm{GU}(n,q)$, 59
$H^2(H,N)$, 32
$H^2_s(H,A)$, 32
HJ, 80
HS, 79
J_1, 80
J_2, 80
$k_n(G)$, 45
$l(g)$, 82
$M(G)$, 34
McL, 80
M_d, 75
$\Omega_q(G)$, 129
$\mathfrak{O}_q(G)$, 129
$\Omega(V,\rho)$, 69
$\omega(c_i)$, 113
$\mathrm{P}\Omega(V,\rho)$, 69
$\mathrm{PSO}(V,\rho)$, 68
$\mathrm{PSU}(V)$, 59
$\mathrm{PSU}(n,q)$, 59
$\mathrm{PSp}(V)$, 53
Φ, 84
Π, 84
$\mathrm{SO}(V,\rho)$, 68
$\mathrm{SU}(V)$, 59
$\mathrm{SU}(n,q)$, 59
σ_v, 84
$\mathrm{Sp}(V)$, 53
$\mathrm{Sz}(q)$, 58
$W^*(G)$, 122
$W(G)$, 122
x_v, 84
$Z^2(H,N)$, 32
$Z^2_s(H,A)$, 32
$\mathrm{Z}^*(G)$, 42
$\partial\varphi$, 32

A

Adian-Rabin, 132
affine plane, 77
alphabet, 4
Anshel-Anshel-Goldfeld protocol, 134
antisymmetry, 53
Artin group, 82
automorphism system, 22

B

Baer, 43, 131
Baker-Campbell-Hausdorff formula, 132
basic commutator, 113
Baumslag, 120
Beyl-Felgner-Schmid, 43
Birkhoff, 123
braid group, 82
branching factor, 44
Burnside group, 108

C

Cayley algebra, 70
Cayley graph, 10
CFSG, 71
Chapman, 76
class of groups, 119
cocycle, 32
cohomology group, 32
commutator
    in a free algebra, 116
conjugacy problem, 132
Conway group, 80
coproduct, 136
coupling, 29
Coxeter, 94
Coxeter graph, 89
Coxeter group, 81
    alternating subgroup, 83
    irreducible, 89
    universal, 82
crossed homomorphism, 137

D

Dehn's problems, 132
deletion condition, 89
Dickson invariant, 71
Dynkin diagram, 94

E

Engel, 124
epicenter, 42
exchange condition, 89
extension, 21
    central, 32

equivalent, 21
split, 21

F
factor set, 22
trivial, 22
Fano plane, 77
Fisher's inequality, 78
free algebra, 116
free product, 100
Frucht, 10

G
GAP, 7, 11, 14–16, 18, 20, 22, 27, 31, 36, 46, 48, 51, 52, 61, 79, 97
Gaschütz, 29, 41, 137
Golay code, 79
Golod, 107
group
capable, 42
complete, 31
finitely presented, 6
Hopfian, 120
isoclinic, 44
metacyclic, 134
minimal nonabelian, 135
orthogonal, 67
polycyclic, 27
symplectic, 53
unitary, 59
Groves, 130
Gupta, 107
Guralnick-Kantor-Kassabov-Lubotzky, 17

H
Hall, 8, 45, 120, 121
Hall's collection process, 128
Hall-Higman lemma, 112
Hall-Janko group, 80
Hall-Petrescu formula, 128
Higman, 126
Higman-Neumann-Neumann, 105
Higman-Sims graph, 79
Higman-Sims group, 79
HNN extension, 106
Hölder, 47
Hopf formula, 40
Hurwitz group, 99
hyperbolic pair, 53

I
isoclinism, 44
isometry, 59
isomorphism problem, 132
Iwasawa, 119

J
Jacobi identity, 98
Johnson-Zassenhaus, 31

K
Keevash, 81
Kirkman, 77
Kurosh, 103

L
lattice, 139
Lazard correspondence, 132
Leech lattice, 80
length, 4, 82
letter, 4
Levi, 109, 110
Lie algebra, 98
Liedahl, 134

M
Magnus, 120
Magnus' Freiheitssatz, 17
Mal'cev, 121
Mathieu, 73, 74
Mathieu group, 75
McLaughlin group, 80
Mendelsohn, 78
Miller, 76, 100
Monster group, 99
Moore, 16, 17

N
Neumann, 8, 122
Newman, 20
Nielsen transformation, 19
Nielsen-Schreier, 12
Norton, 81
Novikov-Adian, 106
Novikov-Boone, 132

O
octonions, 70
Ol'shanskii, 110
orthogonal space, 67

P
parallel postulate, 138
parameter system, 22
equivalent, 23
normalized, 22
Petersen graph, 135
ping-pong lemma, 6
presentation
balanced, 9
projective plane, 77
projective representation, 42

Q
quadratic form, 67
quasideterminant, 71

R
rank, 5
    of a Coxeter group, 81
Read, 41
Rédei, 135
reflection, 69, 82
    complex, 98
reflection group, 82
Reidemeister-Schreier, 15
relation, 6
relator, 6
root, 84
    positive/negative, 84
root system, 84

S
Sanov, 111
schoolgirl problem, 77
Schreier, 25
Schreier transversal, 12
Schreier's formula, 13
Schur, 35, 41
Schur extension, 34
    maximal, 35
    universal, 39
Schur multiplier, 34
Shephard-Todd, 98
Sims-Newman-Seeley, 127
spinor norm, 69
Steinberg, 11
Steiner system, 77
stem group, 45
subdirect product, 119
subgroup
    verbal, 122
    W-marginal, 122
Suzuki group, 58
symplectic basis, 53
symplectic space, 53

T
Tarski monster, 110
Tietze transformation, 18
Tits, 88, 133
Todd-Coxeter algorithm, 15
transvection
    symplectic, 55
    unitary, 63

U
unitary space, 58
universal coefficient theorem, 37

V
Valentiner group, 52
Van der Waerden, 110
variety, 123
von Dyck, 7
von Dyck group, 98

W
Wall, 66
weight
    of a basic commutator, 113
Witt, 62, 72
Witt index, 138
Witt's formula, 116
word, 4
    cyclically reduced, 17
    empty, 4
    equivalent, 4
    reduced, 4
        in a Coxeter group, 82
word problem, 132

Z
Zelmanov, 106
Zorn, 124
Zsigmondy prime, 137