Anthony E. Clement
Stephen Majewicz
Marcos Zyman

The Theory of Nilpotent Groups

Birkhäuser

Birkhäuser

The Theory of Nilpotent Groups

Anthony E. Clement • Stephen Majewicz
Marcos Zyman

The Theory of Nilpotent Groups

Birkhäuser

Anthony E. Clement
Department of Mathematics
CUNY-Brooklyn College
Brooklyn, NY, USA

Stephen Majewicz
Mathematics and Computer Science
CUNY-Kingsborough Community College
Brooklyn, NY, USA

Marcos Zyman
Department of Mathematics
CUNY-Borough of Manhattan
Community College
New York, NY, USA

ISBN 978-3-319-66211-4 ISBN 978-3-319-66213-8 (eBook)
DOI 10.1007/978-3-319-66213-8

Library of Congress Control Number: 2017955372

Printed on acid-free paper

This book is published under the trade name Birkhäuser, www.birkhauser-science.com
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

For Gilbert Baumslag, in memoriam.

Preface

Our foremost objective in writing this book was to present a reasonably self-contained treatment of the classical theory of nilpotent groups so that the reader may later be able to study further topics and perhaps undertake research on his or her own. We have also included some recent work by two of the authors.

The theorems and proofs that appear in this work can be found elsewhere in some shape or form, but they are scattered in the literature. We have tried to include some of the omitted details in the original sources and to offer additional computations and explanations whenever we found it appropriate and useful. It is our hope that the examples, constructions, and computations included herein will contribute to a general understanding of both the theory of nilpotent groups and some of the techniques commonly used to study them. With this in mind, we have attempted to produce a single volume that can either be read from cover to cover or used as a reference. This was our main motivation several years ago, when the idea of writing such a book began to materialize.

We expect both working mathematicians and graduate students to benefit from reading this book. We demand from the reader only a solid advanced undergraduate or beginning graduate background in algebra. In particular, we assume that the reader is familiar with groups, rings, fields, modules, and tensor products. We expect the reader to know about direct and semi-direct products. We also assume knowledge about free groups and presentations of groups. Some topology is certainly useful for Chapter 6.

We declare that the choice of topics is based on what we consider to be a coherent discussion of nilpotent groups and mostly responds to our own mathematical interests. We emphasize that some of the more recent developments in nilpotent group theory (especially from the algorithmic, geometric, and model-theoretic perspectives) have been completely excluded and are well-suited topics for future volumes. Furthermore, major results such as the solution of the isomorphism problem for finitely generated nilpotent groups and the Mal'cev correspondence are only mentioned or discussed briefly.

We adopt certain conventions and notations. We sometimes write "1" for the trivial group, the group identity, and the unity of a ring. All functions and morphisms are written on the left unless otherwise told. For example, we write $\varphi(x)$, rather than $x\varphi$ or x^{φ}. The $n \times n$ identity matrix is always denoted by I_n. Whenever we have a field of characteristic zero, we identity its prime subfield with the rationals. All rings mentioned in this book are associative.

The book is organized in the following manner. In Chapter 1, we discuss the commutator calculus. Chapter 2 is meant to serve as an introduction to nilpotent groups and includes some interesting examples. Chapter 3 deals with the collection process and basic commutators, leading to normal forms in finitely generated free groups and free nilpotent groups. In Chapter 4, we show that finitely generated nilpotent groups are polycyclic, allowing us to obtain another type of normal form in such groups. Chapter 5 is about the theory of isolators, root extraction, and localization. Chapter 6 is a discussion of a classical paper by S. A. Jennings regarding the group ring of finitely generated torsion-free nilpotent groups over a field of characteristic zero. Finally, Chapter 7 contains a selection of additional topics.

We take full responsibility for any errors, mathematical or otherwise, appearing in this work. We have made every effort to accurately cite all pertinent works and do apologize for any omissions.

Brooklyn, NY, USA — Anthony E. Clement
Brooklyn, NY, USA — Stephen Majewicz
New York, NY, USA — Marcos Zyman
October 2017

Acknowledgments

First and foremost, we are deeply indebted to the late Professor Gilbert Baumslag, for the inspiration, encouragement, and support he gave us all.

AEC: I want to express my sincere thanks to the following family members who provided encouragement and emotional support throughout the writing of this book: my wife April Mojica-Clement, my children Mathias and Zayli Clement, and my parents Sir Martin and Lady Margaret Clement, as well as my extended family.

SM: I would like to thank Mrs. Jeanne DeVoy and Mrs. Anne Magliaccio, secretaries in the Department of Mathematics and Computer Science at Kingsborough Community College, CUNY, for assisting me beyond their regular duties at the college.

I thank Ms. Lydia Fischbach for her assistance and in particular for her diligence in proofreading several versions of the manuscript.

Many thanks go to Mr. Igor Melamed, instructor of mathematics at Kingsborough, for translating several articles written in Russian.

Most importantly, I thank my wife Annemarie Majewicz for her patience and support throughout the years.

Lastly, I thank my parents John (in memoriam) and Madeline Majewicz, as well as the rest of my family, for their support.

MZ: I would like to thank Professors Mahmood Sohrabi, Alexei Miasnikov, and Robert Gilman for hosting me at Stevens Institute of Technology during the academic year 2015–2016 and for their input and encouragement. I thank Professor Gretchen Ostheimer for many insightful conversations. I am grateful to Professor Joseph Roitberg (in memoriam), who sparked my interest in nilpotent groups.

Many thanks go to the Department of Mathematical Sciences at Stevens for facilitating me with office space and a pleasant working environment during my sabbatical leave from CUNY in beautiful Hoboken, New Jersey.

Last, but by no means least, I thank my beloved family—my wife Adriana Pérez and my kids María and Clara—for their constant support and patience over the last several years. I also thank my parents Sarita and Salomón and my siblings Jacobo, Sami, and Jayele for always being there.

