de Gruyter Expositions in Mathematics 8

de Gruyter Expositions in Mathematics

1 The Analytical and Topological Theory of Semigroups, *K. H. Hofmann, J. D. Lawson, J. S. Pym* (Eds.)

2 Combinatorial Homotopy and 4-Dimensional Complexes, *H. J. Baues*

3 The Stefan Problem, *A. M. Meirmanov*

4 Finite Soluble Groups, *K. Doerk, T. O. Hawkes*

5 The Riemann Zeta-Function, *A. A. Karatsuba, S. M. Voronin*

6 Contact Geometry and Linear Differential Equations, *V. R. Nazaikinskii, V. E. Shatalov, B. Yu. Sternin*

7 Infinite Dimensional Lie Superalgebras, *Yu. A. Bahturin, A. A. Mikhalev, V. M. Petrogradsky, M. V. Zaicev*

Nilpotent Groups and their Automorphisms

by

Evgenii I. Khukhro

Walter de Gruyter · Berlin · New York 1993

Author

Evgenii I. Khukhro

Institute of Mathematics
Siberian Branch of the Russian Academy of Sciences
630090 Novosibirsk – 90, Russia

1991 Mathematics Subject Classification: Primary: 20-01; 20-02; 17-02.
Secondary: 20D15, 20D45, 20E36, 20F18, 20F40, 17B40

Keywords: Nilpotent group, p-group, operator group, Lie ring, commutator, (regular) automorphism, Hughes subgroup

♾ Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

Khukhro, Evgenii I., 1956–
Nilpotent groups and their automorphisms / by Evgenii I. Khukhro.
p. cm. – (De Gruyter expositions in mathematics ; 8)
Includes bibliographical references and index.
ISBN 3-11-013672-4
1. Nilpotent groups. 2. Automorphisms. I. Title. II. Series.
QA177.K48 1993
512'.2–dc20 93-16401
CIP

Die Deutsche Bibliothek – Cataloging-in-Publication Data

Chuchro, Evgenij I.:
Nilpotent groups and their automorphisms / by Evgenii I. Khukhro. – Berlin ; New York : de Gruyter, 1993
(De Gruyter expositions in mathematics ; 8)
ISBN 3-11-013672-4
NE: GT

Printed in Germany.
Disk Conversion: D. L. Lewis, Berlin. Printing: Gerike GmbH, Berlin.
Binding: Lüderitz & Bauer GmbH, Berlin. Cover design: Thomas Bonnie, Hamburg.

Table of Contents

Notation

Preface

In group theory it is natural to distinguish classes of groups according to the extent to which the group operation is commutative. At one extreme we have the class of commutative (abelian) groups and at the other – free groups or groups close to them, as well as nonabelian simple groups, that is, groups without any nontrivial normal subgroups. Commutativity of elements a and b is equivalent to the triviality of their commutator $[a, b] = a^{-1}b^{-1}ab$. So abelian groups are defined by the identity $[x, y] = 1$ and more complicated commutator identities define classes of groups which are close to abelian, but less commutative. The identity

$$[\ldots[[x_1, x_2], x_3], \ldots, x_{c+1}] = 1$$

defines the variety of nilpotent groups of class c, and the identity $\delta_k = 1$ of 2^k variables, where δ_k is defined recursively by

$$\delta_1 = [x_1, x_2], \quad \delta_{k+1} = [\delta_k(x_1, \ldots, x_{2^k}), \delta_k(x_{2^k+1}, \ldots, x_{2^{k+1}})],$$

defines the variety of soluble groups of derived length k. Another way to define these classes of groups is in terms of the existence of a series of normal subgroups with central or commutative factors, respectively.

Study of nilpotent groups often aims to prove that they possess some degree or other of commutativity. For example, the positive solution of the Restricted Burnside Problem for groups of exponent p^k (Kostrikin, 1959, for $k = 1$, and Zel'manov, 1990, for all k) means that there is a function $f(p, k, m)$ such that the nilpotency class of any m-generated finite p-group of exponent p^k does not exceed $f(p, k, m)$.

The fact that nilpotent groups are close to being commutative means that it is possible to apply linear methods to their study. Using the group operations one can define the structure of a Lie ring on the direct sum of the factors of the lower central series. The action of a group on an invariant commutative section looks like the action of a matrix group on a vector space. However, although for such more linear objects as Lie rings or matrix rings powerful results and highly developed techniques are available, there may be difficulties in using them when it comes to going from groups to rings and back again.

In this book linear methods in the theory of nilpotent groups are applied to the study of automorphisms of nilpotent groups. We prove the analogue of the

positive solution of the Restricted Burnside Problem for groups with a splitting automorphism of prime order. This gives rise to a structural theory of finite p-groups admitting a partition which includes the positive solution of the Hughes problem for almost all (in some precise sense) finite p-groups. The Higman-Kreknin-Kostrikin Theorem on the boundedness of the nilpotency class of Lie rings (or nilpotent groups) admitting a regular automorphism of prime order, is generalized to the case where the number of fixed points is finite: almost regularity of the automorphism of prime order implies almost nilpotency – that is, the existence of a nilpotent subring (or subgroup) of bounded index and of bounded nilpotency class. Kreknin's Theorem on the solubility of Lie rings with regular automorphisms of finite order is used to prove the "almost solubility", in an analogous sense, of a nilpotent p-group with an almost regular automorphism of order p^k. Linear and combinatorial methods are used to prove a theorem of a rather general nature which gives a positive solution to the Restricted Burnside Problem for a variety of operator groups under the hypothesis that this problem has a positive solution for the ordinary variety obtained from this variety by replacing all operators by 1 in its identities.

The first part "Linear methods" is, in fact, a textbook. The existence of many books or chapters of books devoted to nilpotent groups – for instance, those of Baumslag [6], Gorenstein [25], M. Hall [27], P. Hall [28], Huppert and Blackburn [49], Kargapolov and Merzlyakov [51], Kurosh [85] and Warfield [153] – makes it a difficult task to write something new. We have tried to select only that material which is necessary for the exposition of the aforementioned results from the second part of the book, "Automorphisms". Practically every result proved in the first part is in some way used in the second: either there is a reference to it or to its proof, or its statement and proof prepares the reader for more complicated arguments of a similar nature in the second part. But, of course, we have not followed this rule too strictly in the hope that the reader may get to know at least some of the classical methods in the theory of nilpotent groups. However, many important results are only briefly mentioned; for example, Kostrikin's Theorem on Engel Lie algebras is stated here without proof. On the other hand, in the interests of completeness, we reproduce proofs of the theorems of Higman, Kreknin and Kostrikin on regular automorphisms.

The author hopes that the second part is something more than a collection of several research papers under one cover. The material is presented here with more detail and perhaps more intelligibly than originally. Moreover, some proofs are longer in an attempt to make the book more self-contained. Repetition of typical arguments should help with their mastery.

This book is based on a special course given at Novosibirsk University in 1988-90. The author thanks Andrei Vasil'ev and Natasha Makarenko for several valuable remarks.

The book consists of seven chapters.

Chapter 1, as well as containing definitions, notation and some basic facts from group theory and ring theory, also contains the proofs of some useful results. Among them are Higman's Lemma and several theorems about fixed points of automorphisms.

Chapter 2 is devoted to nilpotent groups. It deals with several types – from torsion-free groups to finite p-groups. Many of the results here are, of course, based on commutator calculations.

Chapter 3 introduces associated Lie rings. Here we prove the Magnus-Sanov Theorem on the $(p-1)$-Engel condition for the associated Lie ring of a group of prime exponent p. We also prove the nilpotency of soluble groups of prime exponent. The first section of the chapter contains statements of some Lie ring analogues of group theoretic results.

Chapter 4 opens the second part of the book. First, we prove theorems of Higman, Kreknin and Kostrikin about regular automorphisms of Lie rings. Then the author's theorem about Lie rings admitting an almost regular automorphism of prime order p is proved: if the number of fixed points (dimension of the subspace) is finite then the Lie ring contains a subring of finite index (codimension) which is nilpotent of p-bounded class.

In Chapter 5 theorems on nilpotent groups with regular or almost regular automorphisms of prime order are proved using the results of Chapter 4. Among these results is also a "modular" theorem, dealing with a finite p-group with an automorphism of order p. (The latest results of Shalev and the author on p-groups admitting automorphisms of order p^k with few fixed points are contained in Chapter 8.) We have included also theorems of Makarenko, who has refined estimates for the nilpotency class of subgroups, and of Medvedev, who has generalized our theorem on periodic nilpotent groups with an almost regular automorphism of prime order to the case of an arbitrary nilpotent group. The author is grateful to Makarenko, Medvedev and Shalev for providing proofs of their theorems.

Chapter 6 contains two theorems of a rather general nature concerned with bounding nilpotency classes of nilpotent groups from some varieties of groups with operators. Let Ω be a group and let $\{v_\alpha\}$ be a family of Ω-identities, defining the variety of operator groups $\mathfrak{M}$. We denote by $\{\bar{v}_\alpha\}$ the family of (ordinary) group identities obtained from $\{v_\alpha\}$ by replacing all operators from Ω by 1 and by $\overline{\mathfrak{M}}$ the variety of groups defined by identities $\{\bar{v}_\alpha\}$. First, suppose that there is a constant c bounding the nilpotency class of any nilpotent group in $\overline{\mathfrak{M}}$. The first theorem states that, under this condition, if for an Ω-group $G \in \mathfrak{M}$ the semidirect product $G \leftthreetimes \Omega$ is nilpotent then the nilpotency class of G is bounded by the same number c. The author's theorem from [56] on the nilpotency of a soluble group G with a splitting automorphism φ of prime order p, that is, such that

$$x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1$$

for all $x \in G$, is a prototype of this result. In Chapter 6 this theorem from [56] is deduced from the more general theorem mentioned above; it is related to the structural theory of finite p-groups with a partition (in other terms, finite p-groups which "split") and is applied in Chapter 7. A further consequence is a bound in terms of p only for the nilpotency class of the commutator subgroup of a finite p-group of maximal class.

The second theorem of Chapter 6 gives a positive solution to the Restricted Burnside Problem for a variety of operator groups $\mathfrak{M}$ provided that this problem has a positive solution for the corresponding ordinary variety $\overline{\mathfrak{M}}$. More exactly, suppose that the Restricted Burnside Problem has a positive solution for the variety $\overline{\mathfrak{M}}$ in the sense that locally nilpotent groups from $\overline{\mathfrak{M}}$ constitute a subvariety and, moreover, that the associated Lie ring of a free group of $\overline{\mathfrak{M}}$ satisfies a system of multilinear identities which defines a locally nilpotent variety of Lie rings with a function $f(d)$ bounding the nilpotency class of a d-generated ring. It is proved that if for an Ω-group $G \in \mathfrak{M}$ the semidirect product $G \leftthreetimes \Omega$ is locally nilpotent then the group G belongs to a locally nilpotent variety in which the nilpotency class of a d-generated group is bounded by the function $f\left(d\frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}\right)$. (An example shows that the word "multilinear" in the hypothesis of the theorem is essential.)

One of the main tools in the proofs is Higman's Lemma, proved in Chapter 1.

Note that in both theorems the strong condition that the semidirect product $G \leftthreetimes \Omega$ is (locally) nilpotent is automatically satisfied if both G and Ω are (locally) finite p-groups.

The main theorem from [61], which bounds the nilpotency class of a d-generated nilpotent group with a splitting automorphism of prime order p, is a prototype of the second theorem. This theorem from [61] is the basis for the structural theory of finite p-groups admitting a partition. This theory, expounded in Chapter 7, includes a positive solution of the Hughes problem for almost all finite p-groups – in spite of the fact that there exist counterexamples to the Hughes conjecture. In Chapter 7 we give also the original proof of the theorem from [61], which yields some additional information and illustrates some aspects of Lie ring technique. This proof is reduced by known results to the case of finite p-groups. It uses Kostrikin's Theorem on $(p-1)$-Engel Lie algebras and generalizations of Higman's Theorem on regular automorphisms of prime order to the case of a p-group with an automorphism of order p from Chapter 5. The main technical lemma is an analogue of the Magnus-Sanov Theorem.

The results of Chapters 5 and 7 give, in a certain sense, a complete picture of the structure of nilpotent groups with automorphisms of prime order close to regular (splitting or almost regular). Much less is known about nilpotent groups with such automorphisms of composite order (in contrast to the progress made in the study of finite groups with almost regular automorphisms modulo the structure of nilpotent sections).

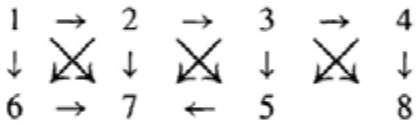
Chapter 8 contains the first major breakthrough in this direction. In the "modular" case where a nilpotent p-group P admits an automorphism of order p^k with p^m fixed points, it is proved that P is almost soluble with a strong bound, in terms of p and k only, on the derived length of a subgroup of bounded index. The proof is based on Kreknin's Theorem on Lie rings from Chapter 4. It uses a group-theoretic corollary to Kreknin's Theorem, obtained with the help of the Mal'cev correspondence given by the Baker-Hausdorff formula, and some techniques from the theory of powerful p-groups, especially, from Shalev's work [129], where a weak bound, in terms of p, k and m, for the derived length of P was obtained.

The necessary preliminary material on the Mal'cev correspondence and on powerful p-groups is included in Chapter 8 without proofs.

Chapters 4–8 contain comments on the present state of the theory, including unsolved problems and connections with other areas.

All propositions, formulae and particular statements referred to subsequently, are numbered by triples a.b.c, where a is the number of the chapter and b is the number of the section.

The interdependence of the chapters is described by the following scheme:



This book was written when the author was in Freiburg (FRG) as a research fellow of the Alexander von Humboldt Foundation. The author thanks the Foundation for its generous support in providing the opportunity to do research work in Germany which facilitated the writing of this book.

The author is also grateful to the Mathematics Institute of the Albert-Ludwigs-University of Freiburg and, in particular, to Professor O.H. Kegel, for their help and hospitality.

The author thanks Professor J.C. Lennox (Cardiff) for his help in language polishing and for many valuable suggestions improving the style of the book.

The author is grateful to the Publishers, Verlag de Gruyter, and especially to Dr. M. Karbe, for their friendly, patient and fast work at all stages from the manuscript to the final galley proofs.

E.I. Khukhro

Part I
Linear Methods

Chapter 1

Preliminaries

We assume that the reader is acquainted with the basic notions of linear algebra, group theory and ring theory at the level of the first 2-3 years of a university course in higher algebra. In particular, we assume as known the notions of group, subgroup, permutation group, automorphism, free group, verbal subgroup, Sylow subgroup, ring, subring and ideal. We also assume the homomorphism theorems and the theorems on the structure of cyclic groups and of finitely generated abelian groups. We take as known also the definition of soluble groups (though the definition and main properties of nilpotent groups are given for completeness sake in Chapter 2). If necessary, this material may be found in university textbooks on higher algebra, for example [77, 86, 105], and in textbooks on group theory, for example [51, 85, 27].

In this chapter we fix some notation and recall some definitions and known facts. A few useful propositions will be proved. Among them are Higman's Lemma and some theorems on fixed points of automorphisms.

The sets of all complex, real, rational, integer and natural numbers are denoted by $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{Z}$, $\mathbb{N}$, respectively.

We shall say that a quantity depending on n is n-bounded (or bounded in terms of n), if it is bounded by some function depending only on n; the analogous expression – $(n_1, n_2, \ldots)$-boundedness (or boundedness in terms of $n_1, n_2, \ldots$) – will also be used where several parameters are involved.

§ 1.1 Groups

The group, generated by a set M, is denoted by $\langle M \rangle$; if M is given by some condition P, that is $M = \{x|\ P(x)\}$, then we write $\langle M \rangle = \langle x|\ P(x) \rangle$.

By $a^g = g^{-1}ag$ we denote the conjugate of the element a under the element g.

For subsets M and N of a group we write

$$M^N = \{m^n|\ m \in M,\ n \in N\}; \quad M \cdot N = \{m \cdot n|\ m \in N,\ n \in N\}.$$

In particular $\langle M^G \rangle$ and $\langle a^G \rangle$ denote the normal closures of the subset M and of the element a in a group G, that is, the smallest normal subgroups of G containing M and a, respectively.

The fact that A is a normal subgroup of a group G is denoted by $A \trianglelefteq G$; if a subgroup does not coincide with the whole group, we say that it is a proper subgroup and use the strict inclusion signs for subgroups and normal subgroups, $<$ and $\triangleleft$, respectively.

By $G = B \leftthreetimes A$ we denote the semidirect product G of groups B and A: that is

$$G = B \cdot A, \quad B \trianglelefteq G \quad \text{and} \quad B \cap A = 1.$$

A subgroup of a group is said to be characteristic if it is invariant under all automorphisms of the group. For example, if the Sylow p-subgroup is normal, then it is unique and therefore characteristic. Since a characteristic subgroup is also invariant under all inner automorphisms, that is, automorphisms induced by conjugation by the elements of the group, it is also normal. A characteristic subgroup of a normal subgroup is normal in the whole group, and a characteristic subgroup of a characteristic subgroup is characteristic in the whole group.

If $N \trianglelefteq G$, we use congruences in a group G modulo N to indicate equality of images of elements or subsets in the factor-group G/N:

$$a \equiv b \pmod{N} \Leftrightarrow aN = bN; \quad H \equiv K \pmod{N} \Leftrightarrow HN = KN.$$

The subsets

$$N_G(M) = \{g \in G \mid M^g = M\}$$

and

$$C_G(M) = \{g \in G \mid gm = mg \text{ for all } m \in M\}$$

are subgroups called, respectively, the normalizer and the centralizer of the subset M in the group G; if the subset $M = \{m\}$ consists of only one element, then we write $C_G(m)$ instead of $C_G(\{m\})$.

A section of a group G is a factor-group M/N where M and N are arbitrary subgroups of G and, of course, $N \trianglelefteq M$. A section is said to be normal if both subgroups are normal: $M \trianglelefteq G$, $N \trianglelefteq G$.

By $[a, b] = a^{-1}b^{-1}ab$ we denote the commutator of the elements a and b in a group. The mutual commutator subgroup $[M, N]$ of arbitrary subsets M and N of a group is defined by

$$[M, N] = \langle [m, n] \mid m \in M, \ n \in N \rangle.$$

It is useful to know that $[M, N] \trianglelefteq \langle M, N \rangle$. We denote the commutator or derived subgroup of the group G by $G' = G^{(1)} = [G, G]$. The members of the derived

series of G are defined by induction:

$$G'' = G^{(2)} = [G', G'], \quad G^{(s+1)} = [G^{(s)}, G^{(s)}].$$

We denote by $\gamma_i(G)$ the members of the lower central series of G:

$$\gamma_1(G) = G, \quad \gamma_{s+1}(G) = [\gamma_s(G), G].$$

We denote the centre of G by $Z(G)$:

$$Z(G) = \{z \in G |\ zg = gz \text{ for all } g \in G\},$$

which is, of course, a normal subgroup of G. By $\zeta_i(G)$ we denote the members of the upper central series of G: $\zeta_1(G) = Z(G)$, and inductively $\zeta_{s+1}(G)$ is the full inverse image in G of the centre $Z(G/\zeta_s(G))$ of the factor-group $G/\zeta_s(G)$.

We define (composite) commutators in the elements of a set X (as formal expressions) inductively by their weight: the elements of X are the commutators of weight 1; if c_1 and c_2 are commutators of weights r_1 and r_2 in elements of X, then $[c_1, c_2]$ is a commutator in elements of X of weight $r_1 + r_2$. We also define the set of elements occurring in the commutator, with appropriate multiplicities, and the set of subcommutators: for commutator of weight 1 both sets consist of the commutator itself as the only element; in the commutator $[c_1, c_2]$ exactly those elements of X occur, which occur either in c_1 or in c_2, their multiplicities being summed, and the set of subcommutators of $[c_1, c_2]$ is the union of the sets of subcommutators of c_1 and c_2 together with c_1 and c_2 themselves. We also say that a commutator c of the elements of X has weight w in the variable $x \in X$, if x occurs in c with multiplicity w.

The commutator

$$[\ldots[[a_1, a_2], a_3], \ldots, a_k]$$

is called simple and is denoted by $[a_1, a_2, \ldots, a_k]$.

A group G is said to have finite exponent (or period) n, if $g^n = 1$ for all $g \in G$. By $G^n = \langle g^n |\ g \in G\rangle$ we denote the subgroup of G, generated by the n-th powers of all of its elements. It is clear, that this is a characteristic subgroup of G and is the smallest normal subgroup of G for which the factor-group has exponent n.

A group G is said to be a periodic group (or torsion group) if each of its elements has finite order. If, in addition, each element has order a power of some fixed prime number p, then G is said to be a p-group, and if the orders of all elements are not divisible by p, then G is said to be a p'-group. If P is a p-group, $\Omega_i(P)$ denotes the subgroup $\langle g \in P |\ g^{p^i} = 1\rangle$.

The minimal number of generators of a finite abelian group is called its rank.

An abelian group of prime exponent p is said to be elementary; it may be regarded as a vector space over the finite field $GF(p) \cong \mathbb{Z}/p\mathbb{Z}$ of order p: vector

addition is the group operation, and multiplying by the residue i modulo p is equivalent to taking the i-th power. The automorphisms of the group are the linear transformations of this vector space.

A chain of nested subgroups

$$1 \leq K_1 \leq K_2 \leq \ldots \leq K_n = G$$

is called a series of G of length n. A series is said to be subnormal, if $K_i \trianglelefteq K_{i+1}$ for all i; it is said to be normal, if $K_i \trianglelefteq G$ for all i. The factor-groups K_{i+1}/K_i of a subnormal series are also called its factors.

Although, as a rule, the group operation is denoted by the multiplication sign (which is often omitted) and the identity element by 1, in the case of commutative groups sometimes additive notation is used: $+$ for the group operation and 0 for the identity element.

The symmetric group of all permutations of n symbols is denoted by $\mathbb{S}_n$.

Schreier's Theorem states that a subgroup of finite index m in a finitely generated group generated by n elements is itself finitely generated by an (m, n)-bounded number of elements.

Let π be a set of prime numbers. Its complement in the set of all primes is denoted by π'. A subgroup of a finite group G, whose order is divisible only by primes from π and its index only by primes from π', is called a Hall π-subgroup of G.

A local covering of a group G is a system of its subgroups $\{H_\alpha | \ \alpha \in \mathrm{A}\}$ such that $G = \bigcup_\alpha H_\alpha$, and for any two subgroups H_α, H_β, $\alpha, \beta \in \mathrm{A}$, there is a subgroup H_γ, $\gamma \in \mathrm{A}$, containing them both: $H_\alpha, H_\beta \subseteq H_\gamma$. Mal'cev's Local Theorem asserts that if all of the subgroups of a local covering of a group satisfy some group-theoretic property, which may be expressed by a quasiuniversal formula in predicate calculus, then the group itself also satisfies this property.

A group G is said to be residually finite if it has a system of normal subgroups $\{N_\alpha\}$ such that all of the factor-groups G/N_α are finite and $\bigcap_\alpha N_\alpha = 1$ (or, equivalently: for each element $g \in G$ there is a normal subgroup $N(g)$ of finite index in G such that $g \notin N(g)$).

§ 1.2 Rings and modules

The finite field of order $q = p^k$, where p is a prime, is denoted by $GF(q) = GF(p^k)$.

A module over a commutative ring K with identity (a K-module) is an additive group M which admits multiplication by the elements of K, satisfying the following

axioms:

$$1a = a, \quad 1 \in K, \ a \in M;$$
$$(\alpha + \beta)a = \alpha a + \beta a, \quad \alpha, \beta \in K, \ a \in M;$$
$$\alpha(a + b) = \alpha a + \alpha b, \quad \alpha \in K, \quad a, b, \in M;$$
$$(\alpha\beta)a = \alpha(\beta a), \quad \alpha, \beta \in K, \ a \in M$$

Thus modules over a field k are precisely the vector spaces over k. Every abelian group may be regarded as a $\mathbb{Z}$-module in a natural way: for $k > 0$

$$ka = \underbrace{a + a + \ldots + a}_{k}, \quad (-k)a = k(-a) \quad \text{and} \quad 0a = 0.$$

The elements $m_1, m_2, \ldots, m_s$ are said to generate the K-module M if each element $m \in M$ may be expressed in the form $m = \sum_i k_i m_i$, where $k_i \in K$. Every s-generated K-module is a homomorphic image of the free s-generated K-module

$$e_1 K \oplus e_2 K \oplus \ldots \oplus e_s K,$$

where for each i the abelian group $e_i K = \{e_i k | \ k \in K\}$ is isomorphic to K, and $k'(e_i k) = e_i(k'k)$ for all $k', k \in K$. In particular, the ring K may be regarded as a free 1-generated K-module with generator $e_1 = 1$.

Let K be a commutative ring with identity and G a group. The group ring

$$KG = \left\{ \sum_g k_g g | \ g \in G, \ k_g \in K \right\}$$

has as its additive group the free K-module whose free generators are the elements of G, and multiplication is defined naturally via the group operation and the distributivity law.

If G is a group of automorphisms of an abelian group V (or if G acts as a group of automorphisms on an abelian group V – see the definition in §1.5), then V may be regarded as a $\mathbb{Z}G$-module:

$$v\left(\sum_g k_g g\right) = \sum_g k_g (v^g).$$

In an analogous way, if G is a group of linear transformations of a vector space V over a field k, then V may be regarded as a kG-module.

We recall here the way in which Maschke's Theorem generalizes to the case of an arbitrary $\mathbb{Z}G$-module. If G is a finite group and V is a $\mathbb{Z}G$-module such

that extraction of unique p-th roots is possible in the additive group of V for all of the prime divisors p of the order of G, then every $\mathbb{Z}G$-submodule U (which is G-invariant by definition) has a direct complement W which is also a $\mathbb{Z}G$-submodule – that is, $V = U \oplus W$, where W is G-invariant. The condition on the additive group of V is automatically satisfied if V is finite and its order is coprime to the order of the group G.

Let A and B be K-modules. Their tensor product $A \otimes_K B$ is defined as the factor-module of the free K-module with free generators $a \otimes b$, $a \in A$, $b \in B$, by the submodule generated by all elements of the form

$$k(a \otimes b) - ka \otimes b, \quad ka \otimes b - a \otimes kb,$$
$$a \otimes (b_1 + b_2) - (a \otimes b_1 + a \otimes b_2),$$
$$(a_1 + a_2) \otimes b - (a_1 \otimes b + a_2 \otimes b).$$

where $k \in K$; $a, a_1, a_2 \in A$; $b, b_1, b_2 \in B$. This is equivalent to taking the set of all formal sums

$$\left\{ \sum k_{a,b} a \otimes b \mid a \in A, \ b \in B \right\}$$

and identifying the elements:

$$k(a \otimes b) = ka \otimes b = a \otimes kb,$$
$$a \otimes (b_1 + b_2) = a \otimes b_1 + a \otimes b_2,$$
$$(a_1 + a_2) \otimes b = a_1 \otimes b + a_2 \otimes b.$$

A mapping $\vartheta\colon A \times B \to C$ induces a homomorphism of the K-module $A \otimes_K B$ into the K-module C by the rule $a \otimes b \to \vartheta(a, b)$ if and only if the following equalities hold:

$$\vartheta(ka, b) = \vartheta(a, kb) = k\vartheta(a, b),$$
$$\vartheta(a, b_1 + b_2) = \vartheta(a, b_1) + \vartheta(a, b_2),$$
$$\vartheta(a_1 + a_2, b) = \vartheta(a_1, b) + \vartheta(a_2, b).$$

If the elements $a_1, a_2, \ldots, a_s$ generate a K-module A and the elements $b_1, b_2, \ldots, b_t$ generate a K-module B, then the st elements $a_i \otimes b_j$ generate the K-module $A \otimes_K B$.

If the K-modules $M = \bigoplus_i M_i$ and $N = \bigoplus_j N_j$ are decomposable into direct sums of K-submodules M_i and N_j, then the tensor product $M \otimes_K N$ is decomposable into the direct sum

$$M \otimes_K N = \bigoplus_{i,j} (M_i \otimes_K N_j)$$

of K-submodules $M_i \otimes_K N_j$.

Corollary. *Let ω be a primitive n-th root of unity. If A is a subgroup of an abelian group B, then*

$$(A \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]) \cap B \otimes 1 = A \otimes 1,$$

where $A \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ is regarded as naturally embedded into $B \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$.

Proof. This follows from the previous assertion and from the fact, that

$$\mathbb{Z}[\omega] = \mathbb{Z} \oplus \mathbb{Z}\omega \oplus \mathbb{Z}\omega^2 \oplus \ldots \oplus \mathbb{Z}\omega^{f(n)-1},$$

where $f(n)$ is Euler's function, so that

$$B \otimes_{\mathbb{Z}} \mathbb{Z}[\omega] = \bigoplus_{i=0}^{f(n)-1} (B \otimes \mathbb{Z}\omega^i) \quad \text{and} \quad A \otimes_{\mathbb{Z}} \mathbb{Z}[\omega] = \bigoplus_{i=0}^{f(n)-1} (A \otimes \mathbb{Z}\omega^i).$$

Note that if an abelian group A has exponent n, then, for any abelian group B, the tensor products $A \otimes B$ and $B \otimes A$ also have exponent n. For example, the abelian group $A \otimes B$ is generated by the elements $a \otimes b$, $a \in A$, $b \in B$, and we have

$$n(a \otimes b) = na \otimes b = 0 \otimes b = 0.$$

Thus, in particular, if abelian groups A and B have coprime exponents m and n, respectively, then $A \otimes B = 0$, since its exponent divides both m and n.

Tensor products are used to extend the ground ring of a module (or a vector space, or any K-algebra). Let A be a K-module and suppose that K is a subring of a ring L. Then L is also a K-module under natural multiplication by elements of K, and one can form the K-module $A \otimes_K L$. This module may also be regarded as an L-module by putting $l_1(a \otimes l_2) = a \otimes l_1 l_2$, where $l_1, l_2 \in L$, $a \in A$. Note that if k is a subring of K, then the tensor product $A \otimes_k L$ is isomorphic as an L-module to the L-module $A \otimes_K L$.

§ 1.3 Lie rings

A Lie ring is a nonassociative ring without identity, whose multiplication, which is usually denoted by brackets $[a, b]$, satisfies the following axioms:

$$[a, a] = 0$$

(anticommutativity)

$$[[a, b], c] + [[b, c], a] + [[c, a], b] = 0$$
(the Jacobi identity).

From anticommutativity and the usual distributivity laws we obtain the identity $[a, b] = -[b, a]$ (indeed, $[a + b,\ a + b] = 0 \Rightarrow [a, a] + [a, b] + [b, a] + [b, b] = 0 \Rightarrow [a, b] + [b, a] = 0$). It is not difficult to deduce from this that for Lie rings the notions of left, right and two-sided ideals coincide.

From the Jacobi identity it is not difficult to deduce that, if I and J are the ideals of a Lie ring, then the additive subgroup, generated by the set of commutators

$$\{[a, b] | \ a \in I,\ b \in J\}$$

is also an ideal of the Lie ring. We denote it by $[I, J]$.

If A and B are subsets of a Lie ring, then $A + B$ will denote the subset

$$A + B = \{a + b | \ a \in A,\ b \in B\}.$$

If A and B are subrings then $A + B$ is also a subring, and if A and B are ideals, then $A + B$ is also an ideal.

Ideals of Lie rings play the same role as normal subgroups in groups – they are kernels of homomorphisms. To indicate that I is an ideal of a Lie ring L we use normal subgroup notation: $I \trianglelefteq L$. The following theorems on homomorphisms of Lie rings are analogous to the homomorphism theorems for groups.

Let L be a Lie ring and N an ideal of L. Then

a) there is a one-to-one correspondence between the set of all subrings of L containing N and the set of all subrings of the factor-ring L/N; this correspondence is given by passing to the images of subrings in the factor-ring L/N. Also, a subring of L containing N is an ideal of L if and only if its image in the factor-ring L/N is an ideal of L/N;

b) if N is the kernel of a homomorphism φ of the Lie ring L then there is an isomorphism $L/N \cong L^{\varphi}$, and, moreover, φ is the composition of the natural homomorphism $L \to L/N$ and an isomorphism of the Lie rings L/N and L^{φ};

c) if A is an ideal of the Lie ring L containing N then there is a Lie ring isomorphism $(L/N)/(A/N) \cong L/A$;

d) if A and B are subrings of the Lie ring L and A is an ideal of B, then there is an isomorphism

$$(B + N)/(A + N) \cong B/(A + (B \cap N)).$$

In particular,

$$(B + N)/N \cong B/(B \cap N).$$

Commutators in elements of a subset Y of a Lie ring and their weights are defined and denoted exactly in the same way as group commutators (see § 1.1).

The ideal of a Lie ring L generated by the set X is denoted by ${}_{id}\langle X\rangle$. The additive group of ${}_{id}\langle X\rangle$ is generated by simple commutators of the form $[\ldots[[x, a_1], a_2], \ldots, a_k]$, $x \in X$, $a_i \in L$, $k \geq 0$.

The additive subgroup generated by the set X is denoted by ${}_{+}\langle X\rangle$, and the subring (or Lie ring) generated by X is denoted by $\langle X\rangle$.

If $L = \langle X\rangle$ is a Lie ring generated by the set X then $L = \sum_i L_i$, where L_i is the homogeneous component of L of weight i (with respect to X), that is, L_i is the additive subgroup, generated by all commutators of weight i in elements of X. In an analogous way, if $L = \langle x_1, x_2, \ldots\rangle$, then $L = \sum L_{i_1,i_2,\ldots}$, where $L_{i_1,i_2,\ldots}$ is the multihomogeneous component of weight i_1 in x_1, of weight i_2 in x_2 , etc., that is, $L_{i_1,i_2,\ldots}$ is the additive subgroup of L generated by all commutators in elements $x_1, x_2, \ldots$ of weight i_1 in x_1, of weight i_2 in x_2, etc. An ideal or an additive subgroup I of a Lie ring $L = \langle X\rangle$ is said to be homogeneous (multihomogeneous) with respect to the generating set X if

$$I = \oplus I \cap L_i \ (I = \oplus I \cap L_{i_1,i_2,\ldots}).$$

The intersections $I \cap L_i$ $(I \cap L_{i_1,i_2,\ldots})$ are called the homogeneous (multihomogeneous) components of weight i (of multiweight $(i_1, i_2, \ldots)$) of a homogeneous (multihomogeneous) ideal (or additive subgroup) I.

The members of the lower central series, $\gamma_i(L)$, of a Lie ring L are defined by induction as follows:

$$\gamma_1(L) = L, \quad \gamma_{s+1}(L) = [\gamma_s(L), L]$$

(sometimes they are also denoted by $L^i = \gamma_i(L)$); the members of the derived series are also defined by induction:

$$L' = \gamma_2(L) = L^2, \quad L'' = L^{(2)} = [L', L'], \quad L^{(s+1)} = [L^{(s)}, L^{(s)}].$$

It is easy to see that $L^{(s)} \subseteq \gamma_{2^s}(L)$ for all $s \in \mathbb{N}$.

It is also easy to see that, if $L = \langle X\rangle$, then $\gamma_k(L) = \sum_{i=k}^{\infty} L_i$, where L_i is the homogeneous component of the Lie ring L of weight i (with respect to X).

If a Lie ring L is also a K-module where multiplication by any element of K is a homomorphism of L, then L is said to be a K-Lie algebra. If K is a subring of R, then the R-module $L \otimes_K R$ may be regarded in a natural way as an R-Lie algebra with Lie multiplication given by:

$$[l_1 \otimes r_1, l_2 \otimes r_2] = [l_1, l_2] \otimes r_1 r_2.$$

As a K-Lie algebra, L is isomorphically embedded into the K-Lie algebra $L \otimes_K R$ by the mapping $l \to l \otimes 1$. It follows easily from the definitions that

$$\gamma_i(L \otimes_K R) = \gamma_i(L) \otimes_K R \quad \text{and} \quad (L \otimes_K R)^{(s)} = L^{(s)} \otimes_K R;$$

in particular, $\gamma_i(L \otimes_K R) = 0 \Leftrightarrow \gamma_i(L) = 0$.

Let G be an abelian group. A Lie ring L is said to have a G-grading or to be G-graded, if to each element $g \in G$ there corresponds a subgroup L_g of the additive group of L such that

$$L = \sum_{g \in G} L_g \quad \text{and} \quad [L_a, L_b] \subseteq L_{a+b} \text{ for all } a, b \in G.$$

For example, if L is a $\mathbb{C}$-Lie algebra, and φ is an automorphism of L of finite order n, then the additive group of L decomposes as a $\mathbb{C}$-vector space into the direct sum of the eigenspaces of the linear transformation φ:

$$L = L_0 \oplus L_1 \oplus \ldots \oplus L_{n-1},$$

where $L_i = \{l \in L \mid l^\varphi = \omega^i l\}$, $i = 0, 1, \ldots, n-1$, and ω is a primitive n-th root of unity. Here we have $[L_a, L_b] \subseteq L_{a+b}$ where $a, b, a+b$ are residues modulo n, so that this decomposition gives a $\mathbb{Z}/n\mathbb{Z}$-grading of L.

The facts of the theory of varieties, common to arbitrary algebraic systems, hold also for Lie rings. There exist free n-generator Lie rings whose elements are linear combinations of commutators in the free generators. More precisely, the free Lie ring $F = \langle x_1, x_2, \ldots \rangle$, freely generated by elements $x_1, x_2, \ldots$, decomposes into the direct sum $F = \oplus F_i$ of its homogeneous components, and also decomposes into the direct sum $F = \oplus F_{i_1, i_2, \ldots}$ of its multihomogeneous components. Two linear combinations of commutators in the free generators are equal only if one may be transformed into another by applications of the Jacobi identity and the anticommutativity and distributivity laws. A more constructive description is given by the theorem which says that the additive group of F is a free abelian group freely generated by the so-called basic commutators in the free generators (the definition of basic commutators will be given in § 2.7 for groups – it is exactly the same for Lie rings). However, we do not need this theorem, although we note that we shall, in fact, prove in § 2.7 that the basic commutators in the generators really generate the additive group of a Lie ring.

It is quite straightforward to see that the additive group of the Lie ring F is generated by simple commutators in the generators. Indeed, it is clear that it is generated by commutators in the generators; so it is sufficient to express every commutator as a linear combination of simple commutators. By an obvious induc-

tion on commutator weight, it suffices to consider commutators of the form

$$[[x_{i_1}, x_{i_2}, \ldots, x_{i_r}], [x_{j_1}, x_{j_2}, \ldots, x_{j_s}]].$$

We proceed by induction on s. For $s = 1$ such commutators are simple, and for $s > 1$ we have by the Jacobi identity

$$\begin{aligned}[[x_{i_1}, x_{i_2}, \ldots, x_{i_r}]], [x_{j_1}, x_{j_2}, \ldots, x_{j_s}] = \\ = [[[x_{i_1}, x_{i_2}, \ldots, x_{i_r}], [x_{j_1}, x_{j_2}, \ldots, x_{j_{s-1}}]], x_{j_s}] - \\ - [[x_{i_1}, x_{i_2}, \ldots, x_{i_r}, x_{j_s}], [x_{j_1}, x_{j_2}, \ldots, x_{j_{s-1}}]].\end{aligned}$$

By the induction hypothesis the second summand on the right-hand side is equal to a linear combination of simple commutators in the generators. In the first summand on the right-hand side the subcommutator

$$[[x_{i_1}, x_{i_2}, \ldots, x_{i_r}], [x_{j_1}, x_{j_2}, \ldots, x_{j_{s-1}}]]$$

is a commutator of smaller weight and so, by induction on the weight, it is a linear combination of simple commutators in the generators. Hence, the entire first summand is also a linear combination of simple commutators in the generators.

Since application of the identities of the free Lie ring to a commutator in the generators transforms it into a linear combination of commutators in those same generators with the same multiplicities of occurrence, the assertion proved above yields the following useful technical lemma: in an arbitrary Lie ring any commutator in elements $y_1, y_2, \ldots, y_n$ is a linear combination of simple commutators in these elements, each one of which has the same weight in each of $y_1, y_2, \ldots, y_n$. Though this fact was established for commutators in the free generators, it is clear that any equation which holds in the free Lie ring also holds in any other Lie ring.

The Jacobi identity allows to "extract" any element, occurring in a simple commutator, to the start. More precisely the following lemma holds: if x_0 occurs in a commutator, then this commutator is a linear combination of simple commutators in the same elements, each one of which starts with x_0. The proof is by repeated application of the Jacobi identity to the simple commutators given by the previous lemma:

$$\begin{aligned}[[a_1, \ldots a_s], x_0, \ldots] = -[x_0, [a_1, \ldots, a_s], \ldots] = \\ = -[x_0, [a_1, \ldots, a_{s-1}], a_s, \ldots] + [x_0, a_s, [a_1, \ldots, a_{s-1}], \ldots] = \ldots\end{aligned}$$

and so on.

If l is an element of the free Lie ring F, then the equation $l = 0$ may be regarded as an identity in the variables occurring in l – that is, the free generators of F. The

class of all Lie rings which satisfy the given family of identities V constitutes a variety of Lie rings. The free ring of this variety is the factor-ring of the free Lie ring F by the verbal ideal $V(F)$ generated by the values of all words from V at arbitrary elements of F. (This is an analogue of a verbal subgroup. It is sometimes called a T-ideal.)

A Lie ring L is said to satisfy the n-th Engel condition (for short, L is an n-Engel Lie ring) if

$$[x, \underbrace{y, y, \ldots, y}_{n}] = 0$$

for all $x, y \in L$. A number of fundamental results on Engel Lie rings have been proved, and we state them here to facilitate reference.

1.3.1 Theorem (Kostrikin [76]). *If a d-generator Lie algebra over a field of characteristic p satisfies the n-th Engel condition where $n < p$ (or n is arbitrary in the case of $p = 0$), then it is nilpotent of (d, n)-bounded nilpotency class.*

Zel'manov [157] proved that in the case of characteristic zero there is even a global bound for the nilpotency class, independent on the number of generators. But in the case of positive characteristic Razmyslov [119] showed that the bound in Kostrikin's Theorem cannot be independent of the number of generators since there exist non-soluble, locally nilpotent, $(p-2)$-Engel Lie algebras of characteristic $p \geq 5$ (and non-soluble locally nilpotent groups of exponent $p \geq 5$).

Kostrikin's Theorem 1.3.1 gives a positive solution to the Restricted Burnside Problem for groups of prime exponent p because the associated Lie rings of such groups are $(p-1)$-Engel by a theorem of Magnus [100] and Sanov [126]. Recently Zel'manov [159, 160] also obtained a positive solution to the Restricted Burnside Problem for groups of prime-power exponent p^k. This follows from his theorem on the local nilpotency of n-Engel Lie algebras of characteristic p for any n and p and from his reduction theorem [158].

The reader, interested in the proofs of these theorems of Kostrikin, Razmyslov and Zel'manov, is referred to the books [78, 121, 145] and to the original papers [76, 119, 157-160]. We shall prove in §3.4 only Higgins' Theorem [39] on the nilpotency of soluble Engel Lie rings and, in §3.3, the Magnus-Sanov Theorem mentioned above.

In many cases there are theorems for Lie rings which are analogous to corresponding theorems about groups. Thus, at the beginning of Chapter 3, we state a few analogues of theorems on groups whose proofs (which we do not give) are mostly based on commutator calculations (for Lie rings such calculations are even easier than for groups). But there may be substantial differences and sometimes Lie

rings in certain sense are worse than groups. For example, the structural constants

$$[e_1, e_2] = e_3, \quad [e_2, e_3] = e_1, \quad [e_3, e_1] = e_2$$

define a 3-dimensional simple Lie algebra over any field including the finite field $GF(p)$ of prime order p – in which case it is a simple Lie algebra consisting of p^3 elements. However, a group of order p^3 is necessarily nilpotent.

§ 1.4 Mappings, homomorphisms, automorphisms

As a rule we use power notation for the action of mappings: a^φ denotes the image of a under the action of φ. But sometimes other notation is used: $a\varphi$ or $\varphi(a)$. The same comment applies to images of subsets.

The automorphisms of a group G comprise a group $\operatorname{Aut} G$, which will be always regarded as a subgroup of the natural semidirect product $G \leftthreetimes \operatorname{Aut} G$. Thus, the image a^φ of an element $a \in G$ under the action of an automorphism $\varphi \in \operatorname{Aut} G$ may be regarded as conjugate of a under φ. We may also write

$$C_G(\varphi) = \{g \in G \mid g^\varphi = g\},$$
$$[G, \varphi] = \langle [g, \varphi] \mid g \in G \rangle = \langle g^{-1} g^\varphi \mid g \in G \rangle$$

and so on.

A section M/N of a group G is said to be φ-invariant, where $\varphi \in \operatorname{Aut} G$, if $M^\varphi = M$ and $N^\varphi = N$.

If M/N is φ-invariant, then φ induces the automorphism $\bar{\varphi}$ of M/N by the rule $(mN)^{\bar{\varphi}} = m^\varphi N$. We shall usually denote the induced automorphism by the same letter, that is, we shall write φ instead of $\bar{\varphi}$, and $C_{M/N}(\varphi)$ instead of $C_{M/N}(\bar{\varphi})$, etc.

An automorphism φ of a group G (or a Lie ring L) is called regular if $C_G(\varphi) = 1$ (or $C_L(\varphi) = 0$).

§ 1.5 Group actions on sets

A group G is said to act on a set Ω, if for every $g \in G$ there is a one-to-one mapping of the set Ω onto itself, usually also denoted by g: $\omega \to \omega g$ (or $\omega \to \omega^g$), such that

$$(\omega g_1) g_2 = \omega (g_1 g_2)$$

for all $\omega \in \Omega$ and $g_1, g_2 \in G$. In particular, $\omega 1 = \omega$ for all $\omega \in \Omega$ and the mapping $\omega \to \omega g^{-1}$ is inverse to the mapping $\omega \to \omega g$. In other words an action of a group G on a set Ω is a homomorphism of G into the group of all one-to-one mappings of the set Ω onto itself. The action is said to be faithful if the kernel of this homomorphism is trivial. The elements of Ω are often called points. For $\omega_0 \in \Omega$ the subset

$$\omega_0 G = \{\omega_0 g |\ g \in G\}$$

is called an orbit under the action of G. The subset

$$G_{\omega_0} = \{g \in G |\ \omega_0 g = \omega_0\}$$

is a subgroup called the stabilizer of the point ω_0. There is a one-to-one correspondence between the set of (right) cosets of the point stabilizer G_{ω_0} in G and the orbit $\omega_0 G$:

$$G_{\omega_0} g \leftrightarrow \omega_0 g.$$

If the set Ω is finite, then $|G : G_{\omega_0}| = |\omega_0 G|$ and, in particular, $|G|$ is a multiple of $|\omega_0 G|$ by Lagrange's Theorem.

For example, every group faithfully acts on itself by right multiplication: $a \to ag$ for $a, g \in G$, that is, the image of an element a under the action of g equals ag. A group G also acts on any of its normal subgroups $N \trianglelefteq G$ by conjugation: $n \to n^g$, where $n \in N$, $g \in G$. The kernel of this action is the centralizer $C_G(N)$, the orbits are the conjugacy classes, the point stabilizers are the centralizers of the elements $C_G(g)$. In an analogous way G acts by conjugation on any of its normal sections M/N, the kernel of this action is the centralizer of the section M/N – that is the largest subgroup H satisfying the property $[H, M] \leq N$.

As an illustrative application of the notion of group action, let us prove the so-called Poincaré's Theorem: if H is a subgroup of finite index n in an arbitrary group G then H contains a normal subgroup of G whose index in G is also finite and does not exceed $n!$. We consider the action of G on the set of right cosets of H, defined by

$$(Hx)g = Hxg.$$

The kernel of this action is the desired normal subgroup; its index is bounded by $n!$ since the corresponding factor-group embeds in the symmetric group of all permutations of the set of n right cosets of H and this group has order $n!$.

Any group of automorphisms $G \leq \mathrm{Aut}\, H$ of a group (or other algebraic system – Lie ring, vector space, etc.) H acts on the set H in a natural way:

$$hg = h^g, \quad h \in H,\ g \in G,$$

where the left-hand side defines the action and the right-hand side is the image of the element h under the automorphism g.

§ 1.6 Fixed points of automorphisms

Here we prove a few well-known results on fixed points of automorphisms which will be used subsequently.

1.6.1 Theorem. *Let G be a finite group, φ an automorphism of G, and N a normal φ-invariant subgroup of G. Then*

$$|C_{G/N}(\varphi)| \leq |C_G(\varphi)|.$$

Proof. Note that, for any group H and automorphism $\psi \in \operatorname{Aut} H$, the number of elements of the form $x^{-1}x^{\psi}$, $x \in H$, is equal to $|H : C_H(\psi)|$, since $x^{-1}x^{\psi}$ depends only on the coset of $C_H(\psi)$ to which x belongs. More precisely, the mapping $xC_H(\psi) \to x^{-1}x^{\psi}$ is a one-to-one correspondence between the set $\{x^{-1}x^{\psi} \mid x \in H\}$ and the set of right cosets of $C_H(\psi)$: it is well defined and injective since

$$x^{-1}x^{\psi} = y^{-1}y^{\psi} \Leftrightarrow yx^{-1} = y^{\psi}(x^{\psi})^{-1} = (yx^{-1})^{\psi} \Leftrightarrow yx^{-1} \in C_H(\psi).$$

Now elements of the form $\bar{g}^{-1}\bar{g}^{\varphi}$ for $\bar{g} = gN \in G/N$ are images of the elements $g^{-1}g^{\varphi}$ of G in the factor-group G/N. Every coset $\bar{g}^{-1}\bar{g}^{\varphi} = g^{-1}g^{\varphi}N$ of N contains at most $|N|$ elements of the form $g^{-1}g^{\varphi}$ and each element of the form $g^{-1}g^{\varphi}$ lies in such a coset $g^{-1}g^{\varphi}N = \bar{g}^{-1}\bar{g}^{\varphi}$. Therefore

$$|N| \cdot |\{\bar{g}^{-1}\bar{g}^{\varphi} \mid \bar{g} \in G/N\}| \geq |\{g^{-1}g^{\varphi} \mid g \in G\}|.$$

Hence,

$$\frac{|G|}{|C_G(\varphi)|} \leq |N|\frac{|G/N|}{|C_{G/N}(\varphi)|} = |N|\frac{|G|}{|N||C_{G/N}(\varphi)|} = \frac{|G|}{|C_{G/N}(\varphi)|},$$

and this implies $|C_{G/N}(\varphi)| \leq |C_G(\varphi)|$ as required.

The theorem is proved.

This result may be strengthened if the order of the automorphism is coprime to the order of the normal subgroup.

1.6.2 Theorem. *Let G be a finite group, φ an automorphism of G, and N a normal φ-invariant subgroup of G whose order is coprime to the order of φ, that*

is $(|N|, |\varphi|) = 1$. *Then*

$$C_{G/N}(\varphi) = C_G(\varphi)N/N.$$

Proof. We shall prove equivalently that each φ-invariant coset gN of N contains an element of $C_G(\varphi)$.

We proceed by induction on $|\varphi|$. Suppose first of all that $|\varphi| = p$ is a prime number. The sizes of the φ-orbits, which constitute a partition of the φ-invariant coset gN, divide $|\varphi| = p$, and hence are equal either to p or to 1. If all orbits were of size p, then p would divide $|gN| = |N|$, and this would contradict the condition $(|N|, |\varphi|) = 1$. Therefore there must be an orbit of size 1, consisting of elements of $C_G(\varphi)$.

Now let the order $|\varphi| = mn$ be a composite number with $m > 1$ and $n > 1$. By the induction hypothesis we have

$$C_{G/N}(\varphi^n) = C_G(\varphi^n)N/N \cong C_G(\varphi^n)/C_G(\varphi^n) \cap N.$$

The centralizer $C_{G/N}(\varphi)$ is contained in $C_{G/N}(\varphi^n)$ and hence embeds isomorphically into $C_G(\varphi^n)/C_G(\varphi^n) \cap N$. Under this natural embedding the image also centralizes φ: indeed, each φ-invariant coset gN is φ^n-invariant and there exists $g_0 \in C_G(\varphi^n) \cap gN$ so that $g_0(C_G(\varphi^n) \cap gN)$ is φ-invariant since $C_G(\varphi^n)$ is φ-invariant. By the induction hypothesis applied to the automorphism φ acting on $C_G(\varphi^n)$ as an automorphism of order n, the φ-invariant coset $g_0(C_G(\varphi^n) \cap gN)$ contains an element of $C_G(\varphi)$. This element is what is required, since $g_0(C_G(\varphi^n) \cap N) \subseteq gN$.

The theorem is proved.

1.6.3 Corollary. *Let G be a finite group, φ – an automorphism of G such that $(|G|, |\varphi|) = 1$. If φ centralizes all factors of some subnormal series of G*

$$1 = K_0 \trianglelefteq K_1 \trianglelefteq K_2 \trianglelefteq \ldots \trianglelefteq K_s = G,$$

then $\varphi = 1$.

Proof. We shall prove by induction on i that $K_i \leq C_G(\varphi)$ for all i. By hypothesis $K_1 \leq C_G(\varphi)$. By Theorem 1.6.2

$$C_{K_{i+1}}(\varphi)K_i/K_i = K_{i+1}/K_i,$$

whence $C_{K_{i+1}}(\varphi)K_i = K_{i+1}$, and, since $K_i \leq C_G(\varphi)$ by the induction hypothesis, we have $K_{i+1} \leq C_G(\varphi)$.

1.6.4 Corollary. *Let G be a finite group and φ be an automorphism of G such that $(|G|, |\varphi|) = 1$. Then*

a) $G = C_G(\varphi)[G, \varphi]$;

b) $[[G, \varphi], \varphi] = [G, \varphi]$.

Proof. a) Since $g^\varphi = g[g, \varphi]$ for any $g \in G$, φ acts trivially on $G/[G, \varphi]$. Hence $G = C_G(\varphi)[G, \varphi]$ by Theorem 1.6.2.

b) We use a). Applying the formula $[ab, c] = [a, c]^b \cdot [b, c]$ we obtain

$$[G, \varphi] = [C_G(\varphi)[G, \varphi], \varphi] \leq [C_G(\varphi), \varphi]^{[G,\varphi]} \cdot [[G, \varphi], \varphi] = [[G, \varphi], \varphi].$$

1.6.5 Corollary. *If the group G of Corollary 1.6.4 is abelian, then (in additive notation)*

$$G = C_G(\varphi) \oplus [G, \varphi].$$

Proof. By Maschke's Theorem the $\mathbb{Z}\langle\varphi\rangle$-submodule $C_G(\varphi)$ has a φ-invariant direct complement U: $G = C_G(\varphi) \oplus U$. Since $C_U(\varphi) = 0$, we have $[U, \varphi] = U$ by Corollary 1.6.4. We now get

$$[G, \varphi] = [C_G(\varphi) \oplus U, \varphi] = [U, \varphi] = U,$$

and $G = C_G(\varphi) \oplus U = C_G(\varphi) \oplus [G, \varphi]$.

Corollary 1.6.3 may be strengthened by dropping the subnormality condition on the series.

1.6.6 Theorem. *Let G be a finite group and φ be an automorphism of G such that $(|G|, |\varphi|) = 1$. If for some series*

$$1 = K_0 \leq K_1 \leq K_2 \leq \ldots \leq K_s = G$$

all cosets kK_i are φ-invariant for all $k \in K_{i+1}$ and for all $i = 0, 1, \ldots, s - 1$, then $\varphi = 1$.

Proof. It is clearly sufficient to show that $gH \subseteq C_G(\varphi)$ for any φ-invariant coset gH where H is a subgroup contained in $C_G(\varphi)$. Then an obvious induction on i will show that $K_i \subseteq C_G(\varphi)$ for all i. We have $g^\varphi = gh$ for some $h \in H$ and, further, $g^{\varphi^k} = gh^k$ for all $k \in \mathbb{N}$, since $h \in H \leq C_G(\varphi)$. Hence, if $|\varphi| = n$, then $g = g^{\varphi^n} = gh^n$, whence $h^n = 1$, and therefore also $h = 1$, since $(|\varphi|, |h|) = 1$ by hypothesis.

The theorem is proved.

§ 1.7 The Jordan normal form of a linear transformation of finite order

This section also contains a few well-known results.

1.7.1 Theorem. *Let φ be an element of finite order n in the group of nondegenerate linear transformations of a vector space V over a field k. If the characteristic of k is coprime to n or is 0, then the Jordan normal form of φ over a suitable field extension of k is diagonal, and the eigenvalues of φ are n-th roots of unity.*

Proof. Straightforward calculation of the n-th power of a Jordan box of size greater than 1×1 shows that the equation

$$\begin{pmatrix} \alpha & 1 & \dots \\ \dots & & \end{pmatrix}^n = \begin{pmatrix} \alpha^n & n\alpha^{n-1} & \dots \\ \dots & & \end{pmatrix} = E$$

is possible only if $\alpha^n = 1$ and $n\alpha^{n-1} = 0$. Since $n \neq 0$ by the hypothesis on the characteristic, the second equation implies $\alpha = 0$, which contradicts the first one. Hence all Jordan boxes have size 1×1 and it is then clear that the eigenvalues of φ are n-th roots of unity.

The theorem is proved.

1.7.2 Theorem. *Let p be a prime number and let φ be an element of order p^k in the group of nondegenerate linear transformations of a vector space V over a field k of characteristic p. Then all eigenvalues of φ are equal to 1 and V has a k-basis with respect to which the matrix of φ is in Jordan normal form, all of its boxes are of size at most $p^k \times p^k$, and there exists at least one box of size greater than $p^{k-1} \times p^{k-1}$.*

Proof. Over a field of characteristic p we have

$$0 = 1 - \varphi^{p^k} = (1 - \varphi)^{p^k},$$

which means that φ is a root of the polynomial $(1 - \lambda)^{p^k}$. Thus the minimal polynomial of φ divides $(1 - \lambda)^{p^k}$ and hence all of the eigenvalues of φ are roots of the polynomial $(1 - \lambda)^{p^k}$. Therefore, all of them are equal to 1. In particular they are contained in k and it follows from the proof of the Jordan Normal Form Theorem that in this case V has a k-basis with respect to which the matrix of φ is in Jordan normal form.

Thus the Jordan normal form of φ consists of boxes of the form

$$\begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & 1 \end{pmatrix}.$$

Let us compute the p^s-th power of such a box:

$$\begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & 1 \end{pmatrix}^{p^s} = \begin{pmatrix} 1 & C_{p^s}^1 & C_{p^s}^2 & \dots & \\ & 1 & C_{p^s}^1 & \ddots & C_{p^s}^2 \\ & & \cdot & \ddots & C_{p^s}^1 \\ & & & \cdot & 1 \end{pmatrix}.$$

Since $\varphi^{p^k} = 1$, we have $C_{p^k}^1 = C_{p^k}^2 = \ldots = 0$. It is easy to see that $C_{p^s}^i$ is divisible by p, if $i < p^k$, and $C_{p^k}^{p^k} = 1$. Hence the size of each box is at most $p^k \times p^k$. The same calculations show that if all boxes have size at most $p^{k-1} \times p^{k-1}$, then $\varphi^{p^{k-1}} = 1$, a contradiction since $|\varphi| = p^k$. Therefore there exists at least one box of size greater than $p^{k-1} \times p^{k-1}$.

The theorem is proved.

1.7.3 Corollary. *Let p be a prime number. If φ is an automorphism of order p^k of an abelian p-group V then $C_V(\varphi) \neq 1$.*

Proof. The subgroup $\Omega_1(V)$ may be regarded as a vector space over the field $GF(p)$ and φ as a linear transformation of $\Omega_1(V)$ (see §1.1). By Theorem 1.7.2, 1 is an eigenvalue of φ. Non-trivial eigenvectors with eigenvalue 1 are clearly non-trivial elements of $C_{\Omega_1(V)}(\varphi) \leq C_V(\varphi)$.

1.7.4 Corollary. *Let p be a prime number and let φ be an automorphism of order p^k of an abelian p-group V, where $|C_V(\varphi)| = p^n$. Then the rank of V is at most $p^k n$.*

Proof. As above we think of the subgroup $\Omega_1(V)$ as a vector space over $GF(p)$ and φ as a linear transformation of $\Omega_1(V)$. By Theorem 1.7.2 all eigenvalues of φ are equal to 1 and the vector space $\Omega_1(V)$ has a $GF(p)$-basis with respect to which the matrix of φ is in Jordan normal form, its boxes being of sizes $\leq p^k \times p^k$. It is easy to see that the subset of this basis, corresponding to a given box,

contains a single eigenvector lying in $C_{\Omega_1(V)}(\varphi)$. It is also clear that eigenvectors, corresponding to different boxes, are linearly independent. Therefore the dimension $d = \dim C_{\Omega_1(V)}(\varphi)$ is equal to the number l of boxes. The rank of the group $\Omega_1(V)$, that is, $\dim \Omega_1(V)$, is equal to the sum of the dimensions of the boxes and hence is not greater than $p^k l$. We observe finally that

$$p^d = |C_{\Omega_1(V)}(\varphi)| \leq |C_V(\varphi)| = p^n,$$

whence $l = d \leq n$.

The corollary is proved.

It is clear, that we have, in fact, also proved the analogous result for dimensions of vector spaces.

1.7.5 Corollary. *Let p be a prime number and let φ be a linear transformation of order p^k of a vector space V over a field of characteristic p. If the dimension of the subspace of fixed elements $C_V(\varphi)$ is finite and equals n, then the dimension of V is at most $p^k n$.*

§ 1.8 Varieties and free groups

A (group) word $v = v(\bar{x})$ in the variables $\bar{x} = (x_1, x_2, \ldots, x_n)$ (or a word in the (group) variables $x_1, x_2, \ldots, x_n$) is an element of the free group with free generators $x_1, x_2, \ldots, x_n$. The value $v(\bar{g})$ of the word $v(\bar{x})$ at the elements $\bar{g} = (g_1, g_2, \ldots, g_n)$ of some group G is the image of the element $v(\bar{x})$ under the homomorphism which extends the mapping $x_i \to g_i$, $i = 1, 2, \ldots, n$.

Let $v(\bar{x}) = v(x_1, x_2, \ldots, x_n)$ be some word. A group G is said to satisfy the identity $v(\bar{x}) = 1$ if the values of the word $v(\bar{x})$ at any elements of the group G are trivial, or, in other words, if the verbal subgroup $v(G)$ is trivial.

Let $V = \{v_\alpha(x)\}$ be a set of words. The class of all groups satisfying all of the identities $v_\alpha \in V$ is called the variety of groups $\mathfrak{M}_V$, defined by the set of identities V. The factor-group of the free group $F = \langle x_1, x_2, \ldots \rangle$ by its verbal subgroup

$$V(F) = \langle v_\alpha(\bar{g}) |\ \bar{g} = (g_1, g_2, \ldots),\ g_i \in F,\ v_\alpha \in V \rangle$$

is called the free group of the variety $\mathfrak{M}_V$. If h_i are elements in an arbitrary group G from $\mathfrak{M}_V$, then the mapping $x_i \to h_i$ induces a homomorphism of $F/V(F)$ into G.

It is not difficult to see that all groups in $\mathfrak{M}_V$ satisfy an identity if and only if the countably-generated free group of $\mathfrak{M}_V$ satisfies that identity.

Birkhoff's Theorem states that varieties are precisely those classes of groups which are closed under taking subgroups, homomorphic images and cartesian products.

For many varieties there is a standard notation. The variety of all abelian groups (defined by the identity $[x, y] = 1$) is denoted by $\mathfrak{A}$, the variety of all soluble groups of derived length k (defined by the identity $\delta_k = 1$, see the Preface) is denoted by $\mathfrak{A}^k$, the variety of all nilpotent groups of nilpotency class c (defined by the identity $[x_1, x_2, \ldots, x_{c+1}] = 1$) is denoted by $\mathfrak{N}_c$, and the variety of all groups of given finite exponent n (defined by the identity $x^n = 1$) is denoted by $\mathfrak{B}_n$.

Many of group-theoretic results may be formulated in terms of varieties. For example, the positive solution of the Restricted Burnside Problem for groups of prime-power exponent p^k means that all locally nilpotent groups of exponent p^k constitute a variety (whose free m-generator groups are denoted by $\bar{B}(m, p^k)$).

Sometimes the language of the theory of varieties is also useful for proving theorems. For example, it may be more convenient to do a calculation first in the free group of a variety and then to apply it later to all groups in the variety. Another typical example is: suppose we succeed in proving that all groups satisfying a certain condition are nilpotent. If this class of groups constitutes a variety, then we automatically get the boundedness of the nilpotency class of these groups. Indeed, if there were groups G_i in this class of unboundedly large nilpotency classes $c_i \to \infty$, then their cartesian product $\times_i G_i$ would be non-nilpotent, although it belongs to the same variety by Birkhoff's Theorem.

The theory of varieties, naturally, has its own problems and specific ways of reasoning. We reproduce here the following simple result, which was independently and more or less simultaneously published in [73, 97, 146].

1.8.1 Theorem. *If every non-trivial group of some variety $\mathfrak{M}$ is distinct from its commutator subgroup, then the variety is soluble (that is $\mathfrak{M} \subseteq \mathfrak{A}^k$ for some k).*

Proof. By a standard argument, like that mentioned above for nilpotency, it is sufficient to show that all groups of the variety $\mathfrak{M}$ are soluble. Suppose, on the contrary, that $\mathfrak{M}$ contains non-soluble groups. It is then clear that the free group F of $\mathfrak{M}$ on a countable set of free generators $x_1, x_2, \ldots$ is also non-soluble.

Let us consider the sequence of isomorphic copies F_i of the group F

$$F_1 \xrightarrow{\tau} F_2 \xrightarrow{\tau} \ldots \xrightarrow{\tau} F_i \xrightarrow{\tau} F_{i+1} \xrightarrow{\tau} \ldots \tag{1.8.2}$$

together with homomorphisms τ, which are defined on the free generators by the rule

$$\tau\colon x_j \to [x_{2j-1}, x_{2j}], \quad j = 1, 2, \ldots$$

(here, on the left-hand side $x_j \in F_i$, and on the right-hand side $x_{2j-1}, x_{2j} \in F_{i+1}$).