Contents

Notations

g^h	The conjugate of g by h
g^{-h}	The conjugate of g^{-1} by h
$Z(G)$	The center of G
$Hom(G, H)$	The group of homomorphisms from G to H
$Aut(G)$	The automorphism group of G
$ker\ \varphi$	The kernel of φ
$im\ \varphi$	The image of φ
$G \cong H$	G is isomorphic to H
$G \times H$	The direct product of G and H
$Inn(G)$	The group of inner automorphisms of G
S_n	The symmetric group on the set $\{1, 2, \ldots, n\}$
A_n	The alternating group on the set $\{1, 2, \ldots, n\}$
D_n	The dihedral group of order $2n$
$\mathscr{H}$	The Heisenberg group
$C_G(X)$	The centralizer of a nonempty subset X of a group G
$[G : H]$	The index of H in G
$[g, h]$	The commutator of g and h
$[g_1, \ldots, g_n]$	Simple commutator of weight $n > 1$
$gp(S)$	The subgroup generated by S
$[X_1, X_2]$	The commutator subgroup of X_1 and X_2
$[X_1, \ldots, X_n]$	$[[X_1, \ldots, X_{n-1}], X_n]$
$Ab(G)$	The abelianization of G
$N_G(S)$	The normalizer of S in G
S^T	The subgroup of G generated by all conjugates of elements of S by elements of T
S^H	The normal closure of S in $gp(S, H)$
$\gamma_i G$	The ith lower central subgroup of G
$\zeta_i G$	The ith upper central subgroup of G
Q	The quaternion group
$\langle X \mid Y \rangle$	The group presented by generators X and relators Y

$G \rtimes_\varphi H$	The semi-direct product of G and H
$G \wr H$	The wreath product of G and H
$\mathbb{Z}_p$	The cyclic group of order p
$\mathbb{Z}_{p^\infty}$	The Prüfer p-group (or p-quasicyclic group)
$UT_n(R)$	The group of $n \times n$ upper unitriangular matrices over R
I_n	The $n \times n$ identity matrix
$\tau_P(G)$	The set of P-torsion elements of G
$\tau(G)$	The set of torsion elements of G
$\|g\|$	The order of g; the length of a groupoid element g
$A \bigotimes B$	The tensor product of A and B
D_∞	The infinite dihedral group
G^n	The subgroup of G generated by the nth powers
$A\ rep\ a$	The rep operation
$lr(X)$	The Lie ring generated by X
$\bigoplus_{j=0}^{\infty} R_j$	The direct sum of R_j
$R[[x_1, \ldots, x_n]]$	The Magnus power series ring in the variables $x_1, \ldots, x_n$ over R.
$\tau_i(\overline{x})$	The ith Hall-Petresco word
G^R	The R-completion of G with respect to a Mal'cev basis
G^*	The Mal'cev completion of G
$H \leq_R G$	H is an R-subgroup of G
$H \trianglelefteq_R G$	H is a normal R-subgroup of G
$gp_R(S)$	The R-subgroup generated by S
$[H_1, H_2]_R$	The commutator R-subgroup of H_1 and H_2
$[H_1, \ldots, H_n]_R$	$[[H_1, \ldots, H_{n-1}]_R, H_n]_R$
$G \cong_R H$	G is R-isomorphic to H
$\tau_\omega(G)$	The set of ω-torsion elements of G
G_π	The π-primary component of G
$ann(g)$	The annihilator of g
G^α	The R-subgroup of G R-generated by αth powers
$I_P(S, G)$	The P-isolator of S in G
$\vartheta(g_1, \ldots, g_n)$	The value of the word ϑ at the r-tuple $(g_1, \ldots, g_n)$
$W(G)$	The verbal subgroup of G
$\vartheta(H_1, \ldots, H_n)$	The generalized verbal subgroup
$H \underset{P}{\sim} K$	H and K are P-equivalent
$\varrho_P(G)$	The maximal P-radicable subgroup of G
$\varrho(G)$	The maximal radicable subgroup of G
G^{p^∞}	$\bigcap_{n=1}^{\infty} G^{p^n}$
$\mathbb{Z}_P$	The ring $\{m/n \in \mathbb{Q} \mid n \neq 0 \text{ is a } P'\text{-number}\}$
RG	The group ring of G over R
$A_R(G), A(G)$	The augmentation ideal of RG
$D_n(R, G), D_n(G)$	The nth dimension subgroup of G over R
$\overline{\gamma}_n G$	The isolator of $\gamma_n G$ in G
$GL(V)$	The group of all invertible R-module endomorphisms

$GL_n(R)$	The group of non-singular $n \times n$ matrices over R
$M_n(\mathbb{Z})$	The ring of $n \times n$ matrices with integral entries
$(U,\ W)$	span$\{(u,\ w) \mid u \in U,\ w \in W\}$, where U and W are Lie subalgebras
$\overline{FG}$	The completion of FG in the A-adic topology
$\overline{A}$	The completion of the augmentation ideal in the A-adic topology
$1 + \overline{A}$	The group $\left\{1 + a \mid a \in \overline{A}\right\}$
$\overline{D}_k$	The group $1 + \overline{A}^{\,k}$
exp	The exponential map
log	The logarithmic map
$\overline{\Lambda} = \mathscr{L}(\overline{A})$	The Lie algebra $\overline{A}$ under commutation
$\log\ G$	The set $\{\log\ g \mid g \in G\}$
$\mathscr{L}_F(G)$	The Lie algebra of G over F
$g \sim h\ (g \nsim h)$	g is conjugate (not conjugate) to h
$UT_n^m(R)$	The normal subgroup of $UT_n(R)$ consisting of those matrices whose $m - 1$ superdiagonals have 0's in their entries
E_{ij}	The $n \times n$ matrix with 1 in the (i, j) entry and 0's elsewhere
$t_{i,j}(r)$	The transvection $I + rE_{ij}$
$Hol(G)$	The holomorph of G
$IA(G)$	The IA-group of G
$\Phi(G)$	The Frattini subgroup of G
$\mathbb{F}_p$	The finite field with p elements
$Fit(G)$	The Fitting subgroup of G

Chapter 1
Commutator Calculus

In this chapter, we introduce the commutator calculus. This is one of the most important tools for studying nilpotent groups. In Section 1.1, the center of a group and other notions surrounding the concept of commutativity are defined. Several results and examples involving central subgroups and central elements are given. Section 1.2 contains the fundamental identities related to commutators of group elements. By definition, the commutator of two elements g and h in a group G is the element $[g, h] = g^{-1}h^{-1}gh$. Clearly, $[g, h] = 1$ whenever g and h commute. This leads to a natural connection between central elements and trivial commutators. The commutator identities allow us to develop properties of commutator subgroups. This is the main focus of Section 1.3.