It is not difficult to see that in the cartesian product

$$F_1 \times F_2 \times \ldots \times F_i \times \ldots,$$

all τ-threads, that is, elements of the form

$$(a, a\tau, a\tau^2, \ldots, a\tau^i, \ldots),$$

constitute a subgroup H, and the set of all elements with only finitely many non-trivial coordinates forms a normal subgroup N. The section HN/N is called the direct limit of the spectrum (1.8.2) and is denoted by $\varinjlim F_i$.

It is important to note that this group is non-trivial: for instance, the element $(x_1, x_1\tau, x_1\tau^2, \ldots) \in H$ does not belong to N, since $x_1\tau^i = \delta_i(x_1, x_2, \ldots, x_{2^i}) \neq 1$ for all i by virtue of the non-solubility of the group F.

By Birkhoff's Theorem the group $\varinjlim F_i$ also belongs to the variety $\mathfrak{M}$. We shall now show that this group coincides with its commutator subgroup. Indeed, the group $\varinjlim F_i$ is generated by the images of the threads

$$(x_j, x_j\tau, x_j\tau^2, \ldots, x_j\tau^i, \ldots), \quad j = 1, 2, \ldots,$$

where the x_j are the generators of F. But, modulo N, for each j the following congruence holds:

$$\begin{aligned}(x_j, x_j\tau, x_j\tau^2, \ldots, x_j\tau^i, \ldots) &\equiv (1, x_j\tau, x_j\tau^2, \ldots, x_j\tau^i, \ldots) = \\ &= (1, [x_{2j-1}, x_{2j}],\ [x_{2j-1}, x_{2j}]\tau, \ldots) = \\ &= \left[(1, x_{2j-1}, x_{2j-1}\tau, \ldots), (1, x_{2j}, x_{2j}\tau, \ldots)\right].\end{aligned}$$

This means that all generators of the group $\varinjlim F_i$ belong to its commutator subgroup, which therefore coincides with the group itself.

Thus the non-trivial group $\varinjlim F_i \in \mathfrak{M}$ coincides with its commutator subgroup, contradicting the hypothesis of the theorem.

§ 1.9 Groups with operators

Let Ω be a group. A group G is said to admit Ω as a group of operators (G is an Ω-group), if Ω acts on G (not necessarily faithfully) as a group of automorphisms, or, in other words, if there is a homomorphism of Ω into $\operatorname{Aut} G$.

Ω-groups may be regarded as algebraic systems which, in addition to the group operations, have unary operations corresponding to the action of the elements of Ω. Ω-subgroups of an Ω-group are its Ω-invariant subgroups; the homomorphisms of an Ω-group commute with the action of the operators from Ω: if σ is a homomorphism of an Ω-group G, then $(g^{\omega})^{\sigma} = (g^{\sigma})^{\omega}$ for all $g \in G$, $\omega \in \Omega$.

The basic properties of the theory of varieties, common to arbitrary algebraic systems, hold also for Ω-groups: there exist free m-generated Ω-groups whose elements, called Ω-group words, or simply Ω-words, define as Ω-identities varieties of Ω-groups (varieties of groups with operators); the free groups of these varieties are factor-groups of the free Ω-groups over verbal Ω-subgroups, etc.

As an abstract ("conventional") group, a free Ω-group with free generators $x_1, x_2, \ldots, x_n$ is a free group on free generators $x_i{}^{\omega}$, $i = 1, 2, \ldots, n$, $\omega \in \Omega$, on which Ω acts as follows:

$$(x_i^{\omega})^{\varphi} = x_i^{(\omega\varphi)}, \quad \omega, \varphi \in \Omega, \ i = 1, 2, \ldots, n.$$

Therefore it is clear that every Ω-word has the form

$$(x_{i_1}^{\omega_1})^{\varepsilon_1} \cdot (x_{i_2}^{\omega_2})^{\varepsilon_2} \cdot \ldots \cdot (x_{i_s}^{\omega_s})^{\varepsilon_s}, \quad \varepsilon_i = \pm 1, \quad \omega_i \in \Omega.$$

Example. Let p be a prime number. All groups G which admit an automorphism φ such that

$$\varphi^p = 1 \quad \text{and} \quad x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1 \text{ for all } x \in G,$$

constitute a variety of $\langle\varphi\rangle$-groups $\mathfrak{M}_p$. It is interesting to note that if $\varphi = 1$ then the group G is of exponent p, and if G is a finite p'-group, then the automorphism φ is regular. This variety will be studied in Chapter 7 where we shall prove an analogue of the positive solution of the Restricted Burnside Problem: locally nilpotent groups from $\mathfrak{M}_p$ form a subvariety of $\langle\varphi\rangle$-groups.

§ 1.10 Higman's Lemma

Let F be a free group (or a free Ω-group) with free generators $x_1, x_2, \ldots, x_n$. For every $i = 1, 2, \ldots, n$ consider the homomorphism ϑ_i of F into itself which extends the mapping

$$x_i \to 1, \quad x_k \to x_k \text{ for } k \neq i.$$

It is clear that $\operatorname{Ker}\vartheta_i = \langle x_i^F \rangle$ (or $\operatorname{Ker}\vartheta_i = \langle x_i^{F\Omega} \rangle$) is the normal closure of x_i in F. For any subset $J \subseteq \{1, 2, \ldots, n\}$ we put

$$D_J = \bigcap_{j \in J} \operatorname{Ker}\vartheta_i.$$

Next we shall consider only commutators in the generators x_i and their inverses, we shall say that such a commutator depends on x_j if it involves x_j or x_j^{-1}.

1.10.1 Lemma. *The subgroup D_J is generated by commutators in the generators x_i and their inverses each of which depends on all of the elements x_j for $j \in J$.*

Proof. We proceed by induction on the cardinality of J. First let $J = \{j\}$, then $D_J = \langle x_j^F \rangle$. Since $x_j^a = x_j[x_j, a]$ for all $a \in F$, then $D_J = \langle x_j, [x_j, a] \mid a \in F \rangle$. Each element of F is a product of the generators x_i and their inverses and so, by repeated application of the standard commutator identities

$$[u, vw] = [u, w][u, v][u, v, w]; \quad [uv, w] = [u, w][u, w, v][v, w],$$

we can express the $[x_j, a]$ as a product of commutators in x_i and x_i^{-1} each of which depends on the element x_j.

Now let $J' = J \cup \{k\}$, that is $D_{J'} = D_J \cap \langle x_k^F \rangle$ and let b be an arbitrary element of $D_J \cap \langle x_k^F \rangle$. By the induction hypothesis the element b is the product

$$b = c_1 c_2 \ldots c_s \tag{1.10.2}$$

of commutators c_i each of which depends on x_j for every $j \in J$. If among the c_i there are some which do not depend on x_k, we subsequently transpose them to the left, preserving the order of their occurrence, by the formula

$$\ldots cc' \ldots = \ldots c'c[c, c'] \ldots,$$

where c is a commutator depending on x_k, c' is a commutator which does not depend on x_k, and the dots denote that part of the product (1.10.2) which is not changed at this step. It is clear that these transpositions preserve (1.10.2) and that all additional factors $[c, c']$ arising are both dependent on all the x_i, $i \in J$ and on x_k. As a result we get

$$b = d_1 d_2 \ldots d_t e_1 e_2 \ldots e_r,$$

where each of the d_i does not depend on x_k, and each of the e_i depends on x_j for all $j \in J'$. Applying ϑ_k to this equation we obtain

$$1 = d_1 d_2 \ldots d_t,$$

since $\vartheta_k(b) = 1$ because of the way we chose b, and because, clearly, $\vartheta_k(e_i) = 1$ for all e_i, depending on x_k. Hence,

$$b = e_1 e_2 \dots e_r,$$

and this proves the assertion of the lemma for J'.

The lemma is proved.

1.10.3 Higman's Lemma ([41]). *An arbitrary element w of the free n-generator group F with free generators $x_1, x_2, \dots, x_n$ may be represented as a product*

$$w = u v_1 v_2 \dots v_{2^n - 1},$$

where $u \in D_{\{1,2,\dots,n\}}$, and each of the elements v_i has the form

$$v_i = w^{(-1)^r} \vartheta_{j_1} \vartheta_{j_2} \dots \vartheta_{j_r},$$

where $1 \leq j_1 < j_2 < \dots < j_r \leq n$.

(The order in which the elements $v_i = w^{(-1)^r} \vartheta_{j_1} \vartheta_{j_2} \dots \vartheta_{j_r}$ occur here is immaterial.)

Proof. It may be easily checked that the following equalities hold for the homomorphisms ϑ_i:

$$\vartheta_k \vartheta_k = \vartheta_k \quad \text{and} \quad \vartheta_i \vartheta_j = \vartheta_j \vartheta_i \text{ for } i \neq j.$$

Let us denote by $1 - \vartheta_i$ the mapping $a \to a(a\vartheta_i)^{-1}$. We have

$$a(1 - \vartheta_i)\vartheta_i = (a(a\vartheta_i)^{-1})\vartheta_i = a\vartheta_i(a\vartheta_i^2)^{-1} = a\vartheta_i(a\vartheta_i)^{-1} = 1,$$

and this means that

$$a(1 - \vartheta_i) \in \operatorname{Ker} \vartheta_i = D_{\{i\}}. \tag{1.10.4}$$

It is also easy to see that

$$b(1 - \vartheta_i) = b, \text{ for any } b \in D_{\{i\}}. \tag{1.10.5}$$

Therefore

$$u = w(1 - \vartheta_1)(1 - \vartheta_2) \dots (1 - \vartheta_n)$$

belongs to the subgroup $D_{\{1,2,\dots,n\}}$. By expanding the brackets and using the definition of the mappings $1 - \vartheta_i$ we express the element w to get the desired result.

The lemma is proved.

This result is often used in combinatorial arguments in group theory. For our purposes it will be convenient to record two of its corollaries. We shall now prove the first of them directly.

1.10.6 Corollary. *Let M and N be verbal subgroups of the free n-generator group F on free generators $x_1, x_2, \ldots, x_n$ and let J be an arbitrary subset of $\{1, 2, \ldots, n\}$. Then, for any element $v \in D_J \cap MN$, there is an element $w \in D_J \cap M$ such that $v \equiv w \pmod N$.*

Proof. For any $a \in M$, $b \in N$ we have

$$ab(1 - \vartheta_k) = ab((ab)\vartheta_k)^{-1} =$$
$$= ab(b\vartheta_k)^{-1}(a\vartheta_k)^{-1} \equiv a(1 - \vartheta_k) \pmod N, \tag{1.10.7}$$

since the verbal subgroup N is ϑ_k-invariant and normal.

By hypothesis $v = ab$, where $a \in M$, $b \in N$. We apply to this equation the mappings $1 - \vartheta_{i_k}$, where $J = \{i_1, i_2, \ldots, i_s\}$. In view of (1.10.5) this does not change the left-hand side, since $v \in D_J$ by hypothesis. Using also (1.10.7), we get

$$v = v(1 - \vartheta_{i_1})(1 - \vartheta_{i_2}) \ldots (1 - \vartheta_{i_s}) =$$
$$= ab(1 - \vartheta_{i_1})(1 - \vartheta_{i_2}) \ldots (1 - \vartheta_{i_s}) \equiv$$
$$\equiv a(1 - \vartheta_{i_1})(1 - \vartheta_{i_2}) \ldots (1 - \vartheta_{i_s}) \pmod N.$$

The element $w = a(1 - \vartheta_{i_1})(1 - \vartheta_{i_2}) \ldots (1 - \vartheta_{i_s})$ satisfies the conclusion of the corollary: it belongs to M, since M is verbal, and to D_J by (1.10.4).

The corollary is proved.

Analogous assertions hold in free nilpotent groups and their proofs are no different from the proofs in this section. We shall recall them in the next chapter (in §2.7) after giving the relevant definitions.

We now derive another corollary of Higman's Lemma the proof of which involves a standard argument using extra generators.

1.10.8 Corollary. *Suppose that p is a natural number and that $f, x_1, x_2, \ldots, x_d$ are arbitrary elements of a group G. Then*

$$(f \cdot x_1^{p-1} x_2^{p-1} \cdot \ldots \cdot x_d^{p-1})^p = h \cdot \prod_{0 \le a_i \le p-1} (x_1^{a_1} x_2^{a_2} \cdot \ldots \cdot x_d^{a_d} \cdot f^{\varepsilon})^{\pm p}, \tag{1.10.9}$$

where $\varepsilon = 0$ or 1, the product involves the factor $f^{\pm p}$ and h is a product of powers of commutators each of which contains at least $p - 1$ occurrences of every element $x_1, x_2, \ldots, x_d$ and at least one occurrence of f.

Proof. Here we have to control the number of occurrences of elements whereas Higman's Lemma indicates only the existence of occurrences. However we can nevertheless use Higman's Lemma here by considering a multiple set of generators. Namely, let

$$f, x_{11}, x_{12}, \ldots, x_{1p-1}, x_{21}, x_{22}, \ldots, x_{2p-1}, \ldots, x_{d1}, x_{d2}, \ldots, x_{dp-1}$$

be free generators of a free group F. We apply Higman's Lemma 1.10.3 to the word

$$w = (f x_{11} x_{12} \ldots x_{1p-1} x_{21} x_{22} \ldots x_{2p-1} \ldots x_{d1} x_{d2} \ldots x_{dp-1})^p$$

to get

$$w = u v_1 v_2 \ldots v_m, \tag{1.10.10}$$

where $u \in \operatorname{Ker} \vartheta_0 \cap \bigcap_{i,j} \operatorname{Ker} \vartheta_{ij}$ and each of the elements v_i has the form

$$v_i = w^{(-1)^{r-1}} \vartheta_{j_1} \vartheta_{j_2} \ldots \vartheta_{j_r},$$
$$\vartheta_{ij} \in \{\vartheta_0, \vartheta_{11}, \ldots, \vartheta_{1p-1}, \vartheta_{21}, \ldots, \vartheta_{2p-1}, \ldots, \vartheta_{d1}, \ldots, \vartheta_{dp-1}\},$$

where the homomorphisms ϑ_0 and ϑ_{ij} are defined as follows:

$$\vartheta_0\colon\ f \to 1; \quad x_{ij} \to x_{ij} \text{ for all } i, j;$$
$$\vartheta_{st}\colon\ x_{st} \to 1; \quad x_{ij} \to x_{ij} \text{ for all } \{i, j\} \neq \{s, t\}; \quad f \to f.$$

Since $v_s^{\pm 1} = w \vartheta_{11} \ldots \vartheta_{1p-1} \vartheta_{21} \ldots \vartheta_{2p-1} \ldots \vartheta_{d1} \ldots \vartheta_{dp-1} = f^p$ for some s, the product clearly involves the factor $f^{\pm p}$.

We now apply to (1.10.10) the homomorphism extending the mapping

$$f \to f; \quad x_{ij} \to x_i \text{ for all } i, j.$$

It is clear that the image of (1.10.10) under this homomorphism is (1.10.9) which satisfies the conclusion of Lemma 1.10.8 (in particular, for each i, occurrences of elements $x_{i1}, x_{i2}, \ldots, x_{ip-1}$ in commutators the product of whose powers is u, guarantee that h (the image of u) contains at least $p - 1$ occurrences of x_i).

The corollary is proved.

Chapter 2

Nilpotent groups

At the beginning of this chapter we use commutator calculations to prove some standard facts connected with nilpotent groups. We then prove some sufficient conditions for soluble groups to be nilpotent.

Schur's Theorem on groups with finite central factor-groups is included for the sake of completeness. We are more interested in converses of this result for finite-by-nilpotent groups, since their proofs prepare the reader for the more complicated analogous arguments which are used in Chapters 4 and 5 for Lie rings and for nilpotent groups with almost regular automorphisms of prime order.

Then we use the language of tensor products of abelian groups to study linear properties of the lower central series of groups. As an application we give an account of the theory of (π-) isolators which is of particular interest for (π-) torsion-free groups.

We give definitions of basic commutators and the collecting process which are important for further applications. We prove the consistency of the collecting process and the fact that the factors of the lower central series of a group are generated by images of basic commutators in its generators. However, we do not prove (nor do we use) the fact that images of these basic commutators are free generators in this case.

The section on finite p-groups contains, as well as elementary properties, both P. Hall's Theorem on the quotients of the lower central series of a normal subgroup of a group and the proof of the so-called Zassenhaus' identity which holds in any group of prime exponent.

§ 2.1 Commutators and commutator subgroups

Definitions of commutators and of commutator subgroups were given in § 1.1. Their properties listed in this section are valid for any group.

2.1.1 Lemma. *The following hold for any elements $a, b, c \in G$ of an arbitrary group G:*

a) $ab = ba[a, b], \quad a^b = a[a, b]$;
b) $[a, b]^{-1} = [b, a]$;
c) $[ab, c] = [a, c]^b \cdot [b, c] = [a, c] \cdot [a, c, b] \cdot [b, c]$;
d) $[a, bc] = [a, c] \cdot [a, b]^c = [a, c] \cdot [a, b] \cdot [a, b, c]$;
e) (*Witt's identity*) $[a, b^{-1}, c]^b \cdot [b, c^{-1}, a]^c \cdot [c, a^{-1}, b]^a = 1$.

Proof is by the direct calculation. For example, we obtain the first part of c) as follows:

$$[a, c]^b \cdot [b, c] = b^{-1}a^{-1}c^{-1}acbb^{-1}c^{-1}bc = b^{-1}a^{-1}c^{-1}abc = [ab, c],$$

and the second part follows from a).

We prove e):

$$\begin{aligned}
&[a, b^{-1}, c]^b \cdot [b, c^{-1}, a]^c \cdot [c, a^{-1}, b]^a = \\
&= b^{-1}[a, b^{-1}]^{-1}c^{-1}[a, b^{-1}]cb\times \\
&\times c^{-1}[b, c^{-1}]^{-1}a^{-1}[b, c^{-1}]ac\times \\
&\times a^{-1}[c, a^{-1}]^{-1}b^{-1}[c, a^{-1}]ba = \\
&= b^{-1}[b^{-1}, a]c^{-1}[a, b^{-1}]cb\times \\
&\times c^{-1}[c^{-1}, b]a^{-1}[b, c^{-1}]ac\times \\
&\times a^{-1}[a^{-1}, c]b^{-1}[c, a^{-1}]ba = \\
&= \underrightarrow{b^{-1}}\,\underleftarrow{b}\,a^{-1}b^{-1}ac^{-1}a^{-1}\,\underrightarrow{b}\,\underrightarrow{a}\,\underrightarrow{b^{-1}}\,\underrightarrow{c}\,\underrightarrow{b}\times \\
&\times \underrightarrow{c^{-1}}\,\underleftarrow{c}\,\underleftarrow{b^{-1}}\underleftarrow{c^{-1}}\,\underleftarrow{b}\,\underleftarrow{a^{-1}}\underleftarrow{b^{-1}}\,\underrightarrow{c}\,\underrightarrow{b}\,\underrightarrow{c^{-1}}\,\underrightarrow{a}\,\underrightarrow{c}\times \\
&\times \underrightarrow{a^{-1}}\,\underleftarrow{a}\,\underleftarrow{c^{-1}}\underleftarrow{a^{-1}}\,\underleftarrow{c}\,\underleftarrow{b^{-1}}\underleftarrow{c^{-1}}aca^{-1}ba = \\
&= \underrightarrow{a^{-1}}\underrightarrow{b^{-1}}\,\underrightarrow{a}\,\underrightarrow{c^{-1}}\underrightarrow{a^{-1}}\,\underleftarrow{a}\,\underleftarrow{c}\,\underleftarrow{a^{-1}}\,\underleftarrow{b}\,\underleftarrow{a} = 1.
\end{aligned}$$

(The arrows denote cancellations.)

2.1.2 Three Subgroup Lemma. *Let N be a normal subgroup of an arbitrary group G. If for subgroups $A, B, C \leq G$ we have*

$$[[A, B], C] \leq N \quad \textit{and} \quad [[B, C], A] \leq N,$$

then we also have $[[C, A], B] \leq N$.

Proof. It is clearly sufficient to prove the lemma in the case when $N = 1$ and we shall assume that fact in what follows. By Witt's identity 2.1.1 e) we have for any

$a \in A$, $b \in B$, $c \in C$

$$[a, b^{-1}, c]^b \cdot [b, c^{-1}, a]^c \cdot [c, a^{-1}, b]^a = 1.$$

The first two factors on the left are trivial since they lie in $[[A, B], C]$ and $[[B, C], A]$, respectively. Hence $[c, a^{-1}, b]^a = 1$ and therefore

$$[c, a^{-1}, b] = ([c, a^{-1}, b]^a)^{a^{-1}} = 1^{a^{-1}} = 1.$$

As a runs through A so does a^{-1} so that we also get $[c, a, b] = 1$ for any $a \in A$, $b \in B$, $c \in C$. Hence, $b \in C_G([c, a])$ for any $b \in B$, and therefore $B \leq C_G([c, a])$, that is, $[c, a] \in C_G(B)$ for any $a \in A$, $c \in C$. But $[C, A]$ is generated by the elements $[c, a]$ and hence $[C, A] \leq C_G(B)$ so that $[[C, A], B] = 1$.

The lemma is proved.

2.1.3 Corollary. *The following hold for any group G for any natural numbers m and n:*

a) $[\gamma_m(G), \gamma_n(G)] \leq \gamma_{m+n}(G)$;

b) $[\gamma_m(G), \zeta_n(G)] \leq \zeta_{n-m}(G)$, *for $n \geq m$ (where $\zeta_0(G) = 1$).*

Proof. Denote for short $\gamma_i(G) = \gamma_i$, $\zeta_j(G) = \zeta_j$. We prove a) by induction on n. For $n = 1$ by definition $[\gamma_m, \gamma_1] = \gamma_{m+1}$ for all m. For $n > 1$ by the induction hypothesis

$$[\gamma_m, \gamma_{n-1}, \gamma_1] \leq [\gamma_{m+n-1}, \gamma_1] = \gamma_{m+n},$$

and also

$$[\gamma_1, \gamma_m, \gamma_{n-1}] = [\gamma_{m+1}, \gamma_{n-1}] \leq \gamma_{m+n}.$$

Therefore, by the Three Subgroup Lemma we deduce that

$$[\gamma_m, \gamma_n] = [\gamma_n, \gamma_m] = [\gamma_{n-1}, \gamma_1, \gamma_m] \leq \gamma_{m+n}.$$

We prove b) by induction on m. For $m = 1$ we have $[\gamma_1, \zeta_n] \leq \zeta_{n-1}$ for every n, since ζ_n/ζ_{n-1} is the centre of G/ζ_{n-1}. For $m > 1$ by the induction hypothesis

$$[\gamma_{m-1}, \zeta_n, \gamma_1] \leq [\zeta_{n-m+1}, \gamma_1] = [\gamma_1, \zeta_{n-m+1}] \leq \zeta_{n-m},$$

and also

$$[\zeta_n, \gamma_1, \gamma_{m-1}] \leq [\zeta_{n-1}, \gamma_{m-1}] = [\gamma_{m-1}, \zeta_{n-1}] \leq \zeta_{n-m}.$$

So, by the Three Subgroup Lemma we also get

$$[\gamma_m, \zeta_n] = [\gamma_1, \gamma_{m-1}, \zeta_n] \leq \zeta_{n-m}.$$

2.1.4 Corollary. *If A, B, C are normal subgroups of an arbitrary group G, then the following holds:*

$$[[A, B], C] \leq [[B, C], A] \cdot [[C, A], B] = [[A, C], B] \cdot [[A, [B, C]].$$

Proof. We only need apply the Three Subgroup Lemma, regarding the product $[[B,C],A] \cdot [[C,A],B]$ as a normal subgroup containing both $[[B, C], A]$ and $[[C, A], B]$.

2.1.5 Proposition. *Let G be an arbitrary group and k any natural number. Then*

a) *$\gamma_k(G)$ contains all commutators of weight $\geq k$ in the elements of G;*

b) *$\gamma_k(G)$ is generated by the simple commutators of weight k in the elements of G;*

c) *if $G = \langle M \rangle$, then the subgroup $\gamma_k(G)$ is generated by the simple commutators of weight $\geq k$ in the elements of M and their inverses.*

Proof. a) Induction on k. For $k = 1$ the assertion is trivial. Now let $k > 1$ and let c be a commutator of weight $r \geq k$. Then $c = [c_1, c_2]$, where c_1 and c_2 are commutators of weight r_1, r_2, respectively, with $r_1 + r_2 = r$. By the induction hypothesis $c_1 \in \gamma_{r_1}(G)$ and $c_2 \in \gamma_{r_2}(G)$. Therefore,

$$c = [c_1, c_2] \in [\gamma_{r_1}(G), \gamma_{r_2}(G)] \leq \gamma_r(G) \leq \gamma_k(G)$$

by 2.1.3 a).

b) Let us set

$$N_k = \langle [x_1, x_2, \ldots, x_k] \,|\, x_i \in G,\ i = 1, 2, \ldots, k \rangle.$$

Since $[x_1, x_2, \ldots, x_k]^y = [x_1^y, x_2^y, \ldots, x_k^y]$, we have $N_k \trianglelefteq G$. It is clear, via a), for example, that $N_k \leq \gamma_k(G)$. We prove the reverse inclusion by induction on k. For $k = 1$ it is trivial. For $k > 1$ the images of all commutators $[x_1, x_2, \ldots, x_{k-1}]$ of weight $k - 1$ lie in the centre of the factor-group G/N_k since $[[x_1, x_2, \ldots, x_{k-1}], x] \in N_k$ for all $x \in G$. Hence, $[N_{k-1}, G] \leq N_k$. By the induction hypothesis we now get

$$\gamma_k(G) = [\gamma_{k-1}(G), G] = [N_{k-1}, G] \leq N_k.$$

c) According to b) we have

$$\gamma_k(G) = \langle [x_1, x_2, \ldots, x_k] \,|\, x_i \in G,\ i = 1, 2, \ldots, k \rangle.$$

We may express each element x_i as a product of elements of M and their inverses. Repeated application of 2.1.1 c), d) yields the desired conclusion.

§ 2.2 Definitions and basic properties of nilpotent groups

The following theorem establishes the equivalence of different definitions of nilpotent groups.

2.2.1 Theorem. *For a group G the following conditions are equivalent:*

a) $\gamma_{c+1}(G) = 1$;

b) $\zeta_c(G) = G$;

c) *G has a central series of length c*

$$G = G_1 \geq G_2 \geq \ldots \geq G_c \geq G_{c+1} = 1,$$

that is, a series such that $[G_i, G] \leq G_{i+1}$ for all $i = 1, 2, \ldots, c$.

d) *G satisfies the identity*

$$[x_1, x_2, \ldots, x_{c+1}] = 1.$$

Proof. Let us assume c). Then induction on i shows that $\gamma_i(G) \leq G_i$: indeed $\gamma_1(G) = G_1 = G$, and if $\gamma_i(G) \leq G_i$, then

$$\gamma_{i+1}(G) = [\gamma_i(G), G] \leq [G_i, G] \leq G_{i+1}.$$

Therefore, $\gamma_{c+1}(G) \leq G_{c+1} = 1$. In an analogous way for all i we have $\zeta_i(G) \geq G_{c+1-i}$: indeed, $\zeta_1(G) \geq G_c$, since $[G_c, G] = 1$, and if $\zeta_i(G) \geq G_{c+1-i}$, then the inclusion $[G_{c-i}, G] \leq G_{c-i+1} \leq \zeta_i(G)$ implies that the image of G_{c-i} in $G/\zeta_i(G)$ is contained in its centre, which means that $\zeta_{i+1}(G) \geq G_{c-i}$. In particular, $\zeta_c(G) = G$.

We have proved that c) $\Rightarrow$ a) and c) $\Rightarrow$ b). Conversely, the upper and lower central series are clearly central series, so that a) $\Rightarrow$ c) and b) $\Rightarrow$ c).

Finally, it is immediate from Proposition 2.1.5 that a) $\Leftrightarrow$ d).

The theorem is proved.

Definition. A group G, satisfying the conditions of Theorem 2.2.1 is called *nilpotent*, and the least natural number c for which these conditions are satisfied is called the *nilpotency class* of the group G.

Often a group is said to be nilpotent of nilpotency class c if it is, in fact, nilpotent of class $\leq c$.

By Theorem 2.2.1 nilpotent groups of class $\leq c$ constitute a variety which we denote by $\mathfrak{N}_c$. By Proposition 2.1.5 the subgroup $\gamma_k(G)$ is verbal in any group G. If F is a free group, then the factor-group $F/\gamma_{c+1}(F)$ is the free group of the variety $\mathfrak{N}_c$.

We now prove the useful fact that the nilpotency identity may be verified by considering only generating elements of a group.

2.2.2 Theorem. *If a group G is generated by the set M, then it is nilpotent of class $\leq c$ if and only if any commutator of weight $c+1$ in elements of M is trivial.*

Proof. For each natural number k define the subgroup

$$M_k = \langle \{[m_1, m_2, \ldots, m_k] \mid m_i \in M\}^G \rangle$$

to be the normal closure of all commutators of weight k in elements of M. The theorem will be proved if we show that $M_k = \gamma_k(G)$ for all $k \in \mathbb{N}$.

Clearly, $M_k \leq \gamma_k(G)$.

To prove the reverse inclusion we use induction on k. For $k = 1$ we have $M_1 = \langle M \rangle = G = \gamma_1(G)$. Let $k > 1$. Notice that in the factor-group G/M_k the images of all commutators of the form $[m_1, m_2, \ldots, m_{k-1}]$, $m_i \in M$, having weight $k-1$, lie in the centre of G/M_k. This follows from the fact that $[m_1, m_2, \ldots, m_{k-1}, m] \in M_k$ for each $m \in M$ and from the fact that an element centralizing all generators of the group, necessarily lies in the centre of that group.

We have therefore proved that $[M_{k-1}, G] \leq M_k$. By the induction hypothesis we now get

$$\gamma_k(G) = [\gamma_{k-1}(G), G] = [M_{k-1}, G] \leq M_k,$$

as required.

The theorem is proved.

Remark. The analogue of Theorem 2.2.2 does not hold for soluble groups: there exist groups $G = \langle M \rangle$ with $\delta_k(m_1, m_2, \ldots, m_{2^k}) = 1$ for all $m_1, m_2, \ldots, m_{2^k} \in M$, which are not, however, soluble of derived length k.

The next theorem collects some elementary properties of nilpotent groups which are easily proved by induction on the nilpotency class.

2.2.3 Theorem. *Let G be a nilpotent group and let N and H be subgroups of G, N being normal: $N \trianglelefteq G$. Then the following hold:*

a) *if $N \neq 1$, then $[N, G] < N$;*

b) *if $N \neq 1$, then $N \cap Z(G) \neq 1$;*

c) *if $H \neq G$, then $N_G(H) > H$;*

d) *if $HG' = G$, then $H = G$;*

e) *H is subnormal in G.*

Proof. a) Since G is nilpotent, we have $[N, \underbrace{G, \ldots, G}_{k}] = 1$ for some k. But if $[N, G] = N$, then substituting we get $[[N, G], G] = [N, G] = N$, and so on, so that $[N, \underbrace{G, \ldots, G}_{i}] = N$ for all i. Taking $i = k$ this contradicts the condition $N \neq 1$.

b) Choose the least $s \geq 0$, for which $[N, \underbrace{G, \ldots, G}_{s}] \neq 1$. Then $[N, \underbrace{G, \ldots, G}_{s}] \leq$ $\leq N \cap Z(G)$, since

$$[[N, \underbrace{G, \ldots, G}_{s}], G] = [N, \underbrace{G, \ldots, G}_{s+1}] = 1.$$

c) By induction on the nilpotency class c of G. For $c = 1$ the group is abelian and $N_G(H) = G > H$. Now let $c > 1$. If $Z(G) \not\leq H$, then $N_G(H) \geq Z(G)H > H$. If, however, $Z(G) \leq H$, then by the induction hypothesis $N_{\bar{G}}(\bar{H}) > \bar{H}$ where $\bar{H}$ is the image of H in $\bar{G} = G/Z(G)$. The full inverse image of $N_{\bar{G}}(\bar{H})$ is, therefore, larger than H and it coincides with $N_G(H)$, since

$$xH \in N_{\bar{G}}(\bar{H}) \Leftrightarrow (HZ(G))^x \leq HZ(G) \Leftrightarrow H^x \leq HZ(G) = H \Leftrightarrow x \in N_G(H).$$

d) By induction on the nilpotency class c of G. For $c = 1$ we have $G' = 1$ and the result is obvious. For $c > 1$ by the induction hypothesis $\bar{H} = \bar{G}$ where bars denote images in $G/Z(G)$. Therefore, $HZ(G) = G$. Then

$$G' = [HZ(G), HZ(G)] = H'$$

using 2.1.1 c), d), so that $G = HG' = HH' = H$.

e) It is easy to see that the series

$$H \leq H\zeta_1(G) \leq H\zeta_2(G) \leq \ldots \leq H\zeta_c(G) = HG = G,$$

where c is the nilpotency class of G, is a subnormal series: for every i, H normalizes $H\zeta_i(G)$ and $\zeta_{i+1}(G)$ normalizes $H\zeta_i(G)$ since

$$[\zeta_{i+1}(G), H\zeta_i(G)] \leq [\zeta_{i+1}(G), G] \leq \zeta_i(G) \leq H\zeta_i(G).$$

The theorem is proved.