1.1 The Center of a Group

The commutator calculus is an essential tool which is used for working with nilpotent groups. In this section, we collect various results on commutators which will be used throughout the book. This material can be found in various places in the literature (see [1–6]).

1.1.1 Conjugates and Central Elements

We begin by defining the conjugate of a group element.

Definition 1.1 Let g and h be elements of a group G. The *conjugate* of g by h, denoted by g^h, is the element $h^{-1}gh$ of G.

A.E. Clement et al., *The Theory of Nilpotent Groups*,
DOI 10.1007/978-3-319-66213-8_1

The conjugate of g^{-1} by h is written as g^{-h}. Notice that

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

Furthermore, if $k \in G$, then

$$(gh)^k = k^{-1}ghk = (k^{-1}gk)(k^{-1}hk) = g^k h^k$$

and

$$\left(g^h\right)^k = \left(h^{-1}gh\right)^k = k^{-1}h^{-1}ghk = (hk)^{-1}g(hk) = g^{hk}.$$

We summarize these in the next lemma.

Lemma 1.1 *Suppose that* $g,\ h,$ *and* k *are elements of any group. Then* $(gh)^k = g^k h^k,\ \left(g^{-1}\right)^h = \left(g^h\right)^{-1}$, *and* $\left(g^h\right)^k = g^{hk}$.

The notion of conjugacy extends to subgroups in a natural way.

Definition 1.2 Two subgroups H and K of a group G are called *conjugate* if $g^{-1}Hg = K$ for some $g \in G$.

In particular, every normal subgroup of G is conjugate to itself.

Definition 1.3 Let G be a group. An element $g \in G$ is called *central* if it commutes with every element of G. The set of all central elements of G is called the *center* of G and is denoted by $Z(G)$. Thus,

$$\begin{aligned} Z(G) &= \{g \in G \mid gh = hg \text{ for all } h \in G\} \\ &= \{g \in G \mid g^h = g \text{ for all } h \in G\}. \end{aligned}$$

It is easy to verify that $Z(G)$ is a normal abelian subgroup of G, and the conjugate of a central element $g \in G$ by any element of G is just g itself.

If G and H are groups, then the (internal and external) direct product of G and H will be written as $G \times H$.

Lemma 1.2 *If* G_1 *and* G_2 *are groups, then* $Z(G_1 \times G_2) = Z(G_1) \times Z(G_2)$.

Proof Suppose that $(g_1,\ g_2) \in Z(G_1 \times G_2)$. Then $(g_1,\ g_2)(x,\ y) = (x,\ y)(g_1,\ g_2)$ for all $(x,\ y) \in G_1 \times G_2$. This implies that $(g_1x,\ g_2y) = (xg_1,\ yg_2)$, and thus $g_1x = xg_1$ and $g_2y = yg_2$. Hence, $g_1 \in Z(G_1)$ and $g_2 \in Z(G_2)$. Therefore, $(g_1,\ g_2)$ is contained in $Z(G_1) \times Z(G_2)$. And so, $Z(G_1 \times G_2) \subseteq Z(G_1) \times Z(G_2)$. In a similar way, one can show that $Z(G_1) \times Z(G_2) \subseteq Z(G_1 \times G_2)$. □

Lemma 1.3 *If* G_1 *and* G_2 *are any two groups, then*

$$\frac{G_1 \times G_2}{Z(G_1 \times G_2)} \cong \frac{G_1}{Z(G_1)} \times \frac{G_2}{Z(G_2)}.$$

Proof The map from $G_1 \times G_2$ to $(G_1/Z(G_1)) \times (G_2/Z(G_2))$ defined by

$$(g_1, g_2) \mapsto (g_1 Z(G), g_2 Z(G))$$

is a surjective homomorphism whose kernel is $Z(G_1 \times G_2)$. The result follows from the First Isomorphism Theorem. □

Let G and H be any two groups. The set of homomorphisms from G to H will be denoted by $Hom(G, H)$, and the group of automorphisms of G by $Aut(G)$. The kernel and image of $\varphi \in Hom(G, H)$ are abbreviated as $ker\ \varphi$ and $im\ \varphi = \varphi(G)$ respectively. If G and H are isomorphic groups, then we write $G \cong H$.

Let G be a group and $h \in G$. Using Lemma 1.1, it is easy to show that the map

$$\varphi_h : G \to G \text{ defined by } \varphi_h(g) = g^h$$

is contained in $Aut(G)$.

Definition 1.4 The map φ_h is called the *conjugation map* or *inner automorphism* induced by h.

It is easy to see that the set of all inner automorphisms of G forms a group under composition. This group is denoted by $Inn(G)$. There is a natural connection between the center of a group and the inner automorphisms of the group.

Theorem 1.1 *Let G be a group and $h \in G$. The map*

$$\varrho : G \to Aut(G) \text{ defined by } \varrho(h) = \varphi_h, \text{ where } \varphi_h(g) = g^h,$$

is a homomorphism with $ker\ \varrho = Z(G)$ and $im\ \varrho = Inn(G)$.

Proof The result follows from Lemma 1.1. □

By Theorem 1.1 and the First Isomorphism Theorem, we have:

Corollary 1.1 *If G is any group, then $G/Z(G) \cong Inn(G)$.*

1.1.2 Examples Involving the Center

In the next few examples, we give the center of various groups.

Example 1.1 A group G is abelian if and only if $Z(G) = G$.

Example 1.2 Let S_n be the symmetric group on the set $S = \{1, 2, \ldots, n\}$, and let "e" denote the identity element of S_n. Clearly, S_1 has trivial center because $S_1 = \{e\}$. Furthermore, $Z(S_2) = S_2$ since S_2 is abelian.

We show that $Z(S_n) = \{e\}$ for $n > 2$. Suppose, on the contrary, that $Z(S_n)$ is nontrivial. Let $\sigma \in Z(S_n)$ be a nonidentity element. There exist distinct elements $a, b \in S$ such that $\sigma(a) = b$. Choose an element $c \in S$ different from a and b, and let

τ be the transposition $(b \;\; c)$. A direct calculation shows that $(\sigma \circ \tau)(a) \neq (\tau \circ \sigma)(a)$, contradicting the assumption that σ is in the center of $Z(S_n)$.

Example 1.3 Let A_n be the alternating group on the set $S = \{1, 2, \ldots, n\}$. This is the subgroup of S_n consisting of all even permutations. Note that $A_1 = A_2 = \{e\}$, and A_3 is cyclic since it has order 3. Thus, $Z(A_n) = A_n$ for $n = 1, 2$, and 3 according to Example 1.1.

The center of A_4 is trivial. The proof is similar to the one used in Example 1.2. Assume that $Z(A_n)$ is nontrivial, and let σ be a nonidentity element of $Z(A_n)$. There exist distinct elements $a, b \in S$ such that $\sigma(a) = b$. Choose two elements c and d in S different from a and b, and let $\tau = (b \;\; c \;\; d)$. It is easy to see that $(\sigma \circ \tau)(a) \neq (\tau \circ \sigma)(a)$, contradicting the assumption that the center is nontrivial.

Using the same argument as above, one can show that A_n has trivial center whenever $n \geq 5$. We provide an alternative proof which uses the fact that A_n is a simple group whenever $n \geq 5$. Since this is the case, either $Z(A_n) = \{e\}$ or $Z(A_n) = A_n$. If it were true that $Z(A_n) = A_n$, then A_n would be abelian by Example 1.1. However, a quick calculation shows that

$$(1 \;\; 2 \;\; 3)(3 \;\; 4 \;\; 5) \neq (3 \;\; 4 \;\; 5)(1 \;\; 2 \;\; 3).$$

Thus, A_n is non-abelian and $Z(A_n) \neq A_n$. We conclude that $Z(A_n) = \{e\}$ for $n \geq 5$.

Example 1.4 Let D_n be the dihedral group of order $2n$, the group of isometries of the plane which preserve a regular n-gon. If y is a reflection across a line through a vertex and x is the counterclockwise rotation by $2\pi/n$ radians, then the elements of D_n are

$$1, x, x^2, \ldots, x^{n-1}, y, xy, x^2y, \ldots, x^{n-1}y,$$

and the equalities

$$x^n = 1, \; y^2 = 1, \; \text{and } xy = yx^{-1}$$

hold in D_n.

Both D_1 and D_2 are abelian, so $Z(D_1) = D_1$ and $Z(D_2) = D_2$. We determine $Z(D_n)$ when $n \geq 3$. Since $xy = yx^{-1}$, we have

$$x^r y = yx^{-r} \qquad (r \in \mathbb{Z}). \tag{1.1}$$

We claim that no element of the form $x^t y$ for any $t \in \{0, 1, \ldots, n-1\}$ is central. Assume, on the contrary, that $x^t y \in Z(D_n)$ for some such t. Then $x^t y$ commutes with x. Hence, $x^{-1}(x^t y)x = x^t y$, and thus $x^{t-1}yx = x^t y$. Applying (1.1) to both sides of this equality yields $yx^{1-t}x = yx^{-t}$. After canceling the y's, we get $x^{2-t} = x^{-t}$. This means that $x^2 = 1$, a contradiction. Therefore, $x^t y \notin Z(D_n)$ for any

$t \in \{0, 1, \ldots, n-1\}$. Consequently, an element of $Z(D_n)$ must take the form x^t for some $t \in \{0, 1, \ldots, n-1\}$. Clearly, $x^0 = 1 \in Z(D_n)$.

Suppose $x^t \in Z(D_n)$ for some $t \in \{1, \ldots, n-1\}$. By (1.1), we have

$$yx^t = x^t y = yx^{-t}.$$

Hence, $x^t = x^{-t}$; that is, $x^{2t} = 1$. Since x has order n, it must be that n divides $2t$. Hence, there exists $k \in \mathbb{N}$ such that $2t = nk$. If $k \geq 2$, then $2t \geq 2n$. This cannot happen since $1 \leq t \leq n-1$. This means that $k = 1$, and thus $2t = n$. Now, if n is odd, then no such t exists. We conclude that $Z(D_n)$ is trivial when n is odd. If n is even, then $t = \frac{n}{2}$, and consequently, $x^{n/2} \in Z(D_n)$. Therefore, $Z(D_n)$ is the cyclic group of order 2 generated by $x^{n/2}$ when n is even.

Example 1.5 Let $\mathscr{H}$ be the group of 3×3 upper unitriangular matrices over $\mathbb{Z}$ with the group operation being matrix multiplication. Thus,

$$\mathscr{H} = \left\{ \begin{pmatrix} 1 & a_{12} & a_{13} \\ 0 & 1 & a_{23} \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; a_{ij} \in \mathbb{Z} \right\}.$$

This group is called the *Heisenberg group*. The identity element in $\mathscr{H}$ is clearly the 3×3 identity matrix and will be denoted by I_3. It is easy to show that

$$Z(\mathscr{H}) = \left\{ \begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; c \in \mathbb{Z} \right\}.$$

1.1.3 Central Subgroups and the Centralizer

Definition 1.5 A subgroup H of a group G is called *central* if $H \leq Z(G)$.