A finite group is nilpotent if and only if it is the direct product of its Sylow subgroups. The necessity of this condition will be obtained as a corollary to a theorem on π-isolators in § 2.5, and its sufficiency will follow from the fact that finite p-groups are nilpotent, which will be proved in § 2.8.

It is possible, however, to prove the necessity by more elementary means using properties of automorphisms of order coprime to the order of the group. Suppose a finite group G is nilpotent; let us prove that it is the direct product of its Sylow subgroups – which is equivalent to the fact that all of the Sylow subgroups are normal. We use induction on the order of the group. Since G is nilpotent, it has a normal subgroup H of prime index p. (One can take for H any of the maximal subgroups of G containing the commutator subgroup.) By the induction hypothesis all of the Sylow subgroups of H are normal in H, and hence all of them are characteristic in H. Therefore all of them are normal in G. So, for each prime number $q \neq p$ the Sylow q-subgroup of G is normal.

It remains to show that the Sylow p-subgroup of G is also normal. If $\bar{a}$ is a non-trivial element of the factor-group G/H, then its order is p, and one can choose its inverse image a lying in a Sylow p-subgroup P of G. It is clear that $G = P \cdot H_{p'}$, where $H_{p'}$ is the Hall p'-subgroup of H. It therefore suffices to prove that $[P, H_{p'}] = 1$, and since $[P \cap H, H_{p'}] = 1$ in the nilpotent group H, it suffices to show that $[a, H_{p'}] = 1$. If this is not so, then the element a, normalizing the subgroup $H_{p'}$, induces on it by conjugation an automorphism of coprime order. By Corollary 1.6.4 b) we get

$$[[H_{p'}, a], a] = [H_{p'}, a] \neq 1.$$

By repeated substitution we therefore obtain

$$[\ldots[[H_{p'}, \underbrace{a], a], \ldots, a}_{n}] = [H_{p'}, a] \neq 1$$

for all $n \in \mathbb{N}$. This obviously contradicts the nilpotency of G.

§ 2.3 Some sufficient conditions for soluble groups to be nilpotent

The following theorem often facilitates the use of induction in proving that a group is nilpotent.

2.3.1 Theorem (P. Hall [30]). *Let N be a normal subgroup of a group G. If N is nilpotent of class k and the $G/[N, N]$ is nilpotent of class c, then G itself is nilpotent and its nilpotency class is bounded by $f(k, c) = (c-1)\frac{k(k+1)}{2} + k$.*

Proof. Using the fact that $G/[N, N]$ is nilpotent of class c we shall prove the formally more general assertion that $\gamma_{f(k,c)+1}(G) \leq \gamma_{k+1}(N)$ for all $k \in \mathbb{N}$. We

proceed by induction on k, simultaneously calculating $f(k, c)$. For $k = 1$ the hypothesis gives $\gamma_{c+1}(G) \leq \gamma_2(N)$, so that we may put $f(1, c) = c$.

Now suppose that $k > 1$ and that $\gamma_{f(k,c)+1}(G) \leq \gamma_{k+1}(N))$. For any $s \in \mathbb{N}$ consider the commutator subgroup

$$\gamma_{f(k,c)+s+1}(G) = [\gamma_{f(k,c)+1}(G), \underbrace{G, \ldots, G}_{s}] \leq$$

$$\leq [\gamma_{k+1}(N), \underbrace{G, \ldots, G}_{s}] = [\underbrace{N, \ldots, N}_{k+1}, \underbrace{G, \ldots, G}_{s}].$$

Repeated application of 2.1.4 gives

$$[\underbrace{N, \ldots, N}_{k+1}, G] \leq [[N, G], N, \ldots, N] \cdot [N, [N, G], N, \ldots, N] \cdot$$

$$\cdot \ldots \cdot [N, \ldots, N, [N, G]]$$

and furthermore

$$[\underbrace{N, \ldots, N}_{k+1}, \underbrace{G, \ldots, G}_{s}] \leq$$

$$\leq \prod_{i_1 + \ldots + i_{k+1} = s} [[N, \underbrace{G, \ldots, G}_{i_1}], \ldots, [N, \underbrace{G, \ldots, G}_{i_{k+1}}]] \tag{2.3.2}$$

(here, by definition, $[N, \underbrace{G, \ldots, G}_{0}] = N$).

For sufficiently large $s \geq (k+1)(c-1)+1$ at least one of the summands i_j of the sum $i_1 + \ldots + i_{k+1} = s$ will be greater than $c - 1$, and each commutator on the right-hand side of (2.3.2) will contain a subcommutator

$$[N, \underbrace{G, \ldots, G}_{i_r}] \text{ with } i_r \geq c. \tag{2.3.3}$$

Now 2.1.4 allows us to "transpose" any of the normal subgroups to the beginning of the commutator subgroups:

$$[[A, B], C] \leq [[A, C], B] \cdot [[A, [B, C]] = [[C, A], B] \cdot [[C, B], A].$$

Repeated application of this enables us to "pull out" the long subcommutators of the form (2.3.3) in the right-hand side of (2.3.2) to the beginning, so that for

$s = (k+1)(c-1)+1$ we get:

$$\prod_{i_1+\ldots+i_{k+1}=s} [[N, \underbrace{G, \ldots, G}_{i_1}], \ldots, [N, \underbrace{G, \ldots, G}_{i_{k+1}}]] \leq [[N, \underbrace{G, \ldots, G}_{c}], \underbrace{N, \ldots, N}_{k}].$$

(Here we have replaced $[N, \underbrace{G, \ldots, G}_{i_r}]$ with $i_r \geq c$ by the larger subgroup $[N, \underbrace{G, \ldots, G}_{c}]$, and $[N, \underbrace{G, \ldots, G}_{i_j}]$ for other i_j by N.)

As a result for $s = (k+1)(c-1)+1$ we have

$$[\gamma_{f(k,c)+1}(G), \underbrace{G, \ldots, G}_{s}] \leq [[N, \underbrace{G, \ldots, G}_{c}], \underbrace{N, \ldots, N}_{k}] \leq$$

$$\leq [[\underbrace{G, G, \ldots, G}_{c+1}], \underbrace{N, \ldots, N}_{k}] \leq [[N, N], \underbrace{N, \ldots, N}_{k}] = \gamma_{k+2}(N).$$

The last step again uses the condition $\gamma_{c+1}(G) \leq [N, N]$.

Thus, if we put $f(k+1, c) = f(k, c) + (k+1)(c-1)+1$, then $\gamma_{f(k+1,c)+1}(G) \leq \leq \gamma_{k+2}(N)$, as required.

It is not difficult to compute, starting from $f(1, c) = c$, that

$$f(k, c) = (c-1) + 1 + 2(c-1) + 1 + 3(c-1) + 1 + \ldots$$
$$\ldots + k(c-1) + 1 = (c-1)\frac{k(k+1)}{2} + k.$$

The theorem is proved.

2.3.4 Corollary. *If in a variety of groups (or in a variety of groups with operators) $\mathfrak{M}$, all metabelian groups are nilpotent of class $\leq c$, then any soluble group in $\mathfrak{M}$ is nilpotent and its nilpotency class is (s, c)-bounded by some function $g(s, c)$, where s is the derived length of the group. (In other words, if $\mathfrak{M} \cap \mathfrak{A}^2 \subseteq \mathfrak{N}_c$ then $\mathfrak{M} \cap \mathfrak{A}^s \subseteq \mathfrak{N}_{g(s,c)}$.)*

Proof. By induction on the derived length s of the group. The case $s = 1$ is trivial, and the case $s = 2$ is covered by hypothesis. For $s > 2$, we consider the free countably-generated group F of the variety $\mathfrak{M} \cap \mathfrak{A}^s$. It is sufficient to prove that F is nilpotent of class $\leq g(s, c)$. By the induction hypothesis F' is nilpotent of class $\leq g(s-1, c)$, and F/F'' is nilpotent of class $\leq c$ by hypothesis. Therefore, by Theorem 2.3.1, F is nilpotent of class

$$\leq (c-1)\frac{g(s-1, c)(g(s-1, c)+1)}{2} + g(s-1, c).$$

It is easy to see that the proof of Corollary 2.3.4, based on Theorem 2.3.1, gives as an explicit expression for $g(s, c)$ a polynomial in c with leading term $2^{2^{s-2}-2^{s-1}+1} \cdot c^{2^{s-1}-1}$. However, this bound may be significantly improved.

2.3.5 Theorem. *If, in a variety of groups (or in a variety of groups with operators) $\mathfrak{M}$, all metabelian groups are nilpotent of class $\leq c$, then any soluble group in $\mathfrak{M}$ is nilpotent and its nilpotency class does not exceed $\frac{c^s-1}{c-1}$ where s is the derived length of the group.*

Proof. We shall at first prove the theorem in the technically simpler case of (ordinary) varieties of groups, and then we shall indicate how the arguments should be modified to include varieties with operators.

We use induction on the derived length s. The case $s = 1$ is trivial and the case $s = 2$ is covered by the hypothesis of the theorem.

Before making the induction step we find conditions under which a semidirect product of two groups belongs to the given variety $\mathfrak{M}$. Let the group word $w(\bar{x}) = w(x_1, x_2, \ldots, x_n)$ be one of the identities, defining $\mathfrak{M}$ (see §1.8). Regarding y_i, z_i $(i = 1, 2, \ldots, n)$ also as free generators of a free group, we represent the value $w(\bar{y}\bar{z}) = w(y_1 z_1, y_2 z_2, \ldots, y_n z_n)$ in the form

$$w(\bar{y}\bar{z}) = w(\bar{z}) \cdot c_w(\bar{y}, \bar{z}), \tag{2.3.6}$$

where $c_w(\bar{y}, \bar{z})$ is an element of the normal closure

$$\langle\langle y_1, y_2, \ldots, y_n\rangle^{\langle z_1, \ldots, z_n\rangle}\rangle$$

of the subgroup $\langle y_1, y_2, \ldots, y_n\rangle$ in the group $\langle y_1, y_2, \ldots, y_n, z_1, z_2, \ldots, z_n\rangle$, so that the word $c_w(\bar{y}, \bar{z})$ has the form

$$y_{i_1}^{v_1} \cdot y_{i_2}^{v_2} \cdot \ldots \cdot y_{i_r}^{v_r}, \quad v_i = \pm t_i(\bar{z}),$$

where $t_i(\bar{z})$ are some words in $z_1, z_2, \ldots, z_n$.

(The presentation (2.3.6) may be obtained by transposing all the powers of the z_i to the beginning, preserving the order of their occurrence, by using the formulae

$$y_i^v \cdot z_j^e = z_j^e \cdot y_i^{vz_j^e},$$

where $e = \pm 1$, $v = \pm t(\bar{z})$ and $t(\bar{z})$ is some word in $z_1, z_2, \ldots, z_n$. Another explanation of (2.3.6) consists in the fact that the image of $w(\bar{y}\bar{z})$ under the homomorphism extending the mapping $y_i \to 1$, $z_i \to z_i$ $(i = 1, 2, \ldots, n)$ is obviously equal to $w(\bar{z})$; the kernel of this homomorphism is the normal closure

$\langle\langle y_1, y_2, \ldots, y_n\rangle^{\langle z_1,\ldots,z_n\rangle}\rangle$, and $w(\bar{z})$ is an element of that coset of the kernel which contains $w(\bar{y}\bar{z})$.)

We fix notation $c_w(\bar{y}, \bar{z})$ from (2.3.6) for the word w.

2.3.7 Lemma. *A semidirect product $A \leftthreetimes B$ of groups A and B belongs to the variety $\mathfrak{M}$ if and only if the group B belongs to $\mathfrak{M}$ and $c_w(\bar{a}, \bar{b}) = 1$ for each identity w of $\mathfrak{M}$ and arbitrary elements $a_1, a_2, \ldots, a_n \in A$ and $b_1, b_2, \ldots, b_n \in B$.*

Proof. If $A \leftthreetimes B \in \mathfrak{M}$, then $B \in \mathfrak{M}$, since B is a factor-group of $A \leftthreetimes B$. Therefore, $w(\bar{b}) = 1$ for any $b_1, b_2, \ldots, b_n \in B$. Furthermore, $w(\bar{a}\bar{b}) = 1$ for any $a_1, a_2, \ldots, a_n \in A$. Hence, by virtue of (2.3.6), we have $c_w(\bar{a}, \bar{b}) = 1$.

Suppose, conversely, that $B \in \mathfrak{M}$ and $c_w(\bar{a}, \bar{b}) = 1$ for any $a_1, a_2, \ldots, a_n \in A$, $b_1, b_2, \ldots, b_n \in B$. Then also $w(\bar{b}) = 1$ and, therefore, by (2.3.6), $w(\bar{a}\bar{b}) = 1$. But arbitrary elements of the semidirect product $A \leftthreetimes B$ have the form $a_1b_1, a_2b_2, \ldots, a_nb_n$. Therefore $A \leftthreetimes B$ satisfies each identity w of $\mathfrak{M}$, that is $A \leftthreetimes B \in \mathfrak{M}$.

The lemma is proved.

2.3.8 Lemma. *Suppose that the group G belongs to the variety $\mathfrak{M}$. Then for any normal subgroup A of G and any normal subgroup $C \leq C_G(A)$ the natural semidirect product $A \leftthreetimes (G/C)$ also belongs to $\mathfrak{M}$.*

Proof. Set $B = G/C$. By Lemma 2.3.7 it is sufficient to show that B belongs to $\mathfrak{M}$ and $c_w(\bar{a}\bar{b}) = 1$ for every identity w of the variety $\mathfrak{M}$ and any elements $a_1, a_2, \ldots, a_n \in A$ and $b_1, b_2, \ldots, b_n \in B$. It is clear that $B \in \mathfrak{M}$ since it is a factor-group of $G \in \mathfrak{M}$. Now let us choose in G arbitrary inverse images $\hat{b}_j$ of the elements b_i, $i = 1, 2, \ldots, n$. Since $w(\bar{a}\hat{b}) = 1$ and $w(\hat{b}) = 1$, by (2.3.6) we have $c_w(\bar{a}, \hat{b}) = 1$. It remains to note, that

$$c_w(\bar{a}, \bar{b}) = c_w(\bar{a}, \hat{b}),$$

since the word $c_w(\bar{a}, \bar{b})$ has the form

$$a_{i_1}^{v_1} \cdot a_{i_2}^{v_2} \cdot \ldots \cdot a_{i_r}^{v_r}, \quad v_i = \pm t_i(\bar{b}),$$

where $t_i(\bar{b})$ are some words in $b_1, b_2, \ldots, b_n \in B$, and, clearly,

$$a_i^{t(\hat{b})} = a_i^{t(\bar{b})}$$

by the definition of the action of $B = G/C$ on A.

The lemma is proved.

Now we proceed with the induction step in the proof of Theorem 2.3.5.

2.3.9 Lemma. *Suppose that the variety $\mathfrak{M}$ satisfies the conditions of Theorem 2.3.5 and let G be a soluble group in $\mathfrak{M}$ of derived length s. Then*

$$[G^{(s-1)}, \underbrace{G, G, \ldots, G}_{c^{s-1}}] = 1.$$

Proof. Induction on the derived length s. The case $s = 1$ is trivial and, if $s = 2$, we need to show that $[G', \underbrace{G, \ldots, G}_{c}] = 1$. This follows from the hypothesis of Theorem 2.3.5, since $\gamma_{c+1}(G) = 1$ for the metabelian group $G \in \mathfrak{M}$.

Suppose that $s > 2$. Then $[G^{(s-1)}, G']$ is normal in G and $G/[G^{(s-1)}, G']$ lies in $\mathfrak{M}$. As G' centralizes the section $V = G^{(s-1)}/[G^{(s-1)}, G']$, the group G/G' acts on it by conjugation. By Lemma 2.3.8 the natural semidirect product $V \leftthreetimes G/G'$ also belongs to $\mathfrak{M}$. The group $V \leftthreetimes G/G'$ is metabelian and hence $[V, \underbrace{G/G', \ldots, G/G'}_{c}] = 1$ by the hypothesis of Theorem 2.3.5. This may be expressed in terms of the inverse images as:

$$[G^{(s-1)}, \underbrace{G, G, \ldots, G}_{c}] \leq [G^{(s-1)}, G'].$$

If we replace $G^{(s-1)}$ by $W = [G^{(s-1)}, \underbrace{G', \ldots, G'}_{i}]$ in this argument, we obtain that

$$[W, \underbrace{G, G, \ldots, G}_{c}] \leq [W, G'] = [G^{(s-1)}, \underbrace{G', G', \ldots, G'}_{i+1}].$$

An obvious induction on j yields

$$[G^{(s-1)}, \underbrace{G, G, \ldots, G}_{c \cdot j}] \leq [G^{(s-1)}, \underbrace{G', G', \ldots, G'}_{j}],$$

and, in particular,

$$[G^{(s-1)}, \underbrace{G, G, \ldots, G}_{c^{s-1}}] \leq [G^{(s-1)}, \underbrace{G', G', \ldots, G'}_{c^{s-2}}].$$

But the right-hand side here is 1 by the induction hypothesis applied to G' since G' is soluble of derived length $s - 1$ (note that $(G')^{(s-2)} = G^{(s-1)}$).

The lemma is proved.

We now finish the proof of Theorem 2.3.5. Suppose that $s > 2$ and let G be a soluble group in $\mathfrak{M}$ of derived length s. By the induction hypothesis applied to $G/G^{(s-1)}$, we have $\gamma_{f(s-1,c)+1}(G) \leq G^{(s-1)}$, where $f(s-1,c) = \frac{c^{s-1}-1}{c-1}$.

Hence, by Lemma 2.3.9,

$$\gamma_{f(s-1,c)+c^{s-1}+1}(G) = [\gamma_{f(s-1,c)+1}(G), \underbrace{G, G, \ldots, G}_{c^{s-1}}] \leq$$

$$\leq [G^{(s-1)}, \underbrace{G, G, \ldots, G}_{c^{s-1}}] = 1.$$

Hence G is nilpotent of class at most

$$f(s-1,c) + c^{s-1} = \frac{c^{s-1}-1}{c-1} + c^{s-1} = \frac{c^s - 1}{c-1},$$

which is the required bound.

The proof of the theorem in the case of ordinary varieties of groups is now complete. In the case of varieties of groups with operators Ω, the proof may be obtained by more or less word for word repetition of the same arguments. The difference lies in the fact that in this case subgroups are Ω-invariant subgroups, the free generators freely generate a free Ω-group, the identities (Ω-words) are elements of the free Ω-group, etc. In particular, the elements $c_w(\bar{y}, \bar{z})$ from (2.3.6) have the form

$$y_{i_1}^{v_1} \cdot y_{i_2}^{v_2} \cdot \ldots \cdot y_{i_r}^{v_r}, \quad v_i = \pm t_i(\bar{z}),$$

where $t_i(\bar{z})$ are elements of the semidirect product $\langle z_1, z_2, \ldots, z_n \rangle \leftthreetimes \Omega$.

§ 2.4 The Schur-Baer Theorem and its converses

Although we are primarily concerned with converses of the Schur-Baer Theorem, we prove it here for the sake of completeness in spite of the fact that the methods used are different from anything we have yet mentioned.

2.4.1 Theorem. *If the index of the centre of a group G is finite (and equals n), then the commutator subgroup of G is also finite (and its order is n-bounded).*

Proof. Put $Z = Z(G)$ for short and decompose G into a union of cosets of Z:

$$G = g_1 Z \cup g_2 Z \cup \ldots \cup g_n Z.$$

As follows from 2.1.1 c), d), all commutators of elements of G are of the form $[g_i, g_j]$, $i, j = 1, 2, \ldots, n$. Therefore the commutator subgroup of G is contained in the commutator subgroup of $\langle g_1, g_2, \ldots, g_n \rangle$ and so we may assume that $G = \langle g_1, g_2, \ldots, g_n \rangle$. Then Z, as a subgroup of finite index n in the finitely generated (n-generated) group G, is also finitely generated by an n-bounded number of elements by Schreier's Theorem. Now for any $g \in G$ and for $i = 1, 2, \ldots, n$ put

$$g_i g = z_i(g) g_{i(g)},$$

where $z_i(g) \in Z$, and $g_{i(g)} \in \{g_1, \ldots, g_n\}$. Straightforward calculation shows that the mapping

$$\vartheta \colon g \to \prod_{i=1}^{n} z_i(g)$$

is a homomorphism of G into Z (the so-called transfer homomorphism). It is easy to see that $\vartheta(z) = z^n$ for all $z \in Z$. Since Z is an abelian group, $\operatorname{Ker} \vartheta \supseteq G'$. As a result we get

$$1 = \vartheta(G' \cap Z) = (G' \cap Z)^n,$$

which means that $G' \cap Z$ has finite exponent dividing n. Since $G' \cap Z$ is a subgroup of the finitely generated abelian group Z, it is itself finitely generated (with the same bound on the number of generators). Hence $G' \cap Z$ is finite (and its order is n-bounded).

Since $|G' : G' \cap Z| \leq |G : Z| = n$, we obtain that G' is also finite of n-bounded order.

The theorem is proved.

The converse of Theorem 2.4.1 is false – it is easy to construct an example of a group with finite commutator subgroup whose centre is small (and coincides with the commutator subgroup). For example, take

$$G = \langle a_i, b_i \mid i \in \mathbb{N}; \quad [a_i, a_j] = [b_i, b_j] = [a_i, b_j] = 1 \text{ for } i \neq j;$$
$$[a_i, b_i] = c, [a_i, c] = [b_i, c] = 1 \text{ for all } i; \ c^2 = 1 \rangle.$$

Then $G' = \langle c \rangle$ is a group of order 2 while $Z(G) = \langle c \rangle$ has infinite index.

However there are some valid weaker assertions of a converse nature.

2.4.2 Theorem. *If the set of all commutators $\{[g, h] \mid g, h \in G\}$ of the elements of a group G is finite (and consists of n elements), then $|G : \zeta_2(G)|$ is also finite (and n-bounded).*

Proof. We shall need the following simple lemma.

2.4.3 Lemma. *For any element a of an arbitrary group G there is a one-to-one correspondence between the set of commutators of the form $\{[g, a] \mid g \in G\}$ and the set of (right) cosets of the centralizer $C_G(a)$.*

Proof. We associate with the coset $C_G(a)g$ the commutator $[g, a]$. This mapping is well defined since, for any other representative of the same coset $g' = cg$, $c \in C_G(a)$, we have

$$[g', a] = [cg, a] = [c, a]^g \cdot [g, a] = 1^g \cdot [g, a] = [g, a].$$

It is clear that this mapping is onto the whole set $\{[g, a] \mid g \in G\}$. Also distinct elements have distinct images: if $[x, a] = [y, a]$, then

$$[xy^{-1}, a] = [x, a]^{y^{-1}}[y^{-1}, a] = [y, a]^{y^{-1}}[y^{-1}, a] = [yy^{-1}, a] = 1,$$

which means that $xy^{-1} \in C_G(x)$.

The lemma is proved.

To make it easier to understand more complicated arguments which will occur in Chapters 4 and 5, we give another version of the proof of this simple lemma. For a fixed element $a \in G$ define the mapping

$$\mu_a \colon x \to [x, a]$$

from G into the set of commutators of the elements of G. Though, of course, this mapping is not a homomorphism, we have shown, in fact, that the full inverse images of the $[x, a]$ are the cosets of $C_G(x)$, and the "kernel" – the full inverse image of the identity element – is $C_G(x)$, so that its index is equal to the cardinality of $\{[x, a] \mid x \in G\}$. (In Chapters 4 and 5 some generalized centralizers of bounded indices are defined in an analogous way as the kernels of certain homomorphisms.)

We now return to the proof of Theorem 2.4.2. By Lemma 2.4.3, the hypothesis of the theorem imply that $|G : C_G(g)| \leq n$ for any $g \in G$.

Now suppose that $c_1, c_2, \ldots, c_n$ are all the commutators of elements of G. For each c_i we fix representatives $b_{i,j}$ of the cosets of $C_G(c_i)$. The total number of elements $b_{i,j}$ is obviously n-bounded, since $|G : C_G(c_i)| \leq n$ for all i. It follows from the proof of Lemma 2.4.3 that for any $g \in G$ and for each c_i there is a representative $b_{i,j(g)}$ such that $[g, c_i] = [b_{i,j(g)}, c_i]$.

Now put

$$M = \bigcap_{i,j} C_G(b_{i,j}) \cap \bigcap_k C_G(c_k).$$

where the intersection is taken over all c_k and over all $b_{i,j}$. Since both the index of each centralizer and the number of centralizers is n-bounded, the index of the intersection $|G : M|$ is also n-bounded.

It remains to prove that $M \subseteq \zeta_2(G)$. It is clear that it is sufficient to show that $[m, g_1, g_2] = 1$ for any $m \in M$ and any $g_1, g_2 \in G$.

But the commutator $[m, g_1]$ coincides with one of the c_i, and hence,

$$[m, g_1, g_2] = [c_i, g_2] = [c_i, b_{i,j}] = [m, g_1, b_{i,j}]$$

for some $b_{i,j}$. We now apply Witt's identity 2.1.1 e):

$$[m, g_1, b_{i,j}]^{g_1^{-1}} \cdot [g_1^{-1}, b_{i,j}^{-1}, m]^{b_{i,j}} \cdot [b_{i,j}, m^{-1}, g_1^{-1}]^m = 1.$$

Here the second factor is 1, since $[g_1^{-1}, b_{i,j}^{-1}]$ coincides with one of the c_s, and $m \in C_G(c_s)$ by the construction of M. For the same reason $m \in C_G(b_{i,j})$, so that the third factor is also 1. Therefore, $[m, g_1, b_{i,j}]^{g_1^{-1}} = 1$, whence $[m, g_1, g_2] = [m, g_1, b_{i,j}] = 1$, as required.

The theorem is proved.

If, as well as satisfying the hypothesis of Theorem 2.4.2, the group G is finitely generated (by s elements), then even the index of the centre of G is finite (and (n, s)-bounded).

2.4.4 Theorem. *Suppose that G is an s-generator group, s finite, and that $\{[g, h] \mid g, h \in G\}$ is finite and consists of n elements. Then $|G : Z(G)|$ is also finite and (n, s)-bounded.*

Proof. By Lemma 2.4.3 we have $|G : C_G(g)| \leq n$ for each $g \in G$. Suppose that $G = \langle a_1, a_2, \ldots, a_s \rangle$. Then $\bigcap_{i=1}^{s} C_G(a_i)$ has finite (n, s)-bounded index and is contained in the centre of G, because it centralizes each generator of G.

The theorem is proved.

We also recall the following general theorems of P. Hall [29]. If, in an arbitrary group G, the subgroup $\gamma_{k+1}(G)$ is finite (of order n), then $|G : \zeta_{2k}(G)|$ is also finite (and (n, k)-bounded). There are examples of groups G with $\gamma_{k+1}(G)$ finite, in which the index $|G : \zeta_{2k-1}(G)|$ is already infinite, so that $2k$ is best possible. If G is in addition finitely generated (by s elements), then $|G : \zeta_k(G)|$ is also finite (and (n, k, s)-bounded).

We prove here only the following simpler theorem which will be used later.

2.4.5 Theorem. *Suppose that G is an arbitrary group with $\gamma_k(G)$ finite (of order n). Then G has a nilpotent subgroup of nilpotency class k whose index is finite (and (k, n)-bounded).*

Proof. The centralizer $C_G(\gamma_k(G))$ has all the desired properties. Indeed, its index is n-bounded, since $G/C_G(\gamma_k(G))$ embeds in the automorphism group of $\gamma_k(G)$ which has order $\leq n!$. We prove that $C_G(\gamma_k(G))$ is nilpotent of class $\leq k$. It suffices to show that $[c_1, c_2, \ldots, c_{k+1}] = 1$ for any $c_1, \ldots, c_{k+1} \in C_G(\gamma_k(G))$. But this follows at once from the fact that $[c_1, c_2, \ldots, c_k] \in \gamma_k(G)$ and $c_{k+1} \in C_G(\gamma_k(G))$.

The theorem is proved.

(With an eye on future generalizations in Chapters 4 and 5 we observe that the subgroup $C_G(\gamma_k(G))$, constructed in the proof of Theorem 2.4.5, may be defined as the intersection of the "kernels" of the mappings $\mu_a\colon x \to [x, a]$, where $a \in \gamma_k(G)$, defining "kernel" to mean the full inverse image of the identity element.)

Theorem 2.4.1 was proved by Schur in [127]. Baer [3] also proved that if in an arbitrary group the index of the k-th member of the upper central series is finite and equals n, then the order of the $(k+1)$-th member of the lower central series is also finite and (n, k)-bounded.

§ 2.5 Lower central series. Isolators

It is convenient to use the language of tensor products of abelian groups (see §1.2) to describe linear properties of lower central series. The same facts may, of course, be obtained by direct commutator calculations, but the use of tensor products provides substantial "economies of thought". An account of the theory of isolators is given as an application at the end of this section.

2.5.1 Lemma. *Let A and B be subgroups of a group such that $[A, B] \leq Z(\langle A, B\rangle)$. Then the mapping*

$$\mu\colon (a, b) \to [a, b]$$

is linear in both arguments, that is $\mu(a_1a_2, b) = [a_1, b][a_2, b]$ and $\mu(a, b_1b_2) = [a, b_1][a, b_2]$, for any $a, a_1, a_2 \in A$ and $b, b_1, b_2 \in B$.

Proof. Straightforward application of 2.1.1 c) and d). For example,

$$\mu(a_1a_2, b) = [a_1, b]^{a_2} \cdot [a_2, b] = [a_1, b][a_2, b],$$

since $[a_1, b] \in Z(\langle A, B\rangle)$ by hypothesis.

2.5.2 Theorem. *Suppose that G is an arbitrary group and $\bar{x}$ denotes the image of an element x in the factor-group G/G'. Then, for any natural number n, the mapping*

$$\vartheta\colon (\bar{x}_1 \otimes \bar{x}_2 \otimes \ldots \otimes \bar{x}_n) \to [x_1, x_2, \ldots, x_n]\gamma_{n+1}(G)$$

induces a homomorphism of the tensor product $\underbrace{G/G' \otimes \ldots \otimes G/G'}_{n}$ *onto* $\gamma_n(G)/\gamma_{n+1}(G)$.

Proof. We set $\gamma_i = \gamma_i(G)$, for short. For some k suppose that $\hat{a}$ and $\bar{b}$ are images of $a \in \gamma_k$ and $b \in G$ in γ_k/γ_{k+1} and G/G' respectively. Consider the mapping

$$\mu\colon \hat{a} \otimes \bar{b} \to [a, b]\gamma_{k+2}$$

from $\gamma_k/\gamma_{k+1} \otimes G/G'$ into $\gamma_{k+1}/\gamma_{k+2}$. This mapping is well defined, since $[\gamma_{k+1}, G] \leq \gamma_{k+2}$ and $[\gamma_k, G'] \leq \gamma_{k+2}$. Since $[\gamma_k, G] = \gamma_{k+1} \leq Z(G/\gamma_{k+2})$, it is linear in each of its arguments by Lemma 2.5.1. Hence (see § 1.2), the mapping μ induces a homomorphism of $\gamma_k/\gamma_{k+1} \otimes G/G'$ into $\gamma_{k+1}/\gamma_{k+2}$. This is a homomorphism onto the whole factor-group $\gamma_{k+1}/\gamma_{k+2}$, since $[\gamma_k, G] = \gamma_{k+1}$.

We now prove the theorem by induction on n. For $n = 1$ the result is trivial. For $n > 1$, it follows from the induction hypothesis that the mapping

$$\vartheta_1\colon (\bar{x}_1 \otimes \ldots \otimes \bar{x}_{n-1} \otimes \bar{x}_n) \to [x_1, \ldots, x_{n-1}]\gamma_n \otimes \bar{x}_n$$

induces a homomorphism of $\underbrace{G/G' \otimes \ldots \otimes G/G'}_{n}$ onto $\gamma_{n-1}/\gamma_n \otimes G/G'$. The composition of ϑ_1 and μ (as defined above with $k = n - 1$) obviously induces the required homomorphism

$$\vartheta\colon (\bar{x}_1 \otimes \bar{x}_2 \otimes \ldots \otimes \bar{x}_n) \xrightarrow{\vartheta_1} [x_1, \ldots, x_{n-1}]\gamma_n \otimes \bar{x}_n \xrightarrow{\mu} \xrightarrow{\mu} [x_1, x_2, \ldots, x_n]\gamma_{n+1}(G)$$

of $\underbrace{G/G' \otimes \ldots \otimes G/G'}_{n}$ onto γ_n/γ_{n+1}.

The theorem is proved.

Now we are in a position to use linear properties of tensor products of abelian groups.

2.5.3 Corollary. *Suppose that G is an arbitrary group such that G/G' has exponent m. Then $\gamma_n(G)/\gamma_{n+1}(G)$ has exponent dividing m, for all n.*

Proof. The result follows from Theorem 2.5.2 and from the fact that the exponent of a tensor product of abelian groups divides the exponents of each of the factors (see § 1.2).

2.5.4 Corollary. *Suppose that G is a nilpotent group of class c such that G/G' has exponent m. Then G has finite exponent dividing m^c.*

Proof. This follows from the preceding corollary.

2.5.5 Corollary. *Suppose that for some set of prime numbers π and some $k \in \mathbb{N}$ the orders of all elements of the factor-group $\gamma_k(G)/\gamma_{k+1}(G)$ are finite and π-numbers. Then the same holds for $\gamma_n(G)/\gamma_{n+1}(G)$, for all $n \geq k$.*

Proof. It is clearly sufficient to consider $\gamma_{k+1}(G)/\gamma_{k+2}(G)$. In the proof of Theorem 2.5.2 we established the existence of a homomorphism of $\gamma_k(G)/\gamma_{k+1}(G) \otimes G/G'$ onto $\gamma_{k+1}(G)/\gamma_{k+2}(G)$. Now $\gamma_k(G)/\gamma_{k+1}(G) \otimes G/G'$ is generated by elements of the form $a \otimes b$, where $a \in \gamma_k(G)/\gamma_{k+1}(G)$, and $b \in G/G'$. By hypothesis we have, in additive notation, $ma = 0$ for some π-number m. Therefore $m(a \otimes b) = ma \otimes b = 0$.

2.5.6 Corollary. *If a group G is finitely generated by d elements, then $\gamma_n(G)/\gamma_{n+1}(G)$ is generated by d^n elements, for each n.*

Proof. This follows from Theorem 2.5.2, since the rank of the tensor product of abelian groups does not exceed the product of the ranks of the factors – see § 1.2.

It is not difficult to show that a subgroup of a finitely generated nilpotent group is also finitely generated, and the number of its generators is bounded in terms of the number of generators of the group and its nilpotency class. This follows from Corollary 2.5.6 and from the following general result.

2.5.7 Proposition. *If a group G has a subnormal series of length n with cyclic factor-groups, then each of its subgroups may be generated by n elements.*

Proof. Induction on n. For $n = 1$, we have that G and all of its subgroups are cyclic. Suppose that $n > 1$, and let

$$G = G_1 \trianglerighteq G_2 \trianglerighteq \ldots \trianglerighteq G_n \trianglerighteq 1$$

be a subnormal series of G with cyclic factor-groups. If H is an arbitrary subgroup of G, then the factor-groups

$$H/(H \cap G_2) \cong HG_2/G_2 \leq G/G_2$$

are cyclic. Hence $H/(H \cap G_2)$ is generated by a single element, and an arbitrary inverse image of this element together with $H \cap G_2$ generate H. By the induction hypothesis $H \cap G_2$ is generated by $n-1$ elements. Hence H is generated by n elements.

The proposition is proved.

We now apply Theorem 2.5.2 to construct isolators and the largest periodic subgroups in nilpotent groups.

Definition. Let π be a subset of the set of all prime numbers. A natural number is said to be a *π-number*, if it is the product of numbers from π.

Definition. For any subset M of a group G the subset

$$I_\pi(M) = \{x \in G \mid x^n \in M \text{ for some } \pi\text{-number } n = n(x)\}$$

is called the *π-isolator* of M (or simply the *isolator* of M if π is the set of all primes). If the set $\pi = \{p\}$ consists of only one element, then we write $I_p(M)$ instead of $I_{\{p\}}(M)$.

The following result shows how important the concept of isolator is in the theory of nilpotent groups.

2.5.8 Theorem. *Suppose that G is a nilpotent group, π is a set of of prime numbers and H is a subgroup of G. Then the π-isolator $I_\pi(H)$ is a subgroup of G.*

Proof. We may clearly assume that $G = \langle I_\pi(H)\rangle$. We note that this assumption is inherited by all factor-groups of G. We therefore have to prove that $G = I_\pi(H)$. We use induction on the nilpotency class c of the group G.

The result is easily verified for abelian groups. Therefore, for any element $g \in G$, there is a π-number m such that $g^m \in \gamma_2(G)H$.

Now consider an arbitrary commutator $[g_1, g_2, \ldots, g_c]$ of weight c, where c is the class of G. To each element g_i there corresponds a π-number m_i and an element $k_i \in \gamma_2(G)$ such that $g_i^{m_i}k_i \in H$. Since $\gamma_c(G)/\gamma_{c+1}(G) = \gamma_c(G)$ is a homomorphic image of $\underbrace{G/G' \otimes \ldots \otimes G/G'}_{c}$ (Theorem 2.5.2), it follows that

$$[g_1^{m_1}k_1, g_2^{m_2}k_2, \ldots, g_c^{m_c}k_c] = [g_1, g_2, \ldots, g_c]^{m_1 m_2 \ldots m_c}.$$

Now $m_1 m_2 \ldots m_c$ is clearly a π-number and the left-hand side of the above equation belongs to H, since $g_i^{m_i}k_i \in H$, for all i. Thus $[g_1, g_2, \ldots, g_c] \in I_\pi(H)$. Therefore, the abelian group $\gamma_c(G)$, generated by the commutators $[g_1, g_2, \ldots, g_c]$ (see 2.1.5), is contained in $I_\pi(H)$.

By the induction hypothesis applied to $G/\gamma_c(G)$, for any element $g \in G$, there is a π-number m such that $g^m \in H\gamma_c(G)$, that is $g^m = hz$ for some $z \in \gamma_c(G)$ and $h \in H$. Since $\gamma_c(G) \subseteq I_\pi(H)$, as was proved above, there is a π-number s such that $z^s \in H$. As a result $g^{ms} = h^s z^s \in H$. But ms is also a π-number so that $g \in I_\pi(H)$, as required.

The theorem is proved.

Definition. The isolator $I_\pi(1)$ of the identity subgroup of a nilpotent group G is also a subgroup by Theorem 2.5.8 called the *π-torsion part* of G (the *torsion* or *periodic part* of the group, if π is the set of all primes).

Obviously, if the isolator $I_\pi(1)$ is a subgroup, then it is characteristic. Therefore we have the following corollary to Theorem 2.5.8.

2.5.9 Corollary. *If a periodic group is nilpotent, then it is the direct product of its Sylow subgroups.*

We now record a result which we shall need later.

2.5.10 Proposition. *The periodic part $T(G)$ of a finitely generated nilpotent group G is finite.*

Proof. It follows from Corollary 2.5.6 that G has a finite normal series with cyclic factors. Hence the subgroup $T(G)$ is finitely generated by Proposition 2.5.7. By Corollary 2.5.6 all factors of the lower central series of $T(G)$ are finitely generated as well. Since they, like $T(G)$, are periodic groups, they are finite. Thus $T(G)$ is also finite.

The proposition is proved.

To conclude this section we remark that to many theorems about lower central series there correspond dual results about properties of upper central series – see, for example, the book of Warfield [153]. Thus, for example, if the centre of a nilpotent group has no π-torsion, then the group itself is π-torsion-free, etc. However, we do not know whether there is any automatic procedure for obtaining such dual results.

§ 2.6 Nilpotent groups without torsion

In this section π will always denote a set of prime numbers. All our results will be proved for nilpotent groups without π-torsion (see § 2.5). Their proofs are not

essentially any more difficult than their proofs in the special case of torsion-free groups, that is, where π is the set of all primes.

The following theorem shows that extracting π-roots in nilpotent groups without π-torsion is a well-defined operation.

2.6.1 Theorem. *Let G be a nilpotent group without π-torsion, that is, such that $I_\pi(1) = 1$. Then for $x, y \in G$*

a) *if $x^n = y^n$ for a π-number n, then $x = y$;*

b) *if $x^m y^n = y^n x^m$ for π-numbers m, n, then $xy = yx$.*

Proof. a) Induction on the nilpotency class c of G. For $c = 1$ the group is abelian and

$$x^n = y^n \Leftrightarrow x^n y^{-n} = 1 \Leftrightarrow (xy^{-1})^n = 1,$$

and, since G has no π-torsion by hypothesis, $xy^{-1} = 1$. Thus $x = y$. For $c > 1$, we have $x \equiv y \pmod{Z(G)}$ by the induction hypothesis, that is, $x = yz$, $z \in Z(G)$. By hypothesis $(yz)^n = x^n = y^n \Rightarrow y^n z^n = y^n \Rightarrow z^n = 1$. But G has no π-torsion and so $z = 1$, which implies that $x = y$, as required.

b) It is clear that it suffices to prove this assertion in the case where one of m, n is 1 (then $x^m y^n = y^n x^m \Rightarrow x^m y = y x^m \Rightarrow xy = yx$). Thus, we shall assume without loss of generality, that $x^m y = y x^m$. This implies that $x^m = y^{-1} x^m y = (y^{-1} x y)^m$. It now follows from a) that $x = y^{-1} x y$, that is $xy = yx$, as required.

The theorem is proved.

2.6.2 Theorem. *For any group G and any natural numbers i, j the following holds*

$$[I_\pi(\gamma_i(G)), I_\pi(\gamma_j(G))] \leq I_\pi(\gamma_{i+j}(G)).$$

(Note that the subset $I_\pi(\gamma_i(G))$ is a subgroup for every i since it coincides with the full inverse image of the π-isolator of the identity subgroup of the nilpotent factor-group $G/\gamma_i(G)$.)

Proof. We put $H = G/I_\pi(\gamma_{i+j}(G))$. The group H is obviously nilpotent of class $\leq i + j - 1$ and has no π-torsion. Since for $k \leq i + j$ it is evident that $I_\pi(\gamma_k(H)) = I_\pi(\gamma_k(G))/I_\pi(\gamma_{i+j}(G))$, it will be sufficient to prove that

$$[I_\pi(\gamma_i(H)), I_\pi(\gamma_j(H))] = 1.$$

For any $a \in I_\pi(\gamma_i(H))$, $b \in I_\pi(\gamma_j(H))$ we have $a^m \in \gamma_i(H)$ and $b^n \in \gamma_j(H)$ for some π-numbers m and n. Therefore, by 2.1.3 a), $[a^m, b^n] \in \gamma_{i+j}(H) = 1$, that is $a^m b^n = b^n a^m$. But H is π-torsion-free, and hence we have $ab = ba$ by Theorem 2.6.1.

Hence $[a, b] = 1$ for any $a \in I_\pi(\gamma_i(H))$, $b \in I_\pi(\gamma_j(H))$, as required.
The theorem is proved.

§ 2.7 Basic commutators and the collecting process

Suppose that a group G has a subnormal series

$$G = G_1 \geq G_2 \geq \ldots \geq G_n \geq G_{n+1} = 1$$

with cyclic factors $\langle \bar{c}_i \rangle = G_i/G_{i+1}$. Let us choose inverse images $c_i \in G$ of the elements $\bar{c}_i$. Then every element $g \in G$ may be uniquely represented in the form

$$g = c_1^{i_1} \cdot c_2^{i_2} \cdot \ldots \cdot c_n^{i_n} \tag{2.7.1}$$

where $0 \leq i_s < |\bar{c}_s|$ if $|\bar{c}_s| < \infty$, and $i_s \in \mathbb{Z}$ if $|\bar{c}_s| = \infty$. We prove this by induction on n: for $n = 1$, the group $G = \langle c_1 \rangle$ is cyclic and the assertion is obvious. For $n > 1$, the image $\bar{g}$ of the element g in the cyclic factor-group G/G_2 is uniquely represented in the form $\bar{c}_1^{i_1}$, where $0 \leq i_1 < |\bar{c}_1^{i_1}|$. Then by the induction hypothesis applied to the element $c^{-i_1} g \in G_2$, it is uniquely represented in the form

$$c^{-i_1} g = c_2^{i_2} \cdot c_3^{i_3} \cdot \ldots \cdot c_n^{i_n},$$

which gives the unique representation in the required form (2.7.1).

A group which has a subnormal series with cyclic factors is called *polycyclic*. For finite groups this is equivalent to solubility, but infinite polycyclic groups constitute a proper subclass of the soluble groups. Many of their properties are similar to the properties of nilpotent groups.

Having a central series with finitely generated factors is, of course, sufficient for a group to be polycyclic. Hence, finitely generated nilpotent groups are polycyclic, by Corollary 2.5.6.

The representation of an element of a group in the form (2.7.1) looks like a decomposition of a vector over a basis of a vector space. The elements $c_1, c_2, \ldots, c_n$ are called a *Mal'cev basis* of the group G. The product of two elements of the form (2.7.1) $a = c_1^{i_1} \cdot c_2^{i_2} \cdot \ldots \cdot c_n^{i_n}$, $b = c_1^{j_1} \cdot c_2^{j_2} \cdot \ldots \cdot c_n^{j_n}$, may be also represented in the form (2.7.1)

$$ab = c_1^{f_1(\bar{i}, \bar{j})} \cdot c_2^{f_2(\bar{i}, \bar{j})} \cdot \ldots \cdot c_n^{f_n(\bar{i}, \bar{j})}.$$

Here the exponents $f_k(\bar{i}, \bar{j})$, are, of course, functions of the vectors $\bar{i} = (i_1, i_2, \ldots, i_n)$, $\bar{j} = (j_1, j_2, \ldots, j_n)$. In fact the $f_k(\bar{i}, \bar{j})$ define the complete multiplication table

of the group, that is, they carry full information about its structure. In many cases it is sufficient to know that the $f_k(\bar{i}, \bar{j})$ are polynomials in the "coordinates" $\bar{i}$, $\bar{j}$. The explicit form of these polynomials depends on the particular Mal'cev basis of the group chosen.

In this section we shall construct a Mal'cev basis for free nilpotent groups, consisting of the so-called basic commutators in the free generators. We observe that it is important to be able to do calculations in free nilpotent groups. Firstly, many results may be regarded as facts about the structure of the free nilpotent groups. For example, Kostrikin's Theorem giving a positive solution to the Restricted Burnside Problem for groups of prime exponent p means that, for the free n-generated nilpotent group F of class c, the (finite) index of the subgroup F^p is (n, p)-bounded (and does not depend on c). Secondly, calculation in free nilpotent groups is also a method of proving theorems.

There are many ways of defining basic commutators for free nilpotent groups. The product of two elements of the form (2.7.1) is usually brought to the same form by means of a sequence of transformations, and this procedure is called a collecting process. We shall describe the basic commutators and the collecting process of P. Hall. (For another more general way of defining basic commutators, see the work of Shirshov [132].)

Let F be a free nilpotent group of class c with free generators $x_1, x_2, \ldots, x_n$. The following definition of basic commutators uses induction on their weight and simultaneously defines a linear order on them.

Definition. The elements $x_1, x_2, \ldots, x_n$ are *basic commutators of weight* 1 *in* $x_1, x_2, \ldots, x_n$. Commutators of weight 1 are linearly ordered arbitrarily. If c_1 and c_2 are basic commutators of weight i_1 and i_2 respectively, then the commutator $[c_1, c_2]$ is a *basic commutator of weight* $i_1 + i_2$, if the following two conditions are satisfied:

a) $c_1 > c_2$ and

b) if $c_1 = [c_{11}, c_{12}]$, where c_{11} and c_{12} are basic commutators, then $c_{12} \leq c_2$.

Basic commutators of greater weight are by definition greater (in the linear order) than basic commutators of smaller weight. The linear order on basic commutators of the same weight is defined arbitrarily.

For example, suppose that $F = \langle x_1, x_2, x_3 \rangle$. Put $x_1 < x_2 < x_3$ for weight 1. Then, according to the definition, the basic commutators of weight 2 will be $[x_2, x_1]$, $[x_3, x_1]$, $[x_3, x_2]$. Let us order them, for instance, in the following way: $[x_2, x_1] < [x_3, x_1] < [x_3, x_2]$. Then the basic commutators of weight 3 will be

$$[x_2, x_1, x_1],\ [x_2, x_1, x_2],\ [x_2, x_1, x_3],\ [x_3, x_1, x_1]$$
$$[x_3, x_1, x_2],\ [x_3, x_1, x_3],\ [x_3, x_2, x_2],\ [x_3, x_2, x_3].$$

Next, the basic commutators of weight 4 will be

$$[x_2, x_1, x_1, x_1],\ [x_2, x_1, x_1, x_2], \ldots,$$
$$[[x_3, x_1], [x_2, x_1]],\ [[x_3, x_2], [x_2, x_1]],\ [[x_3, x_2], [x_3, x_1]],$$

here dots denote the remaining simple basic commutators.

The following important theorem will be proved here only in part.

2.7.2 Theorem. *The factors of the lower central series of the free group F with free generators $x_1, x_2, \ldots, x_n$ are free abelian groups: for each $k \in \mathbb{N}$, the factor-group $\gamma_k(G)/\gamma_{k+1}(G)$ is freely generated by the basic commutators of weight k in $x_1, x_2, \ldots, x_n$.*

We shall only prove here that the $\gamma_k(G)/\gamma_{k+1}(G)$ for all k are generated by basic commutators of weight k in $x_1, x_2, \ldots, x_n$ – this will follow from the collecting process described below.

2.7.3 Corollary. *The basic commutators in the free generators $x_1, x_2, \ldots, x_n$ of the free nilpotent group $F/\gamma_{c+1}(F)$ of class c taken in their order, constitute a Mal'cev basis of $F/\gamma_{c+1}(F)$. If $c_1 < c_2 < \ldots < c_N$ are all basic commutators of weight $\leq c$, then each element $g \in F/\gamma_{c+1}(F)$ may be uniquely represented in the form*

$$g = c_1^{i_1} \cdot c_2^{i_2} \cdot \ldots \cdot c_N^{i_N}, \quad i_j \in \mathbb{Z}. \tag{2.7.4}$$

Now we proceed to describe the collecting process. We shall need the formulae gathered in the following lemma.

We use the abbreviation: $[v, {}_s u] = [v, \underbrace{u, u, \ldots, u}_{s}]$.

2.7.5 Lemma. *For any $u, v \in F$ and for all k the following congruences hold modulo $\gamma_k(F)$:*

a) $vu = uv[v, u]$;

b) $vu^{-1} \equiv u^{-1}v[v, {}_2u][v, {}_4u] \ldots [v, {}_3u]^{-1}[v, {}_1u]^{-1} \pmod{\gamma_k(F)}$;

c) $v^{-1}u = u[v, u]^{-1}v^{-1}$;

d) $v^{-1}u^{-1} \equiv u^{-1}[v, {}_1u][v, u_3] \ldots [v, {}_4u]^{-1}[v, {}_2u]^{-1}v^{-1} \pmod{\gamma_k(F)}$.

Since commutators of the form $[v, {}_s u]$ belong to $\gamma_k(G)$ for $s \geq k - 1$, in the formulae b) and d) the dots denote, of course, only a finite number of factors.

Proof. It is easy to verify a) and c) by direct calculation.

Next we have

$$1 = [v, uu^{-1}] = [v, u^{-1}][v, u][v, u, u^{-1}],$$

whence,

$$[v, u^{-1}] = [v, u, u^{-1}]^{-1} \cdot [v, u]^{-1}.$$

Applying this to $[[v, u], u^{-1}]$, where $[v, u]$ takes the role of v, we get on substituting

$$\begin{aligned}[v, u^{-1}] &= [v, u, u^{-1}]^{-1} \cdot [v, u]^{-1} = \\ &= ([v, u, u, u^{-1}]^{-1} \cdot [v, u, u]^{-1})^{-1} \cdot [v, u]^{-1} = \\ &= [v, u, u][v, u, u, u^{-1}][v, u]^{-1}.\end{aligned}$$

We now repeat the process with $[[v, u, u], u^{-1}]$, etc. Commutators of weight $\geq k$ are trivial modulo $\gamma_k(F)$, and so after a finite number of steps this process yields

$$[v, u^{-1}] \equiv [v, {}_2u][v, {}_4u] \ldots [v, {}_3u]^{-1}[v, {}_1u]^{-1} \pmod{\gamma_k(F)}.$$

But $vu^{-1} = u^{-1}v[v, u^{-1}]$, and so b) is proved. It is also easy to verify that $v^{-1}u^{-1} = u^{-1}[v, u^{-1}]^{-1}v^{-1}$, so that the same congruence gives d).

The lemma is proved.

These formulae enable us to transpose basic commutators in such a way that the additional factors emerging are also basic commutators or their inverses. This is guaranteed by the following lemma, which is an easy consequence of the definition of basic commutators.

2.7.6 Lemma. *If v, u, $[v, u]$ are basic commutators in the free generators of a free group F, then the commutators $[v, {}_ku]$ are also basic commutators in the free generators for all k.*

The collecting process involves considering group words, written in the form

$$c_{j_1}^{e_1} \cdot c_{j_2}^{e_2} \cdot \ldots \cdot c_{j_s}^{e_s}, \tag{2.7.7}$$

where $e_i = \pm 1$, and the c_{j_r} are basic commutators. The collected part of the word (2.7.7) is by definition its initial segment (reading from the left) of the form

$$c_1^{i_1} \cdot c_2^{i_2} \cdot \ldots \cdot c_t^{i_t},$$

where we recall that $c_1 < c_2 < \ldots < c_N$ are all basic commutators and, by contrast with (2.7.7), the exponents i_r belong to $\mathbb{Z}$. Every step of the collecting process transforms the word (2.7.7) identically, lengthening its collected part.

In order to represent any element of F modulo $\gamma_{c+1}(F)$ in the form (2.7.4), that is, in the collected form, it is sufficient to apply the collecting process, which we are about to describe, to the representation of this element as a product of the free

generators $x_1, x_2, \ldots, x_n$ of F and their inverses, that is to its representation in the form (2.7.7), where all commutators c_i are of weight 1.

The collecting process consists of consecutive stages, each of which is devoted to the "collecting" of all the occurrences of a given basic commutator c_{i_0}, while it still occurs in the non-collected part of the product (2.7.7) – under the condition that all basic commutators smaller than c_{i_0} are already "collected", that is, do not occur in the non-collected part of (2.7.7). Here we say that the commutator c_{i_0} occurs in (2.7.7) if this product contains $c_{i_0}^e$ as a factor, where $e = \pm 1$.

Thus, at the first stage we collect all occurrences of c_1 (for example, $c_1 = x_1$, if $x_1 < x_2 < \ldots < x_n$). At the start the first occurrence (from the left) of c_1 is transferred to the first place – the subwords $v^e u^f$, where $e, f = \pm 1$, $u = c_1$ and v is a basic commutator greater than c_1, are replaced by the right-hand sides of formulae 2.7.5. Note that each such substitution transfers the first occurrence of c_1 one step to the left, and that all emerging additional factors are also basic commutators. The latter is true by Lemma 2.7.6, because here we always have $v > u$, so that u, v and $[v, u]$ are basic commutators. After several such substitutions the first occurrence of c_1 will be at the beginning of the product (2.7.7) and will constitute its collected part. Next, the first occurrence of c_1 in the non-collected part is transferred in the same way by 2.7.5 to the left, to the collected part, and so on.

Ultimately, there will be no occurrences of c_1 in the non-collected part. The collected part will be a power of c_1, and all additional factors emerging will be basic commutators greater than c_1. The first stage is complete.

We proceed by induction. Suppose that after the first k stages of the collecting process the collected part has the form

$$c_1^{i_1} \cdot c_2^{i_2} \cdot \ldots \cdot c_k^{i_k},$$

and the non-collected part contains only basic commutators v greater than c_k with the property that if $v = [v_1, v_2]$, then $v_2 \leq c_k$.

At stage $k + 1$ all occurrences of c_{k+1} are consecutively transferred to the left, to the end of the collected part. Namely, the subwords $v^e u^f$, where $e, f = \pm 1$, $u = c_{k+1}$ and v is one of the basic commutators which is not only greater than c_k but also greater than c_{k+1}, since we always transfer the first occurrence of c_{k+1} to the left, are replaced by applying 2.7.5. Here, either the weight of v is 1 and then $[v, u]$ is a basic commutator, or $v = [v_1, v_2]$, and by the induction hypothesis $v_2 \leq c_k < c_{k+1}$, so that $[v, u]$ is also a basic commutator by definition. Therefore, all additional factors appearing are basic commutators by Lemma 2.7.6. It is also clear that all of them are greater than c_{k+1}, and if $v = [v_1, v_2]$ is such an additional factor which appeared at stage $k + 1$, then $v_2 \leq c_{k+1}$.

After a finite number of such steps the collected part will be lengthened by a power of c_{k+1}, and in the non-collected part only basic commutators greater than c_{k+1} will be left. Also, if $v = [v_1, v_2]$ is such a commutator, then $v_2 \leq c_{k+1}$ (we

pointed out above that this property held for commutators appearing at stage $k+1$ and it is valid for other commutators by the induction hypothesis). This completes stage $k+1$.

Since in a free nilpotent group of class c there is only a finite number of basic commutators in the free generators which occur in a given element, a group word, it follows that this collecting process will transform that group word to the collected form (2.7.4) in a finite number of steps.

The collecting process is a way of doing calculations in free nilpotent groups and it may be used to carry out calculations in specific cases. However, it may be also used to prove certain theorems. For example, we have just used the collecting process to prove that, for all k, the factor-groups $\gamma_k(G)/\gamma_{k+1}(G)$ are generated by the basic commutators of weight k in the generators of a group G. In the next chapter we shall use this process to prove Zassenhaus' identity

$$[x, \underbrace{y, y, \ldots, y}_{p-1}] \equiv 1 \pmod{\gamma_{p+1}(G)},$$

which is satisfied by every group of prime exponent p. But here we record only one example of such calculations.

Example. *In any group G of exponent 3 the following congruence holds:*

$$[x, y, y] \equiv 1 \pmod{\gamma_4(G)}$$

for arbitrary elements $x, y \in G$.

Proof. It is clearly enough to prove, that we have $[x, y, y] \in F^3$, where x, y are free generators of a free nilpotent group F of class 3. We consider the power $(xy)^3 \in F^3$ and transform it using the collecting process:

$$(xy)^3 = x\,\underset{\leftarrow}{y}\,\underset{-}{x}yxy = x^2\underset{\leftarrow}{y[y,x]}\underset{-}{y}xy =$$

$$= x^3 y\underset{\leftarrow}{[y,x]}\underset{\cdots\cdots}{[y,x,x]}\underset{-}{y}[y,x]y =$$

$$= x^3y^2\underset{\leftarrow}{[y,x]}\underset{\cdots\cdots}{[y,x,y]}\underset{\leftarrow}{[y,x]}\underset{\cdots\cdots}{[y,x,y]}\underset{\leftarrow}{[y,x,x]}\underset{-}{[y,x]}y =$$

$$= x^3y^3[y,x][y,x,y][y,x,y][y,x][y,x,y]\cdot$$

$$\cdot[y,x,y][y,x,x][y,x][y,x,y] =$$

$$= x^3y^3[y,x]^3[y,x,y]^5[y,x,x] \equiv [y,x,y]^2[y,x,x] \pmod{F^3}.$$

Here the element being transferred at any given step is underlined in bold. The dotted underlining means that the element in question commutes with the transferred element and does not give rise to additional factors.

Since the left-hand side of the resultant congruence lies in F^3, we have $[y, x, y]^2[y, x, x] \in F^3$. Since x and y are free generators and F^3 is a verbal subgroup, this inclusion will hold also after replacing x by x^2 (that is, we apply the homomorphism extending the mapping $x \to x^2$, $y \to y$). As a result we see that

$$F^3 \ni [y, x^2, y]^2 \cdot [y, x^2, x^2] = [y, x, y]^4 \cdot [y, x, x]^4 \equiv$$
$$\equiv [y, x, y][y, x, x] \pmod{F^3}.$$

Multiplying the elements $[y, x, y]^2[y, x, x]$ and $[y, x, y][y, x, x]$ from F^3, we obtain $[y, x, x]^2 \in F^3$, whence, $[y, x, x] \in F^3$, as required.

(In fact it may even be proved that any group of exponent 3 satisfies $[y, x, x] = 1$ and $[x_1, x_2, x_3, x_4] = 1$.)

Concluding this section we extend Higman's Lemma to free nilpotent groups. The statements and proofs of all results analogous to those in § 1.10, may be obtained by simply reproducing the latter, replacing free groups by free nilpotent groups of some class c. Note that, in view of the collecting process formulae, one need only consider commutators in the free generators x_i (and not in x_i and x_i^{-1}) in the statements and proofs of such theorems.

§ 2.8 Finite p-groups

In this section p will always denote a prime number.

Definition. A group P is called a *finite p-group* if $|P| = p^n$ for some n.

It follows immediately from Lagrange's Theorem that the order of any element of a finite p-group is a power of p.

The proof of the following important theorem is also based on the arithmetical argument.

2.8.1 Theorem. *Every finite p-group is nilpotent.*

Proof. By induction on the order of the group it is obviously sufficient to show that every finite p-group P has a non-trivial centre. The cardinality of the conjugacy class of elements containing a given element $g \in P$ is equal to the index of its centralizer, $|P : C_P(g)|$ (because the centralizer $C_P(g)$ is the stabilizer of the

point g and the conjugacy class containing g is an orbit of P acting on itself by conjugation). Clearly, this number is either divisible by p, or equal to 1. In the latter case, evidently, $g \in Z(P)$. But the group P of order p^n is the union of pairwise non-intersecting conjugacy classes. Since at least one of them has cardinality 1, that consisting of the identity element alone, it follows by an obvious divisibility argument that there must be at least $p-1$ more one-element classes, whose elements therefore belong to $Z(P)$.

The theorem is proved.

We observe that Corollary 2.5.9 and Theorem 2.8.1 provide a criterion for a finite group to be nilpotent.

2.8.2 Corollary. *A finite group is nilpotent if and only if it is the direct product of its Sylow subgroups.*

We proceed to give some elementary properties of finite p-groups.

2.8.3 Corollary. *In an arbitrary finite p-group*

a) *every maximal subgroup is normal and has index p;*

b) *every normal subgroup of order p is contained in the centre of the group.*

Proof. a) If M is a maximal subgroup of a finite p-group P, then $N_P(M) > M$ by Theorem 2.2.3 c), since P is nilpotent; therefore $N_P(M) = P$, that is $M \trianglelefteq P$. Now if $\bar{z}$ is an element of order p from $Z(P/M)$ and $z \in P$ is its inverse image, then $\langle z, M\rangle$ clearly has order $p|M|$ and coincides with P, since M is maximal.

b) If N is a normal subgroup of order p of a finite p-group P, then $N \cap Z(P) \neq 1$ by Theorem 2.2.3 b), since P is nilpotent. Since $|N \cap Z(P)|$, which is not 1, divides the prime number $|N| = p$ by Lagrange's Theorem, we have $|N \cap Z(P)| = p = |N|$, that is $N \cap Z(P) = N$.

The theorem is proved.

Definition. The subgroup $\Phi(P) = P'P^p$ of a finite p-group P is called the *Frattini subgroup* of the group P.

2.8.4 Lemma. *Let P be a finite p-group. Then*

a) *the Frattini subgroup $\Phi(P)$ is the intersection of all maximal subgroups of P;*

b) *if $\langle \Phi(P), M\rangle = P$, for some subset M, then $\langle M\rangle = P$.*

Proof. Let us denote by I the intersection of all maximal subgroups of P. For any maximal subgroup M of P we have $M \trianglelefteq P$ and $|P/M| = p$ by Corollary 2.8.3. In particular, P/M is abelian and has exponent p, whence $M \supseteq P'$ and $M \supseteq P^p$. Therefore $I \supseteq \Phi(P)$.

To prove the reverse inclusion we consider $P/\Phi(P)$. This is an elementary abelian group of exponent p and may be regarded as a vector space over the field $GF(p)$ (see §1.1). It is easy to see that the intersection of all subspaces of codimension 1 is $\{\bar{0}\}$. (Indeed, by the replacement theorem each non-zero vector $\bar{a}$ may be included in some basis $\{\bar{a}_1 = \bar{a}, \bar{a}_2, \ldots, \bar{a}_n\}$ of this vector space. The subspace spanned by $\{\bar{a}_2, \ldots, \bar{a}_n\}$ has codimension 1 and does not contain $\bar{a}$.) Therefore, the intersection of all maximal subgroups of the factor-group $P/\Phi(P)$ is 1. The full inverse images of the maximal subgroups of $P/\Phi(P)$ are clearly maximal subgroups of P. So their intersection, which contains I, is contained in $\Phi(P)$.

b) Let us assume that $\langle M\rangle \neq P$. Then the proper subgroup $\langle M\rangle$ is contained in some maximal subgroup H. By a) we have $H \supseteq \Phi(P)$, whence

$$P = \langle M, \Phi(P)\rangle = \langle M\rangle\Phi(P) \subseteq H\Phi(P) = H,$$

a contradiction.

The lemma is proved.

A *minimal system of generators* of a group is a set of generators such that any proper subset generates a strictly smaller subgroup.