Related to the center of a group is the centralizer of a subset of a group.

Definition 1.6 The *centralizer* of a nonempty subset X of a group G is

$$C_G(X) = \{g \in G \mid g^{-1}xg = x \text{ for all } x \in X\}.$$

It is easy to verify that $C_G(X)$ is a subgroup of G. If $X = \{x\}$, then we write $C_G(x)$ for the centralizer of x. Clearly,

$$C_G(G) = \bigcap_{x \in G} C_G(x) = Z(G).$$

Notice that $C_G(x)$ is just the stabilizer of x under the action of G on itself by conjugation. The orbit of x under this action, called the *conjugacy class* of x, is the set $\{g^{-1}xg \mid g \in G\}$. When G is finite, we get the *class equation* of G :

$$|G| = |Z(G)| + \sum_k [G : C_G(x_k)], \tag{1.2}$$

where one x_k is chosen from each conjugacy class containing at least two elements. Here, $|G|$ stands for the order of G and $[G : H]$ is the index of a subgroup H in G. These notations are standard and will be used throughout the book. We will also write $|g|$ for the order of an element $g \in G$.

1.1.4 The Center of a p-Group

Definition 1.7 Let p be any prime. A group G is called a *p-group* if every element of G has order a power of p.

Finite p-groups are the building blocks of finite groups. The next fact regarding their central structure is important in the study of finite groups.

Theorem 1.2 *If G is a nontrivial finite p-group for some prime p, then* $Z(G) \neq 1$.

Proof Suppose that $|G| = n$. Consider the class equation (1.2) of G. If $x_k \in G$ is not central for some $1 \leq k \leq n$, then $C_G(x_k)$ is a proper subgroup of G. Hence, $[G : C_G(x_k)]$ is a positive power of p. Consequently, each summand in the sum

$$\sum_k [G : C_G(x_k)]$$

is divisible by p. Since p divides $|G|$ by hypothesis, p also divides $|Z(G)|$. Therefore, $Z(G)$ contains nontrivial elements. □

Remark 1.1 It is important to emphasize that G must be finite in Theorem 1.2. An infinite p-group does not necessarily have nontrivial center. This notion is discussed in Remark 2.8.

1.2 The Commutator of Group Elements

One can determine whether or not two group elements commute by calculating their commutator.

Definition 1.8 Let g and h be elements of a group G. The *commutator* of g and h, written as $[g, h]$, is

$$[g, h] = g^{-1}h^{-1}gh = g^{-1}g^h.$$

Clearly, g and h commute if and only if $[g,\ h] = 1$. Thus, the center of G can also be characterized as

$$Z(G) = \{g \in G \mid [g,\ h] = 1 \text{ for all } h \in G\}.$$

Definition 1.9 Let $S = \{g_1,\ g_2,\ \ldots,\ g_n\}$ be a set of elements of a group G. A *simple commutator*, or *left-normed commutator*, of *weight* $n \geq 1$ is defined recursively as follows:

1. The simple commutators of weight 1 are the elements of S, written as $g_j = [g_j]$.
2. The simple commutators of weight $n > 1$ are $[g_1,\ \ldots,\ g_n] = [[g_1,\ \ldots,\ g_{n-1}],\ g_n]$.

We collect some commutator identities which are of utmost importance.

Lemma 1.4 *Let $x,\ y$, and z be elements of a group G.*

(i) $xy = yx[x,\ y]$.
(ii) $x^y = x[x,\ y]$.
(iii) $[x,\ y] = [y,\ x]^{-1}$.
(iv) $[x,\ y]^z = [x^z,\ y^z]$.
(v) $[xy,\ z] = [x,\ z]^y[y,\ z] = [x,\ z][x,\ z,\ y][y,\ z]$.
(vi) $[x,\ yz] = [x,\ z][x,\ y]^z = [x,\ z][x,\ y][x,\ y,\ z]$.
(vii) $\left[x,\ y^{-1}\right] = \left([x,\ y]^{y^{-1}}\right)^{-1}$.
(viii) $\left[x^{-1},\ y\right] = \left([x,\ y]^{x^{-1}}\right)^{-1}$.

Proof

(i) $xy = yx\left(x^{-1}y^{-1}xy\right) = yx[x,\ y]$.
(ii) $x^y = y^{-1}xy = x\left(x^{-1}y^{-1}xy\right) = x[x,\ y]$.
(iii) $[x,\ y] = x^{-1}y^{-1}xy = \left(y^{-1}x^{-1}yx\right)^{-1} = [y,\ x]^{-1}$.
(iv) We have

$$\begin{aligned}
[x,\ y]^z &= z^{-1}\left(x^{-1}y^{-1}xy\right)z \\
&= \left(z^{-1}x^{-1}z\right)\left(z^{-1}y^{-1}z\right)\left(z^{-1}xz\right)\left(z^{-1}yz\right) \\
&= \left(z^{-1}xz\right)^{-1}\left(z^{-1}yz\right)^{-1}\left(z^{-1}xz\right)\left(z^{-1}yz\right) \\
&= [x^z,\ y^z].
\end{aligned}$$

(v) Observe that

$$\begin{aligned}[xy,\ z] &= (xy)^{-1}z^{-1}xyz \\ &= y^{-1}x^{-1}z^{-1}xyz \\ &= y^{-1}\left(x^{-1}z^{-1}xz\right)y\left(y^{-1}z^{-1}yz\right) \\ &= y^{-1}[x,\ z]y[y,\ z] \\ &= [x,\ z]^y[y,\ z] \\ &= [x,\ z][x,\ z,\ y][y,\ z] \text{ by (ii).}\end{aligned}$$

A similar computation gives (vi). By (vi), we have

$$1 = \left[x,\ yy^{-1}\right] = \left[x,\ y^{-1}\right][x,\ y]^{y^{-1}}. \tag{1.3}$$

This establishes (vii), and (viii) follows from (v) in a similar way. □

Lemma 1.5 (The Hall-Witt Identities) *If x, y, and z are elements of a group, then*

$$\left[x,\ y^{-1},\ z\right]^y\left[y,\ z^{-1},\ x\right]^z\left[z,\ x^{-1},\ y\right]^x = 1$$

and

$$\left[x,\ y,\ z^x\right]\left[z,\ x,\ y^z\right]\left[y,\ z,\ x^y\right] = 1.$$

Proof By Lemma 1.4 (iii), we have

$$\begin{aligned}\left[x,\ y^{-1},\ z\right]^y &= y^{-1}\left[\left[x,\ y^{-1}\right],\ z\right]y \\ &= y^{-1}\left[x,\ y^{-1}\right]^{-1}z^{-1}\left[x,\ y^{-1}\right]zy \\ &= y^{-1}\left[y^{-1},\ x\right]z^{-1}\left[x,\ y^{-1}\right]zy \\ &= x^{-1}y^{-1}xz^{-1}x^{-1}yxy^{-1}zy \\ &= \left(xzx^{-1}yx\right)^{-1}yxy^{-1}zy.\end{aligned}$$

Similarly,

$$\left[y,\ z^{-1},\ x\right]^z = \left(yxy^{-1}zy\right)^{-1}zyz^{-1}xz$$

and

$$\left[z,\ x^{-1},\ y\right]^{x} = \left(zyz^{-1}xz\right)^{-1} xzx^{-1}yx.$$

It follows that $\left[x,\ y^{-1},\ z\right]^{y}\left[y,\ z^{-1},\ x\right]^{z}\left[z,\ x^{-1},\ y\right]^{x} = 1$. One can prove the other identity in a similar way. □

1.3 Commutator Subgroups

The notion of the commutator of elements of a group can be generalized to the commutator of subsets of a group.

Definition 1.10 Let G be a group with subset $S = \{s_1,\ s_2,\ \ldots\}$. The subgroup of G *generated* by S, denoted by

$$gp(S) = gp(s_1,\ s_2,\ \ldots),$$

is the smallest subgroup of G containing S. We call S a set of *generators* for $gp(S)$.

The subgroup $gp(S)$ of G can be obtained by taking the intersection of all subgroups of G that contain S. A typical element of $gp(S)$ is of the form

$$s_{i_1}^{\varepsilon_1} s_{i_2}^{\varepsilon_2} \cdots s_{i_n}^{\varepsilon_n},$$

where $s_{i_j} \in S$ and $\varepsilon_j \in \{-1,\ 1\}$ for $1 \leq j \leq n$. If $g \in G$, then $gp(g)$ is just the cyclic subgroup of G generated by g. If $S_1,\ \ldots,\ S_n$ are subsets of G, then the subgroup $gp(S_1 \bigcup \cdots \bigcup S_n)$ is written as $gp(S_1,\ \ldots,\ S_n)$.

Definition 1.11 Let X_1 and X_2 be nonempty subsets of a group G. The *commutator subgroup* of X_1 and X_2 is defined as

$$[X_1,\ X_2] = gp\left([x_1,\ x_2] \mid x_1 \in X_1,\ x_2 \in X_2\right).$$

Thus, $[X_1,\ X_2]$ is the subgroup of G generated by *all* commutators $[x_1,\ x_2]$, where x_1 varies over X_1 and x_2 varies over X_2. In particular, $[G,\ G] = G'$ is the *commutator subgroup* or *derived subgroup* of G.

Remark 1.2 The *set* of all commutators

$$S = \{[x_1,\ x_2] \mid x_1 \in X_1,\ x_2 \in X_2\}$$

does not necessarily form a subgroup of G. For instance, $[x_1,\ x_2]^{-1}$ may not be in S for some $[x_1,\ x_2] \in S$.

If $X_1 = X_2 = G$, then the inverse of every element of S is contained in S by Lemma 1.4 (iii). However, it may be that S is not a subgroup of G because the product of two or more commutators in S is not necessarily a commutator in S. Consider, for example, the *special linear group* $SL_2(\mathbb{R})$ whose elements are the 2×2 matrices with real entries and determinant 1 (the group operation is matrix multiplication). Let I_2 denote the 2×2 identity matrix, and set

$$A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

A routine check shows that $-I_2 = (ABA)^2$,

$$A = \left[\begin{pmatrix} 1 & 0 \\ \frac{4}{3} & 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{pmatrix} \right], \text{ and } B = \left[\begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, \begin{pmatrix} 1 & \frac{4}{3} \\ 0 & 1 \end{pmatrix} \right].$$

Thus, $-I_2$ is a product of commutators. However, $-I_2$ is not the commutator of two elements of $SL_2(\mathbb{R})$. To see this, assume, on the contrary, that $-I_2 = [C, D]$ for some $C, D \in SL_2(\mathbb{R})$. Rewriting this gives $C^{-1}DC = -D$, and thus D and $-D$ are similar matrices. Since the trace of a square matrix equals the trace of any matrix similar to it, D and $-D$ have equal trace. Consequently, the trace of D equals 0. Since the determinant of D equals 1, the characteristic polynomial of D is $f(\lambda) = \lambda^2 + 1$. And so, D has eigenvalues $\pm i$. This means that D is similar to the matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Without loss of generality, we may as well assume that $D = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Suppose that $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Since $CD = -DC$ by assumption, a computation shows that $d = -a$ and $c = b$. Using the fact that C has determinant 1, it follows that $-a^2 - b^2 = 1$. This contradicts the fact that $a, b \in \mathbb{R}$.