2.8.5 The Burnside Basis Theorem. *A set of elements of a finite p-group P is a minimal system of generators for it if and only if the images of these elements in the factor-group $P/\Phi(P)$, viewed as a vector space over $GF(p)$, form a basis.*

Proof. Let $a_1, a_2, \ldots, a_n$ be the inverse images of some basis $\{\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n\}$ of $P/\Phi(P)$. It is then clear that $\langle a_1, a_2, \ldots, a_n, \Phi(P)\rangle = P$, whence by Lemma 2.8.4 b) we have $\langle a_1, a_2, \ldots, a_n\rangle = P$. No one of the elements a_i can be deleted in the left-hand side of this equation, because the images of the remaining elements will not generate $P/\Phi(P)$.

Now suppose that $a_1, a_2, \ldots, a_n$ constitute a minimal system of generators of P. Then their images $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n$ obviously generate $P/\Phi(P)$. If they were linearly dependent in the vector space $P/\Phi(P)$, then some proper subset of them – elements $\bar{a}_{i_1}, \bar{a}_{i_2}, \ldots, \bar{a}_{i_s}$, $s < n$, say, – would form a basis. Then, by what was proved above, $a_{i_1}, a_{i_2}, \ldots, a_{i_s}$ would generate P. But this contradicts the minimality of the system of generators $a_1, a_2, \ldots, a_n$.

The theorem is proved.

We remark that it is an immediate consequence of this theorem that the cardinality of a minimal system of generators of a finite p-group is an invariant of the group.

2.8.6 Theorem. *Suppose that P is a finite p-group.*

a) *Every normal subgroup $N \trianglelefteq P$ may be included into some central series of P with factors of order p.*

b) *Every subgroup $H \leq P$ may be included into some subnormal series of P with factors of order p.*

Proof. a) Induction on the order of P. Since P is nilpotent, we have $Z(P) \cap N \neq 1$ by 2.2.3 b). Let z be an element of order p from $Z(P) \cap N$. Then by the induction hypothesis $P/\langle z \rangle$ has a central series containing $N/\langle z \rangle$, with factors of order p. The complete inverse images of the terms of this series together with the subgroup $\langle z \rangle$ will form the desired central series of P containing N.

b) Since P is nilpotent, we have $N_G(H) > H$ for every proper subgroup H by 2.2.3 c). It is therefore easy to construct a subnormal series of P containing H by starting with H, then taking its normalizer, then the normalizer of the normalizer, and so on. The factors of this series are finite p-groups and by a) they have central series with factors of order p. The complete inverse images of the terms of these series taken together form the desired subnormal series of P containing H.

The theorem is proved.

Now we move on from elementary facts about finite p-groups to more advanced results. The first of them is remarkable in that a rather simple condition yields an unexpectedly strong conclusion.

2.8.7 Theorem (P. Hall [28]). *Let N be a normal subgroup of a finite p-group P which is contained in $\gamma_s(P)$. Then all factors of the lower central series of N, with the possible exception of the last, have order $\geq p^s$ (that is $|\gamma_i(N)/\gamma_{i+1}(N)| \geq p^s$, whenever $\gamma_{i+1}(N) \neq 1$).*

Proof. We first prove a generalization of the Three Subgroups Lemma.

2.8.8 Lemma. *For normal subgroups A and B of an arbitrary group and for any k the following inclusion holds:*

$$[A, \gamma_k(B)] \leq [A, \underbrace{B, B, \ldots, B}_{k}].$$

(Note that a more general assertion is also true: for any k and normal subgroups A, $B_1, B_2, \ldots, B_k$ of an arbitrary group the following holds:

$$[A, [B_1, B_2, \ldots, B_k]] \leq \prod_{\pi \in \mathbb{S}_k} [A, B_{\pi(1)}, B_{\pi(2)}, \ldots, B_{\pi(k)}].)$$

Proof. Induction on k. For $k = 1$ the assertion is trivial. For $k > 1$, by the induction hypothesis we have

$$[A, \gamma_{k-1}(B), B] \leq [A, \underbrace{B, B, \ldots, B}_{k-1}, B] = [A, \underbrace{B, B, \ldots, B}_{k}].$$

Also by the induction hypothesis, applied to the subgroups $[A, B]$ and B instead of A and B, we have

$$[[B, A], \gamma_{k-1}(B)] \leq [[B, A], \underbrace{B, B, \ldots, B}_{k-1}] = [A, \underbrace{B, B, \ldots, B}_{k}].$$

Hence, by the Three Subgroups Lemma 2.1.2, we also have

$$[A, \gamma_k(B)] = [\gamma_{k-1}(B), B, A] \leq [A, \underbrace{B, B, \ldots, B}_{k}].$$

The lemma is proved.

We return to the proof of the theorem. Suppose that $\gamma_{i+1}(N) \neq 1$, that is, $[\gamma_i(N), N] \neq 1$. By Theorem 2.8.6 a) there exists a normal subgroup $M \trianglelefteq P$, such that $\gamma_{i+1}(N) > M$, $|\gamma_{i+1}(N) : M| = p$ and $[\gamma_{i+1}(N), P] \leq M$. Clearly, therefore $[\gamma_i(N), N] \not\leq M$. By hypothesis $N \leq \gamma_s(P)$, and so $[\gamma_i(N), \gamma_s(P)] \not\leq M$. By Lemma 2.8.8 we have $[\gamma_i(N), \gamma_s(P)] \leq [\gamma_i(N), \underbrace{P, P, \ldots, P}_{s}]$ and hence

$$[\gamma_i(N), \underbrace{P, P, \ldots, P}_{s}] \not\leq M. \tag{2.8.9}$$

Denoting images mod $\gamma_{i+1}(N)$ by bars, consider the series

$$\overline{\gamma_i(N)} \geq \overline{[\gamma_i(N), P]} \geq \overline{[\gamma_i(N), P, P]} \geq \ldots \geq \overline{[\gamma_i(N), \underbrace{P, P, \ldots, P}_{s-1}]} \geq 1 \tag{2.8.10}$$

for $\bar{P} = P/\gamma_{i+1}(N)$. We have

$$\overline{[\gamma_i(N), \underbrace{P, P, \ldots, P}_{s-1}]} \neq 1,$$

because otherwise

$$[[\gamma_i(N), \underbrace{P, P, \ldots, P}_{s-1}], P] \leq [\gamma_{i+1}(N), P] \leq M,$$

which contradicts (2.8.9). Therefore all inclusions in (2.8.10) are strict inclusions, since for $k \leq s-1$ by Theorem 2.2.3 a), we have

$$[\overline{\gamma_i(N)}, \underbrace{\bar{P}, \ldots, \bar{P}}_{k}] = [\overline{\gamma_i(N)}, \underbrace{\bar{P}, \ldots, \bar{P}}_{k-1}, \bar{P}] < [\overline{\gamma_i(N)}, \underbrace{\bar{P}, \ldots, \bar{P}}_{k-1}]$$

for the ***non-trivial*** normal subgroup $[\overline{\gamma_i(N)}, \underbrace{\bar{P}, \ldots, \bar{P}}_{k-1}]$ of $\bar{P}$. Thus, each of the s factors of the series (2.8.10) has order $\geq p$. The product of these orders is the order of the factor-group $\gamma_i(N)/\gamma_{i+1}(N)$, and hence the latter is at least p^s.

The theorem is proved.

Now we shall prove a result which we need to prove the Magnus-Sanov Theorem on the $(p-1)$-Engel condition for the associated Lie ring of a group of prime exponent.

2.8.11 Theorem (Zassenhaus' identity). *For any group G of prime exponent p and for any elements $x, y, y_1, y_2, \ldots, y_{p-1} \in G$ the following congruences hold:*

a) $\prod_{\pi \in \mathfrak{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] \equiv 1 \pmod{\gamma_{p+1}(G)}$;

b) $[x, \underbrace{y, y, \ldots, y}_{p-1}] \equiv 1 \pmod{\gamma_{p+1}(G)}$.

Proof. We recall the abbreviation $[v, \underbrace{u, u, \ldots, u}_{n}] = [v, {}_n u]$.

We show first that a) implies b). Putting $y_1 = y_2 = \ldots = y_{p-1} = y$ in a), we get

$$[x, {}_{p-1}y]^{(p-1)!} \equiv 1 \pmod{\gamma_{p+1}(G)}.$$

Since $(p-1)! \equiv -1 \pmod{p}$, we deduce b) at once since G is a group of exponent p.

It suffices to prove a) for free generators $x, y, y_1, y_2, \ldots, y_{p-1}$ of a free group F of the variety $\mathfrak{N}_p \cap \mathfrak{B}_p$ of nilpotent groups of class p and of exponent p (we shall use the extra generator y in the proof). We apply the collecting process to the left-hand side of the equation

$$\underbrace{(xy)\ldots(xy)}_{p} = (xy)^p = 1.$$

If $y < x$ for basic commutators of weight 1, then we shall obtain a product of basic commutators

$$y^p \cdot x^p \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_{p-1} \cdot b \cdot [x, {}_{p-1}y]^\gamma = 1, \tag{2.8.12}$$

where a_i denotes a product of powers of basic commutators of weight $i = 2, 3, \ldots, p-1$, and b denotes a product of powers of basic commutators of weight p, different from $[x, {}_{p-1}y]$. Note that all basic commutators occurring in b have weight ≥ 2 in x.

We now compute the power γ to which $[x, {}_{p-1}y]$ occurs in (2.8.12). It is clear from the description of the collecting process (see §2.7) that a simple commutator $[x, {}_{s}y]$ appears only when we pass the element y to the left over commutators $[x, {}_{s-1}y]$ which arose from previous steps. Therefore $[x, {}_{p-1}y]$ appears only

a) at the successive passing over the first element x of any $p-1$ of the p elements y, which originally were to the right of this element x, that is, C_p^{p-1} times, and

b) at the successive passing over the second x of all $p-1$ elements y, which were placed to the right of this second element x, that is, 1 more.

All together we have $\gamma = C_p^{p-1} + 1 = p + 1$.

Since F has exponent p, we may put $\gamma = 1$ and drop the factor $y^p \cdot x^p$ in (2.8.12).

We shall also need the set R of all commutators in $x, y, y_1, y_2, \ldots, y_{p-1}$, which involve at least two equal variables.

2.8.13 Lemma. a) *The subgroup $\langle R \rangle$ is normal in F.*

b) *If c is a commutator in $x, y, y_1, y_2, \ldots, y_{p-1}$ of weight i in x, j in y, k_s in y_s, $s = 1, 2, \ldots, p-1$, then the commutator obtained from c by replacing x by x^λ, y by y^μ, y_s by $y_s^{\nu_s}$ $(\lambda, \mu, \nu_s \in \mathbb{N})$ is equal to $c^\varkappa \cdot r$, where*

$$\varkappa = \lambda^i \cdot \mu^j \cdot \nu_1^{k_1} \cdot \nu_2^{k_2} \cdot \ldots \cdot \nu_{p-1}^{k_{p-1}},$$

and r is a product of commutators from R, each of whose weights is greater than that of c.

Proof. This follows from 2.1.1 c), d).

We now replace all occurrences of y in (2.8.12) by the product $y_1 \cdot y_2 \cdot \ldots \cdot y_{p-1}$. By 2.1.1 c), d) one can transform the resulting equation to the form

$$a_2 a_3 \ldots a_{p-1} b \cdot \prod_{\pi \in \mathfrak{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] = 1 \qquad (2.8.14)$$

(after changing notation) where a_i is a product of powers of basic commutators of weight $i = 2, 3, \ldots, p-1$ and b is a product of powers of basic commutators of weight p, from R.

Now let ν be an integer whose image in $\mathbb{Z}/p\mathbb{Z}$ generates the multiplicative group of the field $GF(p) \cong \mathbb{Z}/p\mathbb{Z}$, that is has order $p-1$ in it. We replace

all elements $x, y_1, y_2, \ldots, y_{p-1}$ in (2.8.14) by their ν-th powers and collect the resulting equation. By Lemma 2.8.13 we obtain

$$a_2^{\nu^2}\hat{a}_3 \ldots \hat{a}_{p-1}\hat{b} \cdot \left(\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] \right)^{\nu^p} = 1,$$

where a_2 is the same as in (2.8.14), and the $\hat{a}_i$, $i = 3, \ldots, p-1$ and $\hat{b}$ have the same properties as a_i and b in (2.8.14).

Taking the $(-\nu^2)$-th power of (2.8.14) and collecting the result, we get

$$a_2^{-\nu^2}a_3' \ldots a_{p-1}'b' \cdot \left(\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] \right)^{-\nu^2} = 1,$$

where the element a_2 is as before, and a_i', b' have the same properties as a_i and b. This follows from the fact that the commutators of commutators occurring in (2.8.14) lie in R, since all of them involve at least two occurrences of x. Thus, after applying 2.1.1 and the collecting process, there appear no commutators of weight 1 in each variable $x, y_1, y_2, \ldots, y_{p-1}$, other than those in the brackets associated with the power $-\nu^2$, which arise from the product

$$\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}].$$

Multiplying the resultant equations in order to get rid of a_2, for the same reasons we get

$$a_3'' \ldots a_{p-1}''b'' \cdot \left(\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] \right)^{\nu^p - \nu^2} = 1,$$

where a_i'' and b'' have the same meaning as before. By our choice of ν, we have $\nu^p - \nu^2 = \nu^2(\nu^{p-2} - 1) \not\equiv 0 \pmod{p}$, whence there exists an integer s such that $(\nu^p - \nu^2)s \equiv 1 \pmod{p}$. Taking the s-th power of the resultant equation, collecting the result and using the same argument as before, we get

$$a_3 \ldots a_{p-1}b \cdot \left(\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] \right)^{(\nu^p - \nu^2)s} = 1,$$

(after changing notation), where a_i and b have the same meaning as before. In a group of exponent p we may drop the exponent $(v^p - v^2)s$, congruent to 1 modulo p, so that we get

$$a_3 \ldots a_{p-1} b \cdot \prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] = 1.$$

Proceeding in the same manner we can successively eliminate all commutators of weight $\leq p-1$, the product of the commutators of weight 1 in each variable $x, y_1, y_2, \ldots, y_{p-1}$ being always equal to

$$\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}].$$

Eventually we obtain

$$b_0 \cdot b_1 \cdot b_2 \cdot \ldots \cdot b_{p-1} = 1, \tag{2.8.15}$$

where b_i denotes the product of the powers of those commutators which contain exactly i occurrences of the element x. The multilinear component of the left-hand side, that is the product of commutators of weight 1 in each variable $x, y_1, y_2, \ldots, y_{p-1}$ (which is, of course, a part of b_1) is equal to

$$\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}].$$

We rewrite (2.8.15) additively in the abelian group $\gamma_p(F)$ and replace x in it by x^i for each $i = 0, 1, 2, \ldots, p-1$. The following system of equations is obtained:

$$\sum_{j=0}^{p-1} i^j b_j = 0, \quad i = 0, 1, \ldots, p-1.$$

The matrix of coefficients (i^j) has determinant of Vandermonde type and is non-degenerate modulo p. This implies that

$$b_0 = b_1 = \ldots = b_{p-1} = 0,$$

and, in particular, $b_1 = 0$.

We explain this in more detail. The system of equations under consideration may be represented in matrix form with coefficients from the field $\mathbb{Z}/p\mathbb{Z} \cong GF(p)$

$$(\alpha_{ij}) \cdot \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

where the coefficients $\alpha_{ij} = (i-1)^{j-1}$ are residues modulo p. The Vandermonde determinant of the matrix (α_{ij})

$$\det(\alpha_{ij}) = \prod_{1 \le i_2 < i_1 \le p} (i_1 - 1) - (i_2 - 1)) = \prod_{1 \le i_2 < i_1 \le p} (i_1 - i_2)$$

is not 0, as $i_1 - i_2 \neq 0$ for all factors. Therefore the inverse matrix $(\alpha_{ij})^{-1}$ exists. Multiplying by it on the left and using the associativity of matrix multiplication we get

$$\begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{pmatrix} = ((\alpha_{ij})^{-1} \cdot (\alpha_{ij})) \cdot \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{pmatrix} = (\alpha_{ij})^{-1} \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

(Another form of the same argument goes like this: the non-degeneracy of the matrix (α_{ij}) implies that there is a solution $(a_0, a_1, \ldots, a_{p-1})$ of the system of linear equations

$$(x_0, x_1, \ldots, x_{p-1}) \cdot (\alpha_{ij}) = (0, 1, 0, \ldots, 0).$$

Then, taking a linear combination of the equations $\sum_{j=0}^{p-1} i^j b_j = 0$ with coefficients a_i gives

$$0 = \sum_{i=0}^{p-1} a_i \cdot \sum_{j=0}^{p-1} i^j b_j = \sum_{j=0}^{p-1} \left(\sum_{i=0}^{p-1} a_i \cdot i^j \right) b_j = b_1.)$$

Hence, $b_1 = 0$. Now we rewrite this equation in the form

$$b_{1,0} + b_{1,1} + b_{1,2} + \ldots + b_{1,p-1} = 0,$$

where $b_{1,i}$ denotes a linear combination of commutators of weight i in y_1 (and, of course, of weight 1 in x). Exactly the same arguments with respect to the variable

y_1 (replacing y_1 by y_1^i, $i = 0, 1, \ldots, p-1$, etc.) yield

$$b_{1,1} = 0,$$

where the multilinear component of the left-hand side is

$$\sum_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}].$$

Repeating these arguments successively with respect to all variables $x, y_1, y_2, \ldots, y_{p-1}$, we obtain

$$\sum_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] = 0,$$

as required.

The theorem is proved.

Chapter 3

Associated Lie Rings

Witt's identity

$$[a, b^{-1}, c]^b \cdot [b, c^{-1}, a]^c \cdot [c, a^{-1}, b]^a = 1,$$

which holds in every group, looks similar to the Jacobi identity

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

which holds in Lie rings. Apart from the multiplicative notation, the difference lies in the fact that in Witt's identity there are additional conjugating factors, or, using the formula $x^y = x \cdot [x, y]$, additional factors which are commutators of greater weight. One can omit these additional factors and thus define a Lie ring using the group operations, but, of course, a lot of information about the group will be lost. In fact, it is only for nilpotent (or residually nilpotent) groups that such a transition to Lie rings makes sense.

Here we give not only the standard definition of the associated Lie ring, based on the lower central series, but also an analogous construction, based on arbitrary, so-called, strongly central series. This construction will be used in Chapter 5.

We prove the nilpotency of soluble groups of prime exponent, as an illustration of the advantages of using Lie rings which are more linear objects than groups.

We also prove the Magnus-Sanov Theorem that the associated Lie ring of a group of prime exponent p is $(p-1)$-Engel. This result reduces the Restricted Burnside Problem for groups of prime exponent to Kostrikin's Theorem 1.3.1 on Engel Lie algebras. The proof given here may well not be the shortest possible, but it may help us see the kind of difficulty which may be encountered in passing from groups to Lie rings and back. We shall also imitate this proof later at an analogous point in the positive solution of the Restricted Burnside Problem for groups with a splitting automorphism of prime order (Chapter 7).

One of the important advantages of studying a more linear object – a Lie ring – is the fact that one can extend the ground ring and decompose the Lie ring into a sum of the analogues of eigenspaces which arise under the action of an automorphism. We give an account of this technique at the beginning of the next chapter which deals with automorphisms of Lie rings.

§ 3.1 Results on Lie rings analogous to theorems about groups

Here we state a few theorems about Lie rings whose proofs may be easily obtained by translating the proofs of the analogous theorems about groups into the language of Lie rings.

The terms of the lower central series $\gamma_i(L)$ of a Lie ring L were defined in § 1.3. Nilpotent Lie rings are defined analogously to nilpotent groups using the following theorem.

3.1.1 Theorem. *The following conditions are equivalent for a Lie ring L:*

a) $\gamma_{c+1}(L) = 0$;

b) *L has a central series of length c*

$$L = L_1 \geq L_2 \geq \ldots \geq L_c \geq L_{c+1} = 0,$$

that is, such that $[L_i, L] \leq L_{i+1}$ for all $i = 1, 2, \ldots, c$;

c) *the Lie ring L satisfies the identity*

$$[x_1, x_2, \ldots, x_{c+1}] = 0.$$

Definition. A Lie ring L, satisfying the conditions of Theorem 3.1.1, is said to be *nilpotent*, and the least natural number c for which they hold, is called the *nilpotency class* of L (sometimes another term – "index of nilpotency" – is used).

One often says that a Lie ring is nilpotent of class c meaning that it is nilpotent of class $\leq c$.

We note that, in order to prove that a Lie ring is nilpotent, it is sufficient to verify the nilpotency identity for its generators.

3.1.2 Theorem. *If a Lie ring is generated by a subset M, then it is nilpotent of class $\leq c$ if and only if every commutator of weight $c+1$ in elements of M is equal to 0.*

The proof of this theorem follows easily from a description of the terms of the lower central series, which we also give here for completeness' sake.

3.1.3 Proposition. *Suppose that L is an arbitrary Lie ring and that k is a positive integer. Then*

a) *the ideal $\gamma_k(L)$ contains all commutators of weight $\geq k$ in elements of L;*

b) *the additive subgroup $\gamma_k(L)$ is generated by the simple commutators of weight k in elements of L;*

c) *if* $L = \langle M \rangle$, *then the additive subgroup* $\gamma_k(L)$ *is generated by the simple commutators of weight* $\geq k$ *in elements of* M, *it is also generated by the basic commutators of weight* $\geq k$ *in elements of* M; *in particular, if* L *is finitely generated (with* s *generators), then for each* $k \in \mathbb{N}$, *the additive group* $\gamma_k(L)/\gamma_{k+1}(L)$ *is also finitely generated (with a* (k, s)*-bounded number of generators).*

The definition of soluble Lie rings is also completely analogous to the definition of soluble groups.

3.1.4 Theorem. *The following conditions are equivalent for a Lie ring* L:

a) $L^{(s)} = 0$;

b) L *has a series of ideals of length* s *with commutative factors*

$$L = L_0 \trianglerighteq L_1 \trianglerighteq \ldots \trianglerighteq L_{s-1} \trianglerighteq L_s = 0,$$

that is, such that $[L_i, L_i] \leq L_{i+1}$ *for all* $i = 0, 1, \ldots, s-1$.

c) L *satisfies the identity*

$$\delta_s(x_1, x_2, \ldots, x_{2^s}) = 0.$$

(For the definition of δ_s see the Preface, and for the definition of the terms of the derived series $L^{(s)}$ see § 1.3.)

Definition. A Lie ring, satisfying the conditions of Theorem 3.1.4, is said to be *soluble*, and the least natural number s for which they hold, is called the *derived length* of L (sometimes the term – "index of solubility" – is used).

It is often said that a Lie ring is soluble of derived length s, meaning that it is soluble of derived length $\leq s$.

Remark. The analogue of Theorem 3.1.2 does not hold for solubility: there exist examples of Lie rings $G = \langle M \rangle$ with $\delta_k(m_1, m_2, \ldots, m_{2^k}) = 1$ for any $m_1, \ldots, m_{2^k} \in M$, which, however, are not soluble of derived length k.

The following two theorems are Lie ring analogues of Theorems 2.3.5 and 2.4.5, and their proofs may be easily obtained by translating the group theoretic commutator calculations into the language of Lie rings.

3.1.5 Theorem. *If all soluble Lie rings of derived length* 2 *in some variety of Lie rings are nilpotent of class* $\leq c$, *then any soluble Lie ring of derived length* s *in this variety is nilpotent of class* $\leq \frac{c^s-1}{c-1}$.

3.1.6 Theorem. *If the ideal* $\gamma_k(L)$ *in an arbitrary Lie ring* L *is finite and has order* n *(or, for Lie algebra* L, *is of finite dimension* n*), then the Lie ring (Lie algebra)*

L contains a nilpotent subring of class k of finite (k, n)-bounded index (of finite (k, n)-bounded codimension).

(Here the index of a subring means its index as a subgroup of the additive group of the ring.)

§ 3.2 Constructing a Lie ring from a group

We first give the definition of the associated Lie ring which is defined in terms of the lower central series of a group. This construction is then generalized to the case of arbitrary so-called strongly central series.

Let G be a group. We shall put $\gamma_k = \gamma_k(G)$, $k \in \mathbb{N}$, in those cases where there is only one ambient group G so that no danger of confusion arises.

3.2.1 Definition. The additive group of the *associated Lie ring* $L(G)$ of a group G is the direct sum

$$L(G) = \bigoplus_{k=1}^{\infty} \gamma_k/\gamma_{k+1},$$

writing the abelian groups γ_k/γ_{k+1} additively.

For each $k \in \mathbb{N}$ the direct summand γ_k/γ_{k+1} is called the *homogeneous component of $L(G)$ of weight k*. Multiplication in $L(G)$ is defined for the elements of the homogeneous components by

$$[a + \gamma_{i+1}, b + \gamma_{j+1}] = [a, b] + \gamma_{i+j+1},$$

where $a + \gamma_{i+1}$ and $b + \gamma_{j+1}$ are the images of the elements $a \in \gamma_i$ and $b \in \gamma_j$ in factor-groups γ_i/γ_{i+1} and γ_j/γ_{j+1}, respectively, and $[a, b] + \gamma_{i+j+1}$ is the image of the group commutator $[a, b]$ in the factor-group $\gamma_{i+j}/\gamma_{i+j+1}$. Multiplication is then extended to $L(G)$ by linearity.

(Note that it may very well happen that $[a + \gamma_{i+1},\ b + \gamma_{j+1}] = 0$ in $L(G)$, although $[a, b] \neq 1$ in G, – we only need $[a, b] \in \gamma_{i+j+1}$.)

Definition. The elements of the homogeneous component γ_k/γ_{k+1} of weight k of the Lie ring $L(G)$ are called *homogeneous elements of $L(G)$ of weight k.*

3.2.2 Theorem. *Definition 3.2.1 defines a Lie ring structure on $L(G)$.*

If G is nilpotent of class c, then $L(G)$ is also nilpotent of exactly the same class c, and if G is also finite, then $|L(G)| = |G|$. If G is soluble of derived length s, then $L(G)$ is also soluble of derived length $\leq s$.

Every automorphism φ of the group G induces an automorphism of the Lie ring $L(G)$ by its action on the factor-groups γ_i/γ_{i+1} and if G is finite and the order of φ is coprime to the order of G, then φ faithfully acts on $L(G)$ and $|C_{L(G)}(\varphi)| = |C_G(\varphi)|$.

(The automorphism induced on $L(G)$ by φ is usually also denoted by φ.)

Proof. The correctness of the definition of multiplication of homogeneous elements in $L(G)$ follows from 2.1.3 a) and 2.1.1: if $a' + \gamma_{i+1} = a + \gamma_{i+1} \in \gamma_i/\gamma_{i+1}$ and $b' + \gamma_{j+1} = b + \gamma_{j+1} \in \gamma_j/\gamma_{j+1}$, then $a' = ag_1$, where $g_1 \in \gamma_{i+1}$, and $b' = bg_2$, where $g_2 \in \gamma_{j+1}$, and then

$$[a', b'] = [ag_1, bg_2] \equiv [a, b] \pmod{\gamma_{i+j+1}},$$

since $[\gamma_i, \gamma_{j+1}] \cdot [\gamma_{i+1}, \gamma_j] \leq \gamma_{i+j+1}$.

Anticommutativity and the Jacobi identity are multilinear and it therefore suffices to verify them for homogeneous elements. Let $\bar{a} = a + \gamma_{i+1} \in \gamma_i/\gamma_{i+1}$, $\bar{b} = b + \gamma_{j+1} \in \gamma_j/\gamma_{j+1}$ and $\bar{c} = c + \gamma_{k+1} \in \gamma_k/\gamma_{k+1}$. Then

$$[\bar{a}, \bar{b}] = [a, b] + \gamma_{i+j+1} = -[b, a] + \gamma_{i+j+1} = -[\bar{b}, \bar{a}],$$

since $[a, b] = [b, a]^{-1}$ for group commutators. Also we have

$$[\bar{a}, \bar{b}, \bar{c}] + [\bar{b}, \bar{c}, \bar{a}] + [\bar{c}, \bar{a}, \bar{b}] = [a, b, c] + [b, c, a] + [c, a, b] + \gamma_{i+j+k+1} = 0,$$

because of Witt's identity 2.1.1 e) for group commutators

$$[a, b^{-1}, c]^b \cdot [b, c^{-1}, a]^c \cdot [c, a^{-1}, b]^a = 1$$

and

$$\begin{aligned}[a, b^{-1}, c]^b = [a, b^{-1}, c] \cdot [[a, b^{-1}, c], b] \equiv [a, b^{-1}, c] \equiv \\ \equiv [[a, b]^{-1}, c] \equiv [a, b, c]^{-1} \pmod{\gamma_{i+j+k+1}}\end{aligned}$$

by 2.1.1 and 2.1.3 a) (analogously for $[b, c^{-1}, a]^c$ and $[c, a^{-1}, b]^a$).

The fact that the nilpotency classes of G and $L(G)$ are the same for G nilpotent, is established by the following lemma.

3.2.3 Lemma. *In the associated Lie ring of an arbitrary group G, the following hold for any k:*

a) $\gamma_k/\gamma_{k+1} = {}_+\langle [a_1, a_2, \ldots, a_k] \mid a_i \in \gamma_1/\gamma_2 \rangle$;

b) $\gamma_k(L(G)) = \bigoplus_{i=k}^{\infty} \gamma_i/\gamma_{i+1}$.

(Here the right-hand side of a) is the additive subgroup, generated by the Lie ring commutators.)

Proof. By Proposition 2.1.5 the factor-group γ_k/γ_{k+1} is generated by the images of the commutators $[g_1, g_2, \ldots, g_k]$ of weight k in elements $g_i \in G$. Using the definition of multiplication in the associated Lie ring, one can easily prove that the image of such a commutator of weight k is equal to the Lie ring commutator $[\bar{g}_1, \bar{g}_2, \ldots, \bar{g}_k]$, where $\bar{g}_i$ is the image of g_i in γ_1/γ_2. So, passing to additive notation, we get $\gamma_k/\gamma_{k+1} = {}_+\langle [\bar{g}_1, \bar{g}_2, \ldots, \bar{g}_k] \mid\mid \bar{g}_i \in \gamma_1/\gamma_2 \rangle$, which proves assertion a).

Since $[\bar{g}_1, \bar{g}_2, \ldots, \bar{g}_k] \in \gamma_k(L(G))$, it follows from a) that $\gamma_k(L(G)) \supseteq \bigoplus_{i=k}^{\infty} \gamma_i/\gamma_{i+1}$; the reverse inclusion follows from the fact that $[\gamma_m, \gamma_n] \leq \gamma_{m+n}$.

The lemma is proved.

The fact that $|G| = |L(G)|$, if G is a finite nilpotent group, follows from the equations

$$|G| = \prod_{i=1}^{c} |\gamma_i/\gamma_{i+1}| \quad \text{and} \quad |L(G)| = \prod_{i=1}^{c} |\gamma_i/\gamma_{i+1}|,$$

where c is the nilpotency class.

If the group G satisfies the solubility identity

$$\delta_s(x_1, x_2, \ldots, x_{2^s}) = 1,$$

then, by the definition of multiplication in the associated Lie ring, the same identity is satisfied by the homogeneous elements of $L(G)$, and, since this identity is multilinear, it is satisfied by $L(G)$ itself.

If $\varphi \in \operatorname{Aut} G$, then for homogeneous elements $\bar{a} = a + \gamma_{i+1} \in \gamma_i/\gamma_{i+1}$ and $\bar{b} = b + \gamma_{j+1} \in \gamma_j/\gamma_{j+1}$, we have

$$[\bar{a}, \bar{b}]^{\varphi} = [a, b]^{\varphi} + \gamma_{i+j+1} = [a^{\varphi}, b^{\varphi}] + \gamma_{i+j+1} = [a^{\varphi} + \gamma_{i+1}, b^{\varphi} + \gamma_{j+1}] = [\bar{a}^{\varphi}, \bar{b}^{\varphi}].$$

By definition we extend the action of the induced automorphism φ to $L(G)$ from the abelian groups γ_i/γ_{i+1} by linearity, and so, for any $l_1 = \sum_s l_{1s}$, $l_2 = \sum_s l_{2s}$, where $l_{is} \in \gamma_s/\gamma_{s+1}$, $i = 1, 2$, $s \in \mathbb{N}$, we have

$$[l_1, l_2]^{\varphi} = \left[\sum_s l_{1s}, \sum_r l_{2r}\right]^{\varphi} = \left(\sum_{s,r} [l_{1s}, l_{2r}]\right)^{\varphi} =$$

$$= \sum_{s,r} [l_{1s}, l_{2r}]^{\varphi} = \sum_{s,r} [l_{1s}^{\varphi}, l_{2r}^{\varphi}] = \left[\sum_s l_{1s}^{\varphi}, \sum_r l_{2r}^{\varphi}\right] = [l_1^{\varphi}, l_2^{\varphi}].$$

This means that φ really is an automorphism of the Lie ring $L(G)$.

In the case where G is a finite nilpotent group, the faithfulness of an induced automorphism of order coprime to the order of G follows from Corollary 1.6.3, since an automorphism acts trivially on $L(G)$ if and only if it centralizes all factors of the lower central series of G. With the same hypothesis it follows from Theorem 1.6.2 that $|C_{L(G)}(\varphi)| = |C_G(\varphi)|$.

The theorem is proved.

Remark. By contrast with nilpotency class, the derived length of the associated Lie ring $L(G)$ may be smaller than the derived length of the group G, even if the group G is nilpotent.