Definition 1.11 can be generalized. If $\{X_1, X_2, \ldots\}$ is a collection of nonempty subsets of G, then

$$[X_1, \ldots, X_n] = [[X_1, \ldots, X_{n-1}], X_n],$$

where $n \geq 2$. Note that $[X_1, \ldots, X_n]$ contains all simple commutators of the form $[x_1, \ldots, x_n]$, where $x_1 \in X_1, \ldots, x_n \in X_n$. Thus,

$$[X_1, \ldots, X_n] \geq gp([x_1, \ldots, x_n] \mid x_1 \in X_1, \ldots, x_n \in X_n).$$

However, $[X_1, \ldots, X_n]$ may not equal $gp([x_1, \ldots, x_n] \mid x_1 \in X_1, \ldots, x_n \in X_n)$ if $n \geq 3$. For example (see [6]), consider the cyclic subgroups

$$H_1 = gp(\,(1 \;\; 2)\,),\; H_2 = gp(\,(2 \;\; 3)\,),\; \text{ and } H_3 = gp(\,(3 \;\; 4)\,)$$

of the symmetric group S_4. A routine check confirms that $[H_1, H_2, H_3]$ equals A_4, while $gp([h_1, h_2, h_3] \mid h_1 \in H_1, h_2 \in H_2, h_3 \in H_3)$ equals $gp(\,(1\ \ 3\ \ 4)\,)$. Thus,

$$[H_1, H_2, H_3] \neq gp([h_1, h_2, h_3] \mid h_1 \in H_1, h_2 \in H_2, h_3 \in H_3).$$

Lemma 1.6 *Let G be any group.*

(i) If $H \leq G$ and $[G, G] \leq H$, then $H \trianglelefteq G$ and G/H is abelian. Thus, $[G, G] \trianglelefteq G$ and $G/[G, G]$ is abelian.
(ii) If $N \trianglelefteq G$ and G/N is abelian, then $[G, G] \trianglelefteq N$.

Thus, the commutator subgroup of a group is the smallest normal subgroup inducing an abelian quotient. The factor group $Ab(G) = G/[G, G]$ is called the *abelianization* of G.

Proof

(i) Let $g \in G$ and $h \in H$. By Lemma 1.4 (ii),

$$h^g = g^{-1}hg = h[h, g] \in H$$

because H contains $[G, G]$. Therefore, $g^{-1}Hg = H$, and thus H is normal in G. If g_1H and g_2H are elements of G/H, then

$$(g_1H)(g_2H) = g_1g_2H = g_2g_1[g_1, g_2]H = g_2g_1H = (g_2H)(g_1H)$$

by Lemma 1.4 (i). Therefore, G/H is abelian.

(ii) If $gN, hN \in G/N$, then $(gN)(hN) = (hN)(gN)$. Hence,

$$(gN)^{-1}(hN)^{-1}(gN)(hN) = N.$$

We thus have $g^{-1}h^{-1}gh = [g, h] \in N$. It follows that $[G, G] \trianglelefteq N$. □

Lemma 1.6 allows one to conveniently calculate the derived subgroup. This is illustrated in the next few examples.

Example 1.6 Any two elements of an abelian group G commute. Thus, $[G, G] = 1$.

Example 1.7 We compute the commutator subgroup of the alternating group A_n on the set $S = \{1, 2, \ldots, n\}$. Clearly, $[A_n, A_n] = \{e\}$ for $n = 1, 2, 3$ by Example 1.6.

We find the commutator subgroup of A_4. It is well known that A_4 contains a unique nontrivial normal subgroup

$$K = \{e, (1\ \ 2)(3\ \ 4), (1\ \ 3)(2\ \ 4), (1\ \ 4)(2\ \ 3)\},$$

which is an isomorphic copy of the Klein 4-group (see [1]). Since $[A_4 : K] = 3$, the quotient A_4/K is abelian. Therefore, $[A_4, A_4] \trianglelefteq K$, and thus $[A_4, A_4] = K$.

Lastly, we consider the case when $n \geq 5$. In this case, A_n is simple. Thus, the only normal subgroups of A_n are $\{e\}$ and A_n. Since A_n is not abelian, $[A_n, A_n] = A_n$.

Example 1.8 We find the commutator subgroup of the symmetric group S_n on the set $S = \{1, 2, \ldots, n\}$. By Example 1.6, $[S_n, S_n] = \{e\}$ for $n = 1, 2$.

In order to find $[S_n, S_n]$ for $n \geq 3$, we use the fact that A_n is a normal subgroup of index 2 in S_n, and thus S_n/A_n is an abelian group. First, we find $[S_3, S_3]$. Since S_3/A_3 is abelian, we know that $[S_3, S_3] \trianglelefteq A_3$. Furthermore, each element of A_3 can be written as a commutator of elements in S_3 (this is obvious for the identity permutation):

$$(1\ \ 2\ \ 3) = [(2\ \ 3), (1\ \ 3\ \ 2)] \text{ and } (1\ \ 3\ \ 2) = [(2\ \ 3), (1\ \ 2\ \ 3)].$$

Therefore, A_3 is contained in $[S_3, S_3]$, and consequently, $[S_3, S_3] = A_3$.

Next, we show that $[S_4, S_4] = A_4$. Let $(a\ \ b\ \ c)$ be any 3-cycle for some distinct elements $a, b, c \in S$. This 3-cycle can be written as a commutator of elements in S_4 as

$$(a\ \ b\ \ c) = [(a\ \ b), (a\ \ c\ \ b)].$$

It follows that $A_4 \leq [S_4, S_4]$ because A_4 is generated by 3-cycles. Since S_4/A_4 is abelian, $[S_4, S_4] \trianglelefteq A_4$. We conclude that $[S_4, S_4] = A_4$.

Finally, consider the case when $n \geq 5$. Once again, $[S_n, S_n] \trianglelefteq A_n$ because S_n/A_n is abelian. Since the only nontrivial normal subgroup of S_n is A_n, it must be that $[S_n, S_n] = A_n$.