Lemma 3.2.3 a) has a useful corollary.

3.2.4 Corollary. *For any group G the associated Lie ring $L(G)$ is generated by its homogeneous component γ_1/γ_2 of weight 1, and if the group G is generated by a set M, then the Lie ring $L(G)$ is generated by the images of the elements of M in the factor-group γ_1/γ_2.*

Although the associated Lie ring may be constructed starting with an arbitrary group, it is clear that it reflects only the properties of the factor-group $G/\bigcap\limits_{i=1}^{\infty}\gamma_i(G)$, since, clearly,

$$L(G) \cong L\left(G/\bigcap_{i=1}^{\infty}\gamma_i(G)\right).$$

For example, $L(B(m, p)) \cong L(\bar{B}(m, p))$, although the free m-generated Burnside group $B(m, p)$ of prime exponent p is infinite (and insoluble) for $m \geq 2$ and for any prime number $p \geq 667$ by the Adian-Novikov Theorem [1, 116], and the group $\bar{B}(m, p) = B(m, p)/\bigcap\limits_{i=1}^{\infty}\gamma_i(B(m, p))$ is finite by Kostrikin's Theorem [76].

A Lie ring may also be constructed using other types of central series.

Definition. A series of a group G

$$G = K_1 \geq K_2 \geq \ldots \geq K_c \geq K_{c+1} = 1 \tag{3.2.5}$$

is said to be *strongly central* if $[K_i, K_j] \leq K_{i+j}$ for all $i, j = 1, 2, \ldots, c$.

It is clear that any strongly central series is automatically normal and central, and if a group G has a strongly central series of length c, then it is nilpotent of class $\leq c$. The next theorem is completely analogous to Theorem 3.2.2.

3.2.6 Theorem. *Suppose that the group G has a strongly central series (3.2.5). Then if we replace the subgroups γ_i by K_i everywhere in Definition 3.2.1, we define on the direct sum*

$$L_K(G) = \bigoplus_i K_i/K_{i+1}$$

the structure of a Lie ring which is nilpotent of class not greater than the nilpotency class of G; furthermore, the derived length of this Lie ring is not greater than the derived length of G.

If G is also finite, then $|L_K(G)| = |(G)|$.

If φ is an automorphism of the group G such that all subgroups K_i are φ-invariant, then its induced action on the factor-groups K_i/K_{i+1} defines an automorphism of the Lie ring $L_K(G)$. If, further, the order of φ as a group automorphism is coprime to the order of G then φ acts faithfully on $L_K(G)$ and $|C_{L_K(G)}(\varphi)| = |C_G(\varphi)|$.

Proof. The fact that $L_K(G)$ is given a Lie ring structure by the definition is verified by simply repeating the first part of the proof of Theorem 3.2.2, with the letter γ replaced by K.

It follows from the definition of multiplication in $L_K(G)$, that if some commutator identity is satisfied by G, then this identity is also satisfied by the homogeneous elements of $L_K(G)$. As the nilpotency and solubility identities are multilinear, this implies, that the nilpotency class and the derived length of $L_K(G)$ are not greater than the nilpotency class and the derived length of G, respectively.

The rest of the proof is dealt with precisely as in Theorem 3.2.2.

The theorem is proved.

We now give some examples. If G is a nilpotent group and π any set of primes, then, according to Theorem 2.6.2 the subgroups $I_\pi(\gamma_i(G))$ form a strongly central series of $G/I_\pi(1)$, and this gives rise to a Lie ring via Theorem 3.2.6. Note that in general we cannot speak of a strongly central series of G itself, since $I_\pi(1)$ is not necessarily central. But if $I_\pi(1) = 1$, that is, G has no π-torsion, then the $I_\pi(\gamma_i(G))$ form a strongly central series of the group G.

Another important example is the so-called lower central p-series $\{\lambda_i(G)\}$ of a finite p-group G which is defined inductively as follows:

$$\lambda_1(G) = G; \quad \lambda_{i+1}(G) = [\lambda_i(G), G] \cdot (\lambda_i(G))^p$$

(this is the most rapidly descending central series whose factors have exponent p).

We remark that unlike the associated Lie ring defined using the lower central series, the Lie ring $L_K(G)$ may have a smaller nilpotency class than the nilpotency class of G. The simplest example is the group $G = D \times C$, where $D = \langle a, b \mid a^4 = b^2 = 1,\ a^b = a^{-1}\rangle$ is a dihedral group, and $C = \langle c\rangle$ is a cyclic group of order 2.

We put $K_1 = G$, $K_2 = \langle a^2 \rangle \times \langle c \rangle = D' \times C$, $K_3 = \langle a^2 \rangle = D'$. It is easy to see that the subgroups K_i form a strongly central series of G, and that the corresponding Lie ring is commutative, since $[K_1, K_1] \leq K_3$ and $K_2, K_3 \leq Z(G)$.

Nevertheless, for a π-torsion-free nilpotent group G, the Lie ring constructed using the strongly central series of π-isolators $I_\pi(\gamma_i(G))$ behaves well from this point of view.

3.2.7 Theorem. *Let G be a π-torsion-free nilpotent group for some set of primes π. Then the nilpotency class of the Lie ring L constructed via Theorem 3.2.6 using the strongly central series of π-isolators $I_\pi(\gamma_i(G))$ is the same as the nilpotency class of G as a group.*

Proof. Let k be the nilpotency class of the Lie ring L. It is sufficient to prove that the nilpotency class of G is not greater than k since the reverse inequality follows from Theorem 3.2.6. Translating Lie ring multiplication into the language of the group G we get

$$[g_1, g_2, \ldots, g_{k+1}] \in I_\pi(\gamma_{k+2}(G))$$

for any $g_1, g_2, \ldots, g_{k+1} \in G$. In other words, for arbitrary $g_1, g_2, \ldots, g_{k+1} \in G$ there is a π-number n such that

$$[g_1, g_2, \ldots, g_{k+1}]^n \in \gamma_{k+2}(G).$$

The simple commutators of weight $k+1$ generate the subgroup $\gamma_{k+1}(G)$ (see 2.1.5 b)), and hence the orders of all elements of the abelian factor-group $\gamma_{k+1}(G)/\gamma_{k+2}(G)$ are finite π-numbers. It follows from Corollary 2.5.5 that $\gamma_{k+1}(G)$ itself has the same property. But G has no π-torsion by hypothesis and so, $\gamma_{k+1}(G) = 1$, as was required.

The theorem is proved.

§ 3.3 The Lie ring of a group of prime exponent

In this section we prove the Magnus-Sanov Theorem on the $(p-1)$-Engel condition for the associated Lie ring of a group of prime exponent p. This theorem reduces the Restricted Burnside Problem for groups of prime exponent p to the analogous problem for $(p-1)$-Engel Lie rings, which was solved positively by Kostrikin (Theorem 1.3.1).

Here the situation is more transparent than it is in the case of groups of composite exponent p^k, since the $(p-1)$-Engel identity is, in fact, multilinear.

3.3.1 Lemma. *The following two conditions are equivalent for any Lie algebra L over a field of characteristic $p > 0$:*

a) $[x, \underbrace{y, y, \ldots, y}_{p-1}] = 0$ *for any* $x, y \in L$;

b) $\sum_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] = 0$ *for any* $x, y_1, y_2, \ldots, y_{p-1} \in L$.

Proof. b) $\Rightarrow$ a). Putting $y_1 = y_2 = \ldots = y_{p-1} = y$ in b), we get $(p-1)! \times$ $\times [x, \underbrace{y, y, \ldots, y}_{p-1}] = 0$, whence $[x, \underbrace{y, y, \ldots, y}_{p-1}] = 0$, as $(p-1)!$ is a non-zero element of the field.

a) $\Rightarrow$ b). On substituting $y = y_1 + y_2 + \ldots + y_{p-1}$ in a) we see that the multihomogeneous component of weight 1 in each variable $x, y_1, y_2, \ldots, y_{p-1}$ of the resultant equation coincides with the left-hand side of b). In order to extract it we now substitute in a) $y = iy_1 + y_2 + \ldots + y_{p-1}$ for $i = 0, 1, 2, \ldots, p-1$. This gives the system of equations

$$\sum_{k=0}^{p-1} i^k M_k = 0, \quad i = 0, 1, 2, \ldots, p-1,$$

where M_k is the sum of all multihomogeneous elements of weight k in y_1. The matrix (i^k) has a Vandermonde determinant which is not equal to 0. Therefore $M_k = 0$ for all $k = 0, 1, 2, \ldots, p-1$. (See the more detailed explanation in the proof of Theorem 2.8.11.)

Thus, in particular, $M_1 = 0$. By replacing y_2 by iy_2 in this equation, using the same argument as above, one can extract the sum of all homogeneous components of weight 1 in y_1 and in y_2, and so on. This process leads to b).

The lemma is proved.

3.3.2 Theorem (Magnus [100], Sanov [126]). *The associated Lie ring $L(P)$ of any group P of prime exponent p satisfies the identities*

$$pa = 0 \quad \text{and} \quad [a, \underbrace{b, b, \ldots, b}_{p-1}] = 0.$$

Proof. Since the factor-groups $\gamma_i(P)/\gamma_{i+1}(P)$ have exponent p, the Lie ring $L(P)$ satisfies the identity $pa = 0$. Thus, the Lie ring $L(P)$ may be regarded as a Lie algebra over $GF(p)$. By Lemma 3.3.1 it suffices to prove that

$$\sum_{\pi \in \mathbb{S}_{p-1}} [a, b_{\pi(1)}, b_{\pi(2)}, \ldots, b_{\pi(p-1)}] = 0 \tag{3.3.3}$$

for any $a, b_1, b_2, \ldots, b_{p-1} \in L(P)$. Since this identity is multilinear, the elements $a, b_1, b_2, \ldots, b_{p-1}$ may be taken to be homogeneous elements of $L(P)$.

It is convenient at this point to exploit the fact that groups of exponent p form a variety in which there are free groups. Namely, let F be the free group with free generators $x, y_1, y_2, \ldots, y_{p-1}$. By Theorem 2.8.11 a), we have

$$\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] \equiv 1 \pmod{F^p \cdot \gamma_{p+1}(F)}. \tag{3.3.4}$$

It is natural to try to apply the homomorphism ϑ of the group F into P which extends the mapping

$$x \to \hat{a}, \quad y_i \to \hat{b}_i \quad i = 1, 2, \ldots, p-1,$$

where $\hat{a}$ and $\hat{b}_i$ are the inverse images of the elements $a \in \gamma_{s_0}(P)/\gamma_{s_0+1}(P)$, $b_i \in \gamma_{s_i}(P)/\gamma_{s_i+1}(P)$, respectively. However, though the image of the left-hand side of (3.3.4) under ϑ is certainly equal to

$$\prod_{\pi \in \mathbb{S}_{p-1}} [\hat{a}, \hat{b}_{\pi(1)}, \hat{b}_{\pi(2)}, \ldots, \hat{b}_{\pi(p-1)}]$$

and resembles the left-hand side of (3.3.3), and the image of the subgroup F^p is 1, the argument is not quite valid. In fact (3.3.3), which we want to prove, is equivalent, in terms of the group P, to the congruence

$$\prod_{\pi \in \mathbb{S}_{p-1}} [\hat{a}, \hat{b}_{\pi(1)}, \hat{b}_{\pi(2)}, \ldots, \hat{b}_{\pi(p-1)}] \equiv 1 \pmod{\gamma_{s_0+s_1+\ldots+s_{p-1}+1}(P)},$$

but the image of $\gamma_{p+1}(F)$ under ϑ is not necessarily contained in $\gamma_{s_0+s_1+\ldots+s_{p-1}+1}(P)$. Note that the original paper of Magnus [100] contained such a gap; it was filled by him a bit later [101], but the same mistake was reproduced in the book [102].

This difficulty is overcome with the help of Higman's Lemma from § 1.10. The congruence (3.3.4) is equivalent to requiring

$$\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] \in F^p \cdot \gamma_{p+1}(F),$$

where we see that the product is contained in the intersection of the normal closures of all the elements $x, y_1, y_2, \ldots, y_{p-1}$. Since the subgroups F^p and $\gamma_{p+1}(F)$ are verbal, Corollary 1.10.6 may be applied to yield

$$\prod_{\pi \in \mathbb{S}_{p-1}} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(p-1)}] \equiv w \pmod{F^p}, \tag{3.3.5}$$

where $w \in \gamma_{p+1}(F) \cap \langle x^F \rangle \cap \langle y_1^F \rangle \cap \langle y_2^F \rangle \cap \ldots \cap \langle y_{p-1}^F \rangle$.

By Lemma 1.10.1 the element w is a product

$$w = c_1 \cdot c_2 \cdot \ldots \cdot c_r \tag{3.3.6}$$

of commutators c_i of weight $\geq p+1$, each depending on all of the elements x, $y_1, y_2, \ldots, y_{p-1}$. Applying ϑ to (3.3.5), we get

$$\prod_{\pi \in \mathbb{S}_{p-1}} [\hat{a}, \hat{b}_{\pi(1)}, \hat{b}_{\pi(2)}, \ldots, \hat{b}_{\pi(p-1)}] = w^{\vartheta},$$

since $\vartheta(F^p) = 1$.

It remains to show that $w^{\vartheta} \in \gamma_{s_0+s_1+\ldots+s_{p-1}+1}(P)$. Indeed, $w^{\vartheta} = c_1^{\vartheta} \cdot c_2^{\vartheta} \cdot \ldots \cdot c_r^{\vartheta}$ and each c_i^{ϑ} is a commutator of weight $\geq p+1$ in the elements $\hat{a}, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_{p-1}$ or their inverses and each c_i^{ϑ} is depending on all of the elements $\hat{a}, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_{p-1}$. Using the fact that $[\gamma_i(P), \gamma_j(P)] \leq \gamma_{i+j}(P)$, it is clear that single occurrences of the elements $\hat{a}, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_{p-1}$ in c_i^{ϑ} contribute the sum $s_0 + s_1 + s_2 + \ldots + s_{p-1}$ to the index of that member of the lower central series of P, which must contain the commutator c_i^{ϑ}. But the weight of c_i^{ϑ} as a commutator in $\hat{a}, \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_{p-1}$ is at least $p+1$ – therefore there must be at least one additional occurrence which increases that index by at least one. As a result we have

$$c_i^{\vartheta} \in \gamma_{s_0+s_1+\ldots+s_{p-1}+1}(P)$$

for all i, and therefore $w^{\vartheta} \in \gamma_{s_0+s_1+\ldots+s_{p-1}+1}(P)$.

The theorem is proved.

§ 3.4 The nilpotency of soluble Lie rings satisfying the Engel condition

3.4.1 Theorem (Higgins [39]). *If a soluble Lie algebra L of derived length s over a field of characteristic p satisfies the n-th Engel condition, that is if*

$$[x, \underbrace{y, y, \ldots, y}_{n}] = 0$$

for all $x, y \in L$ and either $p = 0$, or $n < p$, then L is nilpotent of class $\leq \frac{(n+1)^s - 1}{n}$.

Proof. All Lie algebras of characteristic p satisfying the n-th Engel condition, form a variety of Lie rings. Therefore, by virtue of Theorem 3.1.5, it is sufficient

to prove the theorem in the case where $s = 2$, because the desired bound on the nilpotency class follows from the statement of Theorem 3.1.5, provided that one can prove for $s = 2$ that the nilpotency class is at most $n + 1$. We may therefore assume in what follows that $s = 2$, that is, that L is soluble of derived length 2.

By repeating the proof of Lemma 3.3.1 more or less word for word, we find that L satisfies the identity

$$\sum_{\pi \in \mathbb{S}_n} [x, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(n)}] = 0. \tag{3.4.2}$$

It follows from the Jacobi identity $[[a, b], c] = -[[b, c], a] - [[c, a], b] = [a, [b, c]] + [[a, c], b]$, that a soluble Lie ring of derived length 2 satisfies the identity $[[a, b], c] = [[a, c], b]$ for $a \in \gamma_2(L)$, and hence also satisfies

$$[a, b_{\pi(1)}, b_{\pi(2)}, \ldots, b_{\pi(k)}] = [a, b_1, b_2, \ldots, b_k]$$

for any $a \in \gamma_2(L)$, $k \in \mathbb{N}$ and any permutation $\pi \in \mathbb{S}_k$.

Thus on substituting $x = [x_1, x_2]$ in (3.4.2), we obtain the identity

$$0 = \sum_{\pi \in \mathbb{S}_n} [x_1, x_2, y_{\pi(1)}, y_{\pi(2)}, \ldots, y_{\pi(n)}] = \\ = n! \cdot [x_1, x_2, y_1, y_2, \ldots, y_n].$$

Since $n! \neq 0$ by hypothesis, this is equivalent to the identity yielding nilpotency of class $n + 1$, namely

$$[z_1, z_2, \ldots, z_{n+2}] = 0,$$

(just substitute $z_1 = x_1$, $z_2 = x_2$, $z_3 = y_1$, $z_4 = y_2, \ldots, z_{n+2} = y_n$).

The theorem is proved.

The associated Lie ring of a soluble group of derived length s of prime exponent p is also soluble of derived length $\leq s$ by Theorem 3.2.2 and satisfies the identities $pa = 0$ and $[a, \underbrace{b, b, \ldots, b}_{p-1}] = 0$ by Theorem 3.3.2. This enables us to apply Theorem 3.4.1 to soluble groups of prime exponent.

3.4.3 Corollary. *Every soluble group G of derived length s of prime exponent p is nilpotent and its nilpotency class is at most $\frac{p^s - 1}{p - 1}$.*

Proof. It is clear that we may assume G to be finitely generated. We use induction on s. By the induction hypothesis $G/G^{(s-1)}$ is nilpotent. Hence G is abelian-by-nilpotent and therefore residually finite by a theorem of P. Hall [31]. Let $\{N_\alpha\}$ be

a family of normal subgroups of finite index in G intersecting trivially. It is clear that it is sufficient to prove that each of the factor-groups G/N_α is nilpotent of the required nilpotency class. We may therefore assume that G is finite. As a finite p-group it is nilpotent, and it only remains to apply Theorem 3.4.1 to its associated Lie ring.

The corollary is proved.

Note that Corollary 3.4.3 may be also proved without using Theorem 3.1.5, which was given in § 3.1 without proof. For we have shown in the proof of Theorem 3.4.1 that soluble Lie algebras of derived length 2 and characteristic p, satisfying the $(p-1)$-Engel condition, are nilpotent of class $\leq p$. This implies that soluble groups of derived length 2 of prime exponent p are also nilpotent of class $\leq p$. We may now apply Theorem 2.3.5 directly, since groups of exponent p form a variety. This yields the conclusion of Corollary 3.4.3.

It was Meier-Wunderli [112] who proved in 1951 that metabelian groups of prime exponent p are nilpotent of class $\leq p$; his work, however, says nothing about the nilpotency of soluble groups of exponent p of arbitrary derived length.

The bound for the nilpotency class in Theorem 3.4.1 may be improved to $\frac{n^s-1}{n-1}$ (and, hence, to $\frac{(p-1)^s-1}{p-2}$ in Corollary 3.4.3). We leave this for the reader as an exercise simply outlining the way by stating the following proposition.

3.4.4 Proposition (Higgins [39]). *If a soluble Lie algebra L of derived length s of characteristic p satisfies the n-th Engel condition:*

$$[x, \underbrace{y, y, \ldots, y}_{n}] = 0$$

for all $x, y \in L$ and either $p = 0$, or $n < p$, then for all $k \geq 1$

a) $[\gamma_k(L'), \underbrace{L, L, \ldots, L}_{n}] \subseteq \gamma_{k+1}(L')$,

b) $\gamma_{kn+2}(L) \subseteq \gamma_{k+1}(L')$ *and*

c) L *is nilpotent of class* $\leq \frac{n^s-1}{n-1}$.

Here c) follows from b) by an obvious induction on the derived length s and b) is an easy consequence of a). One can prove a) noting first that if $a \in \gamma_k(L')$ then

$$[a, b_{\pi(1)}, b_{\pi(2)}, \ldots, b_{\pi(m)}] \equiv [a, b_1, b_2, \ldots, b_m] \pmod{\gamma_{k+1}(L')}$$

for any $m \in \mathbb{N}$ and any permutation $\pi \in \mathbb{S}_m$. Now apply arguments similar to those used in the proof of Theorem 3.4.1.

Part II
Automorphisms

Chapter 4

Lie rings admitting automorphisms with few fixed points

In this chapter we prove theorems of Higman, Kreknin and Kostrikin on regular automorphisms of Lie rings, including Kreknin's Theorem on regular automorphisms of arbitrary finite order. Then a theorem on almost regular automorphisms of prime order is proved: if the number of fixed elements is finite (or has finite dimension) then there is a nilpotent subring of bounded nilpotency class and of bounded index (or codimension).

§ 4.1 Extending the ground ring

Let L be a Lie ring, and φ an automorphism of L of finite order n. Let ω be a primitive n-th root of unity. We shall consider the Lie ring $\tilde{L} = L \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ with φ acting naturally, where $\mathbb{Z}[\omega]$ is regarded as a trivial $\mathbb{Z}\langle\varphi\rangle$-module.

We introduce analogues of eigenspaces.

Definition. The additive subgroup of the Lie ring $\tilde{L}$

$$^{i}\tilde{L} = \{l \in \tilde{L} \mid l^{\varphi} = \omega^{i} l\}.$$

is called a *φ-component* of $\tilde{L}$ with respect to ω^{i}. The elements of φ-components are called *φ-homogeneous.*

Definition. An ideal I of the Lie ring $\tilde{L}$ is said to be *φ-homogeneous*, if $I = \sum_{i=0}^{n-1} I \cap {}^{i}\tilde{L}$.

Though $\tilde{L}$ is not a vector space, "almost all" of $\tilde{L}$ decomposes into an "almost direct" sum of φ-components.

4.1.1 Lemma. a) *The following inclusion holds*

$$n\tilde{L} \subseteq {}^0\tilde{L} + {}^1\tilde{L} + \ldots + {}^{n-1}\tilde{L}.$$

b) *If* $l_0+l_1+\ldots+l_{n-1}=0$, *where* $l_i \in {}^i\tilde{L}$, *then* $nl_j = 0$ *for all* $j = 0, 1, \ldots, n-1$.

c) *If the additive group of* L *has no* n*-torsion, which means that for any* $l \in L$ *the equality* $nl = 0$ *implies* $l = 0$, *then the sum of* φ*-components* ${}^0\tilde{L} \oplus {}^1\tilde{L} \oplus \ldots \oplus {}^{n-1}\tilde{L}$ *is direct.*

Proof. a) For each $a \in \tilde{L}$ and each $i = 0, 1, \ldots, n-1$ we define ${}^ia = \sum_{i=0}^{n-1} \omega^{-is} a^{\varphi^s}$. It is easy to see that ${}^ia \in {}^i\tilde{L}$:

$$({}^ia)^{\varphi} = \left(\sum_{s=0}^{n-1} \omega^{-is} a^{\varphi^s}\right)^{\varphi} = \sum_{s=0}^{n-1} \omega^{-is} a^{\varphi^{s+1}} =$$

$$= \omega^i \cdot \sum_{s=0}^{n-1} \omega^{-i(s+1)} a^{\varphi^{s+1}} = \omega^i \cdot \sum_{r=0}^{n-1} \omega^{-ir} a^{\varphi^r} = \omega^i \cdot {}^ia,$$

since $\omega^n = 1$ and $\varphi^n = 1$.

On summing we get na:

$$\sum_{i=0}^{n-1} {}^ia = \sum_{i=0}^{n-1} \sum_{s=0}^{n-1} \omega^{-is} a^{\varphi^s} = \sum_{s=0}^{n-1} a^{\varphi^s} \sum_{i=0}^{n-1} \omega^{-is} = na^{\varphi^0} = na,$$

because for $k = 0$ it is clear that $\sum_{i=0}^{n-1} \omega^0 = n$, and for $k \not\equiv 0 \pmod{n}$ we have $\sum_{i=0}^{n-1} \omega^{ik} = 0$. (Indeed, for $k \not\equiv 0 \pmod{n}$ we have

$$\omega^k \cdot \sum_{i=0}^{n-1} \omega^{ik} = \sum_{i=0}^{n-1} \omega^{(i+1)k} = \sum_{j=0}^{n-1} \omega^{jk},$$

since $\omega^n = 1$, that is, the sum $\sum_{i=0}^{n-1} \omega^{ik}$ does not change on being multiplied by ω^k; therefore $\sum_{i=0}^{n-1} \omega^{ik} = 0$, since $\omega^k \neq 1$.)

b) Applying the automorphisms φ^k, $k = 0, 1, \ldots, n-1$, to $\sum_{i=0}^{n-1} l_i = 0$ (where $l_i \in {}^i\tilde{L}$) we obtain the following n equations

$$\begin{array}{rrrr} l_0 + & l_1 + & l_2 + \ldots + & l_{n-1} = 0 \\ l_0 + & \omega^1 \cdot l_1 + & \omega^2 \cdot l_2 + \ldots + & \omega^{n-1} \cdot l_{n-1} = 0 \\ l_0 + & \omega^2 \cdot l_1 + & \omega^4 \cdot l_2 + \ldots + & \omega^{2(n-1)} \cdot l_{n-1} = 0 \\ \ldots & & & \\ l_0 + \omega^{n-1} \cdot l_1 + & \omega^{2(n-1)} \cdot l_2 + \ldots + & \omega^{(n-1)(n-1)} \cdot l_{n-1} = 0 & \end{array}$$

In order to show that $n \cdot l_i = 0$ for some i, we multiply each of these equations by an appropriate power of ω to make the coefficient of l_i equal to 1 and we then sum all of these equations. It is easy to see that after cancellation each l_j, $j \neq i$, will have coefficient $\sum_{k=0}^{n-1} \omega^{(j-i)k}$ which is 0 (see above). Therefore we get $n \cdot l_i = 0$, as required.

c) This is an immediate consequence of b).

The lemma is proved.

4.1.2 Lemma. *For any i and j*

$$[{}^i\tilde{L}, {}^j\tilde{L}] \subseteq {}^{i+j}\tilde{L},$$

where $i + j$ is calculated modulo n. In particular, the sum of the φ-components ${}^0\tilde{L} + {}^1\tilde{L} + \ldots + {}^{n-1}\tilde{L}$ is a subring of the Lie ring $\tilde{L}$.

Proof. For $a \in {}^i\tilde{L}$ and $b \in {}^j\tilde{L}$ we have

$$[a, b]^\varphi = [a^\varphi, b^\varphi] = [\omega^i \cdot a, \omega^j \cdot b] = \omega^{i+j} \cdot [a, b],$$

and so $[a, b] \in {}^{i+j}\tilde{L}$.

The lemma is proved.

We set

$${}^i a = \sum_{i=0}^{n-1} \omega^{-is} a^{\varphi^s} \in {}^i\tilde{L}$$

and call ${}^i a$ *φ-component* of the element $a \in \tilde{L}$. As we have already seen, we have

$$na = {}^0a + {}^1a + \ldots + {}^{n-1}a. \tag{4.1.3}$$

§ 4.2 Regular automorphisms of soluble Lie rings

Here we prove the Kreknin-Kostrikin Theorem on the nilpotency of a soluble Lie ring with a regular automorphism of prime order. We identify the combinatorial formulation of this theorem as a fact on free $\mathbb{Z}/n\mathbb{Z}$-graduated Lie rings, which will be useful later.

4.2.1 Theorem (Kreknin, Kostrikin [82]). *If a soluble Lie ring L of derived length s admits a regular automorphism φ of prime order p, then it is nilpotent and its nilpotency class is at most $\frac{(p-1)^s-1}{p-2}$.*

We first prove a formally more general assertion for the subring pL.

4.2.2 Theorem. *If φ is an automorphism of prime order p of a Lie ring L then, for any s, we have*

$$\gamma_{f(p,s)+1}(pL) \subseteq {}_{id}\langle C_L(\varphi)\rangle + L^{(s)},$$

where $f(p,s) = \frac{(p-1)^s-1}{p-2}$.

Proof. We extend the ground ring by a primitive p-th root of unity ω, setting $\tilde{L} = L \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ and specify the action of φ on $\tilde{L}$ by regarding $\mathbb{Z}[\omega]$ as a trivial $\mathbb{Z}\langle\varphi\rangle$-module. It suffices to prove the assertion of the theorem for $\tilde{L}$, since

$$ {}_{id}\langle C_{\tilde{L}}(\varphi)\rangle = {}_{id}\langle C_L(\varphi)\rangle \otimes_{\mathbb{Z}} \mathbb{Z}[\omega], \quad \tilde{L}^{(s)} = L^{(s)} \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$$
$$\text{and} \quad \gamma_f(p\tilde{L}) = \gamma_f(pL) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$$

(these equations follow easily from the definitions). Therefore, if

$$\gamma_{f(p,s)+1}(p\tilde{L}) \subseteq {}_{id}\langle C_{\tilde{L}}(\varphi)\rangle + \tilde{L}^{(s)},$$

then

$$\gamma_{f(p,s)+1}(pL) \otimes 1 = \gamma_{f(p,s)+1}(p\tilde{L}) \cap L \otimes 1 \subseteq$$
$$\subseteq ({}_{id}\langle C_L(\varphi)\rangle + L^{(s)}) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega] \cap L \otimes 1 = ({}_{id}\langle C_L(\varphi)\rangle + L^{(s)}) \otimes 1.$$

(see § 1.2). So we may assume from the outset that the ground ring contains ω, that is, $\tilde{L} = L$.

Let H denote the subring ${}^0L + {}^1L + \ldots + {}^{p-1}L$ (see Lemma 4.1.2). By Lemma 4.1.1 a) we have $pL \subseteq H$ and hence the result follows from the last part of the following proposition.

4.2.3 Proposition. *For every s we have*

a) $[\gamma_n(H'), \underbrace{H, H, \ldots, H}_{p-1}] \subseteq \gamma_{n+1}(H') + {}_{id}\langle {}^0L\rangle, \quad n \geq 1;$

b) $\gamma_{(p-1)n+2}(H) \subseteq \gamma_{n+1}(H') +_{id} \langle {}^0L\rangle, \quad n \geq 0;$

c) $\gamma_{f(p,s)+1}(H) \subseteq H^{(s)} + {}_{id}\langle {}^0L\rangle$, *where*

$$f(p,s) = 1 + (p-1) + (p-1)^2 + \ldots + (p-1)^{s-1} = \frac{(p-1)^s - 1}{p-2}.$$

Proof. a) It is clear from the definitions that the ideals H' and $\gamma_k(H')$ are φ-homogeneous. So, in order to prove a) it is sufficient to show that

$$[{}^{i_0}c, {}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{p-1}}y_{p-1}] \in \gamma_{n+1}(H') + {}_{id}\langle {}^0L\rangle \tag{4.2.4}$$

for any φ-homogeneous elements ${}^{i_0}c \in \gamma_n(H') \cap {}^{i_0}L$ and ${}^{i_k}y_k \in {}^{i_k}L$. This is obvious if $i_r = 0$ for some $r = 0, 1, \ldots, p-1$; we therefore assume from now on that $i_r \neq 0$, for all $r = 0, 1, \ldots, p-1$.

Note that any permutation of the elements ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{p-1}}y_{p-1}$ does not change the commutator (4.2.4) modulo the subgroup $\gamma_{n+1}(H')$. This follows from the Jacobi identity – for $[a,b,c] = [a,c,b] + [a,[b,c]]$ and thus, if $a \in \gamma_n(H')$ then $[a,[b,c]] \in \gamma_{n+1}(H')$, since $[b,c] \in H'$ for $b, c \in H$.

Our aim is to rearrange the elements ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{p-1}}y_{p-1}$ in the left-hand side of (4.2.4) by means of a permutation $\pi \in \mathbb{S}_{p-1}$ in order to obtain the congruence

$$i_0 + i_{\pi(1)} + i_{\pi(2)} + \ldots + i_{\pi(s)} \equiv 0 \pmod{p},$$

where the left-hand side is the sum of several first upper indices of the resultant commutator. Then by Lemma 4.1.2

$$[{}^{i_0}c, {}^{i_{\pi(1)}}y_{\pi(1)}, {}^{i_{\pi(2)}}y_{\pi(2)}, \ldots, {}^{i_{\pi(s)}}y_{\pi(s)}] \in {}^0L,$$

for an initial segment of the resultant commutator so that the original commutator (4.2.4) lies in $\gamma_{n+1}(H') + {}_{id}\langle {}^0L\rangle$.

We now prove the following number-theoretic lemma.

4.2.5 Lemma. *Let p be a prime number and let $i_1, \ldots, i_k$ be non-zero elements of $GF(p)$ (not necessarily distinct). We form the set*

$$M = \left\{ \sum_{s \in S} i_s \,\middle|\, S \subseteq \{1, 2, \ldots, k\} \right\},$$

where, by definition, the sum is 0 for $S = \phi$. Then either $M = GF(p)$, or $|M| \geq k+1$.