Example 1.9 We find the derived subgroup of the dihedral group D_n. Recall from Example 1.6 that

$$D_n = \{1, x, x^2, \ldots, x^{n-1}, y, xy, x^2y, \ldots, x^{n-1}y\},$$

where

$$x^n = 1,\ y^2 = 1, \text{ and } xy = yx^{-1}. \tag{1.4}$$

It follows from the last equality in (1.4) that

$$x^r y = yx^{-r} \text{ and } x^r y = (x^r y)^{-1} \quad (r \in \mathbb{Z}). \tag{1.5}$$

Now, it is clear that $[D_1, D_1] = [D_2, D_2] = 1$ by Example 1.6 because D_1 and D_2 are abelian. We claim that $[D_n, D_n] = gp\left(x^2\right)$ for $n \geq 3$.

From this point on, suppose $n \geq 3$ and let r and s denote integers. Choose $x^{2r} \in gp\left(x^2\right)$, and observe that this element can be written as a commutator as follows:

$$x^{2r} = x^r y^{-1} y x^r = x^r y^{-1} x^{-r} y = [x^{-r}, y],$$

where the second equality is a consequence of (1.5). Thus, $gp\left(x^2\right) \le [D_n,\ D_n]$. To prove that $[D_n,\ D_n] \le gp\left(x^2\right)$, we use (1.4) and (1.5). Suppose that $[a,\ b] \in [D_n,\ D_n]$. There are four possible cases for a and b.

- If $a = x^r$ and $b = x^s$, then $[x^r,\ x^s] = 1 \in gp\left(x^2\right)$.
- If $a = x^r$ and $b = x^s y$, then

$$\begin{aligned}[x^r,\ x^s y] &= x^{-r}\left(x^s y\right)^{-1} x^r x^s y = x^{-r} y^{-1} x^{-s} x^r x^s y \\ &= x^{-r} y^{-1} x^r y = x^{-r} y^{-1} y x^{-r} = x^{-2r} \in gp\left(x^2\right).\end{aligned}$$

- Suppose $a = x^r y$ and $b = x^s$. Then

$$[x^r y,\ x^s] = [x^s,\ x^r y]^{-1} \in gp\left(x^2\right)$$

  by the previous case and Lemma 1.4 (iii).
- Suppose $a = x^r y$ and $b = x^s y$. Then

$$\begin{aligned}[x^r y,\ x^s y] &= \left(x^r y\right)^{-1}\left(x^s y\right)^{-1} x^r y x^s y = x^r y x^s y x^r y x^s y \\ &= x^r x^{-s} y y x^r x^{-s} y y = x^{2r-2s} \in gp\left(x^2\right).\end{aligned}$$

It follows that $[D_n,\ D_n] \le gp\left(x^2\right)$. And so, $[D_n,\ D_n] = gp\left(x^2\right)$ for $n \ge 3$ as claimed. In fact, $[D_n,\ D_n] = gp\left(x^2\right) = gp\left(x\right)$ whenever $n \ge 3$ is odd.

Example 1.10 We show that the derived subgroup of the Heisenberg group $\mathscr{H}$ equals its center. By Example 1.5, the center of $\mathscr{H}$ is

$$Z(\mathscr{H}) = \left\{ \begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \;\middle|\; c \in \mathbb{Z} \right\} = gp\left(\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right). \tag{1.6}$$

Let

$$a = \begin{pmatrix} 1 & a_1 & a_2 \\ 0 & 1 & a_3 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } b = \begin{pmatrix} 1 & b_1 & b_2 \\ 0 & 1 & b_3 \\ 0 & 0 & 1 \end{pmatrix}$$

be elements of $\mathscr{H}$. A simple calculation shows that

$$[a,\ b] = \begin{pmatrix} 1 & 0 & a_1 b_3 - b_1 a_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, each commutator of elements of $\mathscr{H}$ is central, and thus $[\mathscr{H},\ \mathscr{H}] \le Z(\mathscr{H})$. In addition, the generator of $Z(\mathscr{H})$ in (1.6) is a commutator of elements of $\mathscr{H}$:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \left[\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \right].$$

It follows that $[\mathscr{H},\ \mathscr{H}] = Z(\mathscr{H})$.

1.3.1 Properties of Commutator Subgroups

We collect several properties of commutator subgroups.

Definition 1.12 Let G be any group, and let S be a nonempty subset of G. The *normalizer* of S in G, denoted by $N_G(S)$, is

$$N_G(S) = \{g \in G \mid gS = Sg\}.$$

If H is a subgroup of G, then $N_G(H)$ is the largest subgroup of G in which H is normal. If K is another subgroup of G, then K *normalizes* H if $K \le N_G(H)$. Clearly, $N_G(H) = G$ if and only if $H \trianglelefteq G$.

Theorem 1.3 *Let G be a group and $H \le G$. Then $C_G(H) \trianglelefteq N_G(H)$ and the factor group $N_G(H)/C_G(H)$ is isomorphic to a subgroup of $Aut(H)$.*

In particular, we obtain Corollary 1.1 when $H = G$.

Proof By Theorem 1.1, the map

$$\varrho : G \to Aut(H) \ \text{ defined by } \ \varrho(h) = \varphi_h, \ \text{ where } \ \varphi_h(g) = g^h,$$

is a homomorphism. Thus, $\varrho|_{N_G(H)}$, the restriction of ϱ to $N_G(H)$, is a homomorphism. It is easy to verify that $\varrho|_{N_G(H)}$ has kernel $C_G(H)$. The result follows from the First Isomorphism Theorem. □

Proposition 1.1 *Let G be any group with subgroups H and K.*

(i) $[H,\ K] = [K,\ H]$.
(ii) $[H,\ K] \le H$ *if and only if K normalizes H. In particular,* $[H,\ G] < H$ *if and only if* $H \trianglelefteq G$.
(iii) *If $H_1 < G$ and $K_1 < G$ such that $H_1 \le H$ and $K_1 \le K$, then* $[H_1,\ K_1] \le [H,\ K]$.

We point out that (i) is valid for any two *subsets* H and K of G.