Proof. We proceed by induction on k. For convenience we denote by $M(s)$ the set of all sums involving $\{i_1, \ldots, i_s\}$. For $k = 1$, we have $|M(1)| = |\{0, i_1\}| = 2$, since $i_1 \neq 0$. Next, if any of the sums $\sigma + i_{k+1}$, where $\sigma \in M(k)$, does not belong to $M(k)$ then $|M(k+1)| \geq |M(k)| + 1 \geq k + 2$ by the induction hypothesis. But if $\sigma + i_{k+1} \in M(k)$ for all $\sigma \in M(k)$, then, starting from 0 we find that 0, $i_{k+1}, 2i_{k+1}, \ldots, (p-1)i_{k+1}$ lie in $M(k)$. These elements are all distinct, because $i_{k+1} \neq 0$ by hypothesis, hence $|M(k)| = |M(k+1)| = p$, that is $M(k+1) = GF(p)$.

The lemma is proved.

We now have $p - 1$ upper indices $i_1, i_2, \ldots, i_{p-1}$ all of which are distinct from 0 modulo p. By Lemma 4.2.5 every residue modulo p may be represented as a sum of some subset of these indices. In particular, the $-i_0$ may be represented in this way. Transposing the elements ${}^{i_k}y_k$ with the corresponding upper indices in order to place them immediately after the element ${}^{i_0}c$ we obtain a commutator with an initial segment from ${}^0L$, which lies in ${}_{id}\langle C_L(\varphi)\rangle$ and equals (4.2.4) modulo $\gamma_{n+1}(H')$.

b) This follows from a) by induction on n. For $n = 0$ we have $\gamma_2(H) = \gamma_1(H')$, and for $n > 0$ we have

$$\gamma_{(p-1)n+2}(H) = \gamma_{(p-1)(n-1)+2+p-1}(H) =$$
$$= [\gamma_{(p-1)(n-1)+2}(H), \underbrace{H, H, \ldots, H}_{p-1}] \subseteq [\gamma_n(H'), \underbrace{H, H, \ldots, H}_{p-1}]$$

by the induction hypothesis. An application of a) to the right-hand side completes the proof.

c) Induction on s. For $s = 1$, obviously, $\gamma_2(H) = H'$. For $s > 1$, according to b) we have

$$\gamma_{f(p,s)+1}(H) = \gamma_{(p-1)f(p,s-1)+1+1}(H) =$$
$$= \gamma_{(p-1)f(p,s-1)+2}(H) \subseteq \gamma_{f(p,s-1)+1}(H') +_{id} \langle {}^0L\rangle.$$

We now apply the induction hypothesis on the soluble Lie ring H' of derived length $s - 1$ to the right-hand side of the above inclusion:

$$\gamma_{f(p,s-1)+1}(H') \subseteq (H')^{(s-1)} +_{id} \langle {}^0L\rangle = H^{(s)} +_{id} \langle {}^0L\rangle.$$

The proof of the proposition, and therefore of Theorem 4.2.2 is complete.

Proof of Theorem 4.2.1. By hypothesis ${}^0L = 0$ and $L^{(s)} = 0$. Therefore, by Theorem 4.2.2, we have

$$0 = \gamma_{f(p,s)+1}(pL) = p^{f(p,s)+1}\gamma_{f(p,s)+1}(L).$$

In particular, this implies that the abelian group $\gamma_{f(p,s)+1}(L)$ is a p-group. This group is clearly φ-invariant, and, if it is not trivial, then the automorphism φ of order p must have non-trivial fixed points in it (Corollary 1.7.3). This contradicts the hypothesis of the theorem and hence $\gamma_{f(p,s)+1}(L) = 0$, as required.

The theorem is proved.

We point out that we have also proved the following combinatorial fact.

4.2.6 Theorem. *Let p be a prime, s a natural number and let $f = f(p,s) = \frac{(p-1)^s-1}{p-2}$. If ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f+1}}y_{f+1}$ are any elements of an arbitrary Lie ring with arbitrary formal upper indices $i_1, i_2, \ldots, i_{f+1} \in \mathbb{Z}$ attached to them, then the simple commutator*

$$[{}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f+1}}y_{f+1}]$$

may be represented as a linear combination of commutators, each of which has the same entry set ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f+1}}y_{f+1}$ and either belongs to $L^{(s)}$ or contains a subcommutator with its upper index sum zero modulo p.

Proof. We can say that Theorem 4.2.6 has already been proved because in the course of proving Proposition 4.2.3 we were using only the Jacobi identity and were actually transforming commutators of φ-homogeneous elements without changing the entry set, and the subcommutators from ${}^0L$ clearly had upper index sums zero (mod p). This metamathematical argument may be made more rigorous by formalizing it within a free Lie ring as follows.

Let ω be a primitive p-th root of unity and let F be a free Lie ring over $\mathbb{Z}[\omega]$ with free generators ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f+1}}y_{f+1}$. For each $i = 1, 2, \ldots, p-1$ denote by ${}^iF$ the additive subgroup of F generated by all commutators in the free generators ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f+1}}y_{f+1}$, such that the upper index sum of their entry set is i modulo p.

Note that the additive group of the ideal ${}_{id}\langle {}^0F\rangle$ is generated by all commutators in elements ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f+1}}y_{f+1}$ which have subcommutators with upper index sum zero (mod p).

We define an automorphism φ of order p of F by putting $l^\varphi = \omega^i \cdot l$ for elements $l \in {}^iF$, $i = 0, 1, \ldots, p-1$, and extending the action of φ to the sum $F = \sum_i {}^iF$ by linearity. It is clear that the additive subgroups ${}^iF$ are the φ-components. Since

the additive group F has no torsion, by Lemma 4.1.1 c), we have

$$F = {}^0F \oplus {}^1F \oplus \ldots \oplus {}^{p-1}F.$$

By Proposition 4.2.3 c) (here $H = F$) the following holds:

$$[{}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f-1}}y_{f+1}] \in F^{(s)} +_{id} \langle {}^0F \rangle. \tag{4.2.7}$$

But the ideals $F^{(s)}$ and ${}_{id}\langle {}^0F \rangle$ are evidently multihomogeneous with respect to the free generators ${}^{i_s}y_s$. Hence the left-hand side of (4.2.7) belongs to the multihomogeneous component of the right-hand side of weight 1 in each of the ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f+1}}y_{f+1}$ which means that the equation required by Theorem 4.2.6 holds. Since the elements ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f-1}}y_{f+1}$ are the free generators of F, exactly the same holds for an arbitrary Lie ring L.

The theorem is proved.

§ 4.3 Regular automorphisms of Lie rings

The main result of this section is Kreknin's Theorem on the solubility of a Lie ring with a regular automorphism of finite order. Higman's Theorem on the nilpotency of a Lie ring with a regular automorphism of prime order will follow from this theorem of Kreknin and from Theorem 4.2.1 of Kreknin and Kostrikin on the nilpotency of soluble Lie rings with a regular automorphism of prime order. This alternative proof of Higman's Theorem also provides an explicit upper estimate for Higman's function bounding the nilpotency class. Here we also single out the combinatorial formulations of these theorems as facts about free $\mathbb{Z}/n\mathbb{Z}$-graduated Lie rings which will be referred to in the next section.

4.3.1 Theorem (Kreknin [83]). *If a Lie ring admits a regular automorphism of finite order n, then it is soluble and its derived length is not greater than $2^n - 2$.*

First we prove the following assertion about the subring nL.

4.3.2 Theorem. *If φ is an automorphism of finite order n of a Lie ring L, then*

$$(nL)^{(f(n))} \subseteq {}_{id}\langle {}^0L \rangle,$$

where ${}^0L = \{l + l^{\varphi} + l^{\varphi^2} + \ldots + l^{\varphi^{n-1}} \mid l \in L\}$ *and* $f(n) = 2^{n-1} - 1$.

By contrast with the method of argumentation in the preceding section, we first prove the corresponding combinatorial result and then deduce Theorem 4.3.2

from it. There is an advantage in doing this – one may consider the Lie ring to be decomposed into a direct sum of φ-components which helps avoid certain technical complications.

4.3.3 Proposition. *Suppose that ω is a primitive n-th root of unity and that the Lie ring H over $\mathbb{Z}[\omega]$ admits an automorphism φ of order n, where H is decomposable into the direct sum of φ-components:*

$$H = {}^0H \oplus {}^1H \oplus {}^2H \oplus \ldots \oplus {}^{n-1}H.$$

Then, for all $k = 1, 2, \ldots, n-1$,

a) $H^{(2^{k-1})} \cap {}^kH \subseteq \langle {}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H\rangle + {}_{id}\langle {}^0H\rangle$;

b) $H^{(2^k-1)} \subseteq \langle {}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H\rangle + {}_{id}\langle {}^0H\rangle$;

c) $H^{(2^{n-1}-1)} \subseteq {}_{id}\langle {}^0H\rangle$.

(Here, as usual, $\langle {}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H\rangle$ denotes the subring generated by the φ-components ${}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H$.)

Proof. Note first of all that under the hypothesis of the proposition the following clearly hold for all s:

$$\begin{aligned} H^{(s)} &= \bigoplus_{i=0}^{n-1} H^{(s)} \cap {}^iH, \\ H^{(s)} \cap {}^wH &= \sum_{u+v=w} [H^{(s-1)} \cap {}^uH,\ H^{(s-1)} \cap {}^vH], \end{aligned} \tag{4.3.4}$$

where $w = 0, 1, \ldots, n-1$, $u, v \geq 0$.

We shall need the following simple lemma.

4.3.5 Lemma. *Suppose that a, b, c are natural numbers such that $1 \leq a \leq n-1$, $1 \leq b \leq n-1$ and $1 \leq c \leq n-1$. If $a + b \equiv c \pmod{n}$, then either both $a > c$ and $b > c$, or both $a < c$ and $b < c$.*

Proof. Since $a < n$ and $b < n$ we also have $a+b < 2n$. Therefore, either $a+b = c$, or $a + b = c + n$. In the first case, clearly $a < c$ and $b < c$ In the second case both numbers are greater than c, because if any of them was less than c, then their sum would be less than $c + n$, since the other is less than n.

The lemma is proved.

We prove parts a) and b) of Proposition 4.3.3 simultaneously by induction on k. For $k = 1$ part a) means that

$$H' \cap {}^1H \subseteq \langle {}^2H, {}^3H, \ldots, {}^{n-1}H\rangle + {}_{id}\langle {}^0H\rangle.$$

According to (4.3.4) the additive subgroup $H' \cap {}^1H$ is generated by commutators of the form $[x, y]$, where $x \in {}^iH$, $y \in {}^jH$ and $i + j \equiv 1 \pmod{n}$. If any of the residues i, j modulo n is 0, then $[x, y] \in {}_{id}\langle {}^0H\rangle$. If, however, both are greater than 0, then by Lemma 4.3.5 both of them are greater than 1, so that $[x, y] \in \langle {}^2H, {}^3H, \ldots, {}^{n-1}H\rangle$. For $k = 1$ part b) means that

$$H' \subseteq \langle {}^2H, {}^3H, \ldots, {}^{n-1}H\rangle +_{id} \langle {}^0H\rangle$$

which follows from a) and (4.3.4).

Now suppose that $k > 1$. We prove a) first of all using the induction hypothesis for both a) and b). By (4.3.4) the additive subgroup $H^{(2^{k-1})} \cap {}^kH$ is generated by commutators of the form $[x, y]$, where $x \in H^{(2^{k-1}-1)} \cap {}^iH$, $y \in H^{(2^{k-1}-1)} \cap {}^jH$ and $i + j \equiv k \pmod{n}$. If any of the residues i, j is $0 \pmod{n}$, then $[x, y] \in {}_{id}\langle {}^0H\rangle$. If both of them are greater than 0, then by Lemma 4.3.5 either both are greater than k, or both are less than k. In the first case it is clear that $[x, y] \in \langle {}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H\rangle$.

In the second case we apply b) to the subring $H^{(2^{k-1}-1)}$ which contains x and y:

$$H^{(2^{k-1}-1)} \subseteq \langle {}^kH, {}^{k+1}H, \ldots, {}^{n-1}H\rangle + {}_{id}\langle {}^0H\rangle.$$

Therefore the element $y \in H^{(2^{k-1}-1)}$ is equal modulo ${}_{id}\langle {}^0H\rangle$ to a linear combination of commutators of the form

$$[u_1, u_2, \ldots, u_q], \quad u_s \in {}^{i_s}H, \ i_s \geq k, \quad \sum_{i=1}^{q} i_s \equiv j \pmod{n}.$$

By repeatedly applying the Jacobi identity $[a, [b, c]] = [a, b, c] - [a, c, b]$, every commutator of the form

$$[x, [u_1, u_2, \ldots, u_q]]$$

may be expressed as a linear combination of simple commutators of the form

$$[x, u_{\pi(1)}, u_{\pi(2)}, \ldots, u_{\pi(q)}],$$

where $\pi \in \mathbb{S}_q$. Therefore, the commutator $[x, y]$ is equal modulo ${}_{id}\langle {}^0H\rangle$ to a linear combination of simple commutators of the form

$$[x, v_1, v_2, \ldots, v_q], \quad v_s \in {}^{j_s}H, \ j_s \geq k, \quad \sum_{s=1}^{q} j_s \equiv j \pmod{n},$$

each of which certainly satisfies

$$i + j_1 + j_2 + \ldots + j_q \equiv k \pmod{n}.$$

If for such a commutator we have $j_q = k$, then

$$i + j_1 + j_2 + \ldots + j_{q-1} \equiv 0 \pmod{n},$$

which means that $[x, v_1, v_2, \ldots, v_{q-1}] \in {}^0H$ whence

$$[x, v_1, v_2, \ldots, v_q] \in {}_{id}\langle {}^0H\rangle.$$

If, however, $j_q > k$, then clearly

$$i + j_1 + j_2 + \ldots + j_{q-1} \equiv t \not\equiv 0 \pmod{n},$$

and $t > k$ by Lemma 4.3.5. In this case

$$[x, v_1, \ldots, v_{q-1}, v_q] \in \langle {}^tH, {}^{j_q}H\rangle \subseteq \langle {}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H\rangle.$$

So, in any case the commutators $[x, v_1, v_2, \ldots, v_q]$ are contained in

$$\langle {}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H\rangle + {}_{id}\langle {}^0H\rangle.$$

This subring therefore also contains $[x, y]$, as required.

We now prove b). Using the induction hypothesis for $k - 1$ we apply b) to the Lie ring $H^{(2^{k-1})}$:

$$(H^{(2^{k-1})})^{(2^{k-1}-1)} \subseteq$$
$$\subseteq \langle H^{(2^{k-1})} \cap {}^kH,\ H^{(2^{k-1})} \cap {}^{k+1}H, \ldots, H^{(2^{k-1})} \cap {}^{n-1}H\rangle + {}_{id}\langle {}^0H\rangle.$$

The additive subgroups $H^{(2^{k-1})} \cap {}^{k+1}H, \ldots, H^{(2^{k-1})} \cap {}^{n-1}H$ are clearly contained in the subring $\langle {}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H\rangle$ which also contains the additive subgroup $H^{(2^{k-1})} \cap {}^kH$ by a) as proved above. Therefore,

$$H^{(2^k-1)} = (H^{(2^{k-1})})^{(2^{k-1}-1)} \subseteq \langle {}^{k+1}H, {}^{k+2}H, \ldots, {}^{n-1}H\rangle + {}_{id}\langle {}^0H\rangle,$$

as required.

Part c) follows from b) on putting $k = n - 1$.

The proposition is proved.

Before we turn to the proof of Theorem 4.3.2, we state a combinatorial consequence of the proposition just proved. We recall the definition of the identities for

soluble varieties:

$$\delta_1 = [x_1, x_2], \quad \delta_{k+1} = [\delta_k(x_1, \ldots, x_{2^k}),\ \delta_k(x_{2^k+1}, \ldots, x_{2^{k+1}})].$$

4.3.6 Theorem. *Suppose that n is a natural number and set $f(n) = 2^{2^{n-1}-1}$. If ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}$ are any elements of an arbitrary Lie ring with arbitrary formal upper indices $i_1, i_2, \ldots, i_{f(n)} \in \mathbb{Z}$ attached to them, then the commutator*

$$\delta_{2^{n-1}-1}({}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)})$$

may be represented as a linear combination of commutators each of which has the same entry set ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}$ and contains a subcommutator with its upper index sum zero modulo p.

Proof. All we need do is to repeat the proof of Theorem 4.2.6 with obvious modifications. We can, in fact, say that Theorem 4.3.6 has already been proved because in the course of proving Proposition 4.3.3 we were using only the Jacobi identity and were, in fact, transforming commutators of φ-homogeneous elements without changing the entry set, and the subcommutators from ${}^0L$ clearly had upper index sums zero (mod p). This metamathematical argument may be made more rigorous by formalizing it within a free Lie ring as follows.

Let ω be a primitive p-th root of unity and let F be a free Lie ring over $\mathbb{Z}[\omega]$ with free generators ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}$. For each $i = 1, 2, \ldots, n-1$ denote by ${}^iF$ the additive subgroup of the Lie ring F generated by all commutators in the free generators ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}$, such that the upper index sum of their entry set is i modulo p.

Note that the additive group of the ideal ${}_{id}\langle {}^0F\rangle$ is generated by all commutators in elements ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}$ such that they have subcommutators with upper index sum zero (mod p).

We define an automorphism φ of order p of F by putting $l^\varphi = \omega^i \cdot l$ for elements $l \in {}^iF$, $i = 0, 1, \ldots, n-1$, and extending the action of φ to the sum $F = \sum_i {}^iF$ by linearity. It is clear that the additive subgroups ${}^iF$ are the φ-components. Since the additive group F has no torsion, by Lemma 4.1.1 c), we have

$$F = {}^0F \oplus {}^1F \oplus \ldots \oplus {}^{n-1}F.$$

By Proposition 4.3.3 c) we have

$$\delta_{2^{n-1}-1}({}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}) \in {}_{id}\langle {}^0F\rangle.$$

But the ideal ${}_{id}\langle {}^0F\rangle$ is clearly multihomogeneous with respect to the free generators ${}^{i_s}y_s$. Hence the left-hand side belongs to the multihomogeneous component

of the right-hand side of weight 1 in each of the ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}$ which means that the equation required in Theorem 4.3.6 holds. Since the elements ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}$ are the free generators of F, exactly the same holds for an arbitrary Lie ring L.

The theorem is proved.

Proof of Theorem 4.3.2. We extend the ground ring by a primitive n-th root of unity ω by putting $\tilde{L} = L \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ and defining the action of φ on $\tilde{L}$ by regarding $\mathbb{Z}[\omega]$ as a trivial $\mathbb{Z}\langle\varphi\rangle$-module. It is sufficient to prove the theorem for $\tilde{L}$ since

$$_{id}\langle {}^0\tilde{L}\rangle = {}_{id}\langle {}^0L\rangle \otimes_{\mathbb{Z}} \mathbb{Z}[\omega] \quad \text{and} \quad (n\tilde{L})^{(f)} = (nL)^{(f)} \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$$

(these equations easily follow from the definitions), and therefore if $(n\tilde{L})^{(f)} \subseteq {}_{id}\langle {}^0\tilde{L}\rangle$, then

$$(nL)^{(f)} \otimes 1 = (n\tilde{L})^{(f)} \cap L \otimes 1 \subseteq {}_{id}\langle {}^0L\rangle \otimes_{\mathbb{Z}} \mathbb{Z}[\omega] \cap L \otimes 1 = {}_{id}\langle {}^0L\rangle \otimes 1.$$

We may therefore assume from the very beginning that the ground ring contains ω, that is, $\tilde{L} = L$.

Put $f(n) = 2^{2^{n-1}-1}$. By Lemma 4.1.1 a) we have $nL \subseteq {}^0L + {}^1L + \ldots + {}^{n-1}L$ and it is therefore sufficient to show that

$$\delta_{2^{n-1}-1}({}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{f(n)}}y_{f(n)}) \in {}_{id}\langle {}^0L\rangle$$

for any φ-homogeneous elements ${}^{i_s}y_s \in {}^{i_s}L$, $s = 1, 2, \ldots, f(n)$. But this easily follows from Theorem 4.3.6.

Theorem 4.3.2 is proved.

4.3.7 Corollary. *Suppose that the Lie ring L admits a regular automorphism of finite order n.*

a) *If the additive group of L has no n-torsion, that is, $nl = 0$ implies $l = 0$, then L is soluble of derived length at most $2^{n-1} - 1$.*

b) *If n is a prime number, then L is soluble of derived length at most $2^{n-1} - 1$.*

Proof. a) By Theorem 4.3.2 we have

$$n^{2^{2^{n-1}-1}} L^{(2^{n-1}-1)} = (nL)^{(2^{n-1}-1)} \subseteq {}_{id}\langle {}^0L\rangle \subseteq {}_{id}\langle C_L(\varphi)\rangle = 0.$$

Restriction on the additive group of L now implies that $L^{(2^{n-1}-1)} = 0$.

b) If n is a prime, then the Sylow n-subgroup of the additive group of L is trivial, since otherwise an automorphism of prime order n acting on it would have

non-trivial fixed points by Corollary 1.7.3. Hence the additive group of L has no n-torsion and the result now follows by applying a).

Proof of Theorem 4.3.1. Let L be a Lie ring, $\varphi \in \operatorname{Aut} L$, $|\varphi| = n$ and $C_L(\varphi) = 0$. We define

$$T = \{a \in L \mid n^k \cdot a = 0 \text{ for some } k = k(a) \in \mathbb{N}\}.$$

It is clear that T is a φ-invariant ideal of L. By Theorem 4.3.2

$$0 = (nL)^{(2^{n-1}-1)} = n^{f(n)} \cdot L^{(2^{n-1}-1)},$$

where $f(n) = 2^{2^{n-1}-1}$, and so $L^{(2^{n-1}-1)} \subseteq T$. Therefore, in order to prove Theorem 4.3.1 it is sufficient to show that $T^{(2^{n-1}-1)} = 0$, since then

$$L^{(2^n-2)} = (L^{(2^{n-1}-1)})^{(2^{n-1}-1)} \subseteq T^{(2^{n-1}-1)} = 0.$$

We decompose the abelian periodic group T into a direct sum of its Sylow subgroups

$$T = T_{p_1} \oplus T_{p_2} \oplus \ldots \oplus T_{p_r},$$

where $n = p_1^{k_1} p_2^{k_2} \ldots p_r^{k_r}$ is the decomposition of n as a product of prime-powers. It is clear that these Sylow subgroups are in fact φ-invariant subrings. It therefore suffices to prove that each of the subrings T_{p_i} is soluble of derived length $\leq 2^{n-1}-1$.

Suppose that for some prime $p \in \{p_1, p_2, \ldots, p_r\}$ we have $n = p^k \cdot s$, where $(p, s) = 1$. We decompose the finite cyclic group $\langle\varphi\rangle$ into the direct product of Hall subgroups

$$\langle\varphi\rangle = \langle\varphi\rangle_p \times \langle\varphi\rangle_{p'},$$

where, of course, $\langle\varphi\rangle_p = \langle\varphi^s\rangle$ and $\langle\varphi\rangle_{p'} = \langle\varphi^{p^k}\rangle$. Then $C_{T_p}(\varphi^{p^k}) = 0$ (where T_p is the Sylow p-subgroup of the additive group T), since otherwise the automorphism φ^s of order p^k normalizing the non-trivial abelian p-subgroup $C_{T_p}(\varphi^{p^k})$, would have in it non-trivial fixed points (see Corollary 1.7.3), which would lie in

$$C_{T_p}(\varphi^{p^k}) \cap C_{T_p}(\varphi^s) = C_{T_p}(\varphi),$$

contrary to the condition $C_L(\varphi) = 0$.

Hence the T_p admits a regular automorphism φ^{p^k} whose order is coprime to the orders of all elements of its additive periodic group. Therefore, by Corollary 4.3.7, T_p is soluble of derived length $\leq 2^{n-1} - 1$ since the order of φ^{p^k} is obviously not greater than n.

Theorem 4.3.1 is proved.

4.3.8 Corollary (Higman [40], Kreknin [83], Kreknin and Kostrikin [82]). *If a Lie ring admits a regular automorphism of prime order p, then it is nilpotent, and its nilpotency class is bounded by some number $h(p)$ which depends only on p and is not greater than* $\frac{(p-1)^{2^{p-1}+1}-1}{p-2}$.

Proof. Such a Lie ring is soluble of derived length $\leq 2^{p-1}-1$ by Corollary 4.3.7, and therefore, by Theorem 4.2.1, it is nilpotent of class $\leq \frac{(p-1)^{2^{p-1}+1}-1}{p-2}$.

We also give a combinatorial formulation of this theorem.

4.3.9 Corollary. *For any prime number p there exists a natural number $h(p)$ depending only on p and not greater than* $\frac{(p-1)^{2^{p-1}+1}-1}{p-2}$, *such that if* ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{h(p)+1}}y_{h(p)+1}$ *are any elements of an arbitrary Lie ring with arbitrary formal upper indices $i_1, i_2, \ldots, i_{h(p)+1} \in \mathbb{Z}$ attached to them, then the simple commutator*

$$[{}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{h(p)+1}}y_{h(p)+1}]$$

may be represented as a linear combination of commutators, each of which has the same entry set ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{h(p)+1}}y_{h(p)+1}$ *and contains a subcommutator with upper index sum zero modulo p.*

Proof. This may be deduced from Corollary 4.3.8 by using free Lie rings in exactly the same way as the proofs of Theorems 4.2.6 and 4.3.6 are deduced from Propositions 4.2.3 and 4.3.3 respectively.

By the Jacobi identity $[a,[b,c]] = [a,b,c] - [a,c,b]$ any commutator may be expressed as a linear combination of simple commutators, each having the same entry set. Hence the ideal of the free Lie ring on free generators ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{h(p)+1}}y_{h(p)+1}$ whose additive group is generated by commutators in the generators, each having a subcommutator with upper index sum zero (mod p), may also be generated by simple commutators, each having an initial segment with upper index sum zero (mod p). This remark allows us to strengthen the conclusion of Corollary 4.3.9.

4.3.10 Corollary (Higman-Kreknin-Kostrikin Theorem). *For any prime number p there exists a natural number $h(p)$ depending only on p and not greater than* $\frac{(p-1)^{2^{p-1}+1}-1}{p-2}$, *such that if* ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{h(p)+1}}y_{h(p)+1}$ *are any elements of an arbitrary Lie ring with arbitrary formal upper indices $i_1, i_2, \ldots, i_{h(p)+1} \in \mathbb{Z}$ attached to them, then the simple commutator*

$$[{}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_{h(p)+1}}y_{h(p)+1}]$$

may be represented as a linear combination of ***simple*** *commutators, each of which has the same entry set* $^{i_1}y_1, ^{i_2}y_2, \ldots, ^{i_{h(p)-1}}y_{h(p)+1}$ *and contains a subcommutator with upper index sum zero modulo* p.

For every prime number p there clearly exists a least number $h(p)$ satisfying Corollary 4.3.8 (or, equivalently, Corollaries 4.3.9 and 4.3.10). We call this *Higman's function* and from now on we reserve for it the notation $h(p)$.

The fact that Higman's function turns out to be the same whether it is defined in terms of Corollary 4.3.8 or in terms of Corollaries 4.3.9 and 4.3.10, warrants explanation. As we saw in the proof of Corollary 4.3.9 (or in the proofs of Theorems 4.2.6 and 4.3.6), the value of $h(p)$ which fits Corollary 4.3.8, also fits Corollaries 4.3.9 and 4.3.10.

To see the converse, we suppose that $h(p)$ is a number satisfying Corollary 4.3.10 and then prove that any Lie ring L with a regular automorphism φ of prime order p is necessarily nilpotent of class $\leq h(p)$. For $\tilde{L} = L \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ where ω is a primitive p-th root of unity, we have $p\tilde{L} \subseteq {}^1L + {}^2L + \ldots + {}^{p-1}L$. Then by Corollary 4.3.10

$$p^{h(p)+1}\gamma_{h(p)+1}(\tilde{L}) = \gamma_{h(p)+1}(p\tilde{L}) \subseteq {}_{id}({}^0L) = 0.$$

Thus the additive group of the subring $\gamma_{h(p)+1}(L)$ is a p-group – therefore it is trivial, since otherwise the automorphism φ of order p acting on it would have non-trivial fixed points by Corollary 1.7.3.

§ 4.4 Almost regular automorphism of prime order

Here we generalize the Higman-Kreknin-Kostrikin Theorem on Lie rings with a regular automorphism of prime order to the case where either the set of fixed points is finite, or, for Lie algebras, the fixed points constitute a finite-dimensional subspace. This "almost regularity" of the automorphism implies that the Lie ring is "almost nilpotent" in the sense that there is a subring of bounded nilpotency class, the bound depending only on the order of the automorphism, and of bounded index (or codimension), the bound depending also on the number (or dimension) of fixed points.

For convenience we shall use the terms like "(p, m)-bounded quantity" to mean that for any natural p and m there exists a natural number $f(p, m)$, such that this quantity does not exceed $f(p, m)$.

4.4.1 Theorem. *Let* φ *be an automorphism of prime order* p *of a Lie ring (algebra)* L. *If the number of fixed points* $C_L(\varphi)$ *is finite and equal to* $|C_L(\varphi)| = q$ *(the dimension of the subalgebra* $C_L(\varphi)$ *is finite and equal to* q*), then* L *has a*

subring (subalgebra) of (p, q)-bounded index (codimension), which is nilpotent of p-bounded class.

(Here the index of a subring is its index as a subgroup of the additive group of the ring.)

Proof. This exploits the Higman-Kreknin-Kostrikin Theorem on Lie rings with regular automorphisms of prime order in its combinatorial form as a result about Lie rings with a $\mathbb{Z}/p\mathbb{Z}$-graduation (Corollary 4.3.10). The method of constructing the nilpotent subring of bounded index resembles the proofs of the converses of the Schur-Baer Theorems in § 2.4. Instead of centralizers it uses generalized centralizers relative to the commutation of φ-components of the Lie ring.

We note that it seems impossible to prove that a Lie ring with an almost regular automorphism of prime order necessarily contains a subring of bounded index with a regular automorphism of prime order.

At first we prove a consequence of the Higman-Kreknin-Kostrikin Theorem, which is also of combinatorial nature.

In what follows p is the prime occurring in the hypothesis of the theorem.

4.4.2 Proposition. *For any m and n there exists an (m, n, p)-bounded number $f = f(m, n, p)$ such that any simple commutator of weight f in the elements ${}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_f}x_f$ of an arbitrary Lie ring with arbitrary upper indices $i_s \not\equiv 0 \pmod{p}$ attached to them, is equal to a linear combination of commutators, each having the same entry set $X = \{{}^{i_s}x_s \mid 1 \leq s \leq f\}$ and each either containing a subcommutator of the form*

$$[{}^{k_1}w_1, {}^{k_2}w_2, \ldots, {}^{k_t}w_t], \quad {}^{k_i}w_i \in X, \tag{4.4.3}$$

which has m initial segments with upper index sum zero modulo p:

$$k_1 + k_2 + \ldots + k_{r_i} \equiv 0 \pmod{p}, \quad i = 1, 2, \ldots, m$$
$$1 < r_1 < r_2 < \ldots < r_m = t,$$

or containing a subcommutator of the form

$$[{}^{k_0}w, c_1, c_2, \ldots, c_n], \tag{4.4.4}$$

where ${}^{k_0}w \in X$ and each of the n simple commutators c_i has the form

$$[{}^{k_1}u_1, {}^{k_2}u_2, \ldots, {}^{k_s}u_s], \quad {}^{k_j}u_j \in X$$

with upper index sum zero modulo p:

$$k_1 + k_2 + \ldots + k_s \equiv 0 \pmod{p}.$$

We can put $f(m, n, p) = 1 + \sum_{i=1}^{m-1} (h(p)+1)^{i+1} \cdot n^i$, *where* $h(p)$ *is Higman's function as in § 4.3.*

Proof. It is clearly sufficient to prove this proposition regarding the elements ${}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_f}x_f$ as the free generators of a free Lie ring L. We define on L the $\mathbb{Z}/p\mathbb{Z}$-graduation

$$L = L_0 \oplus L_1 \oplus L_2 \oplus \ldots \oplus L_{p-1},$$

where for each s the additive subgroup L_s is generated by all commutators in the generators ${}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_f}x_f$ having upper index sum congruent to s modulo p.

For brevity we write $h = h(p)$ for the Higman's function.

We recall that Corollary 4.3.10 states that for any elements ${}^{j_i}a_i \in L_{j_i}$, $i = 1, 2, \ldots, h+1$, the simple commutator

$$[{}^{j_1}a_1, {}^{j_2}a_2, \ldots, {}^{j_{h+1}}a_{h+1}]$$

of weight $h+1$ is equal to a linear combination of simple commutators in the same elements, each having the same entry set and each having an initial segment from L_0 which is an initial segment with upper index sum zero modulo p.

In fact, the proof of the proposition consists in multiple applications of this assertion. Before giving it in full using the inevitable formalism of induction, we highlight its basic ideas by doing the first steps. The initial segment of weight $h+1$ of the commutator under consideration,

$$[{}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_f}x_f], \tag{4.4.5}$$

is equal to a linear combination of commutators with the same entry set and with initial segments from L_0 so that the whole commutator (4.4.5) is a linear combination of commutators of the form

$$[{}^{k_1}x_{i_1}, {}^{k_2}x_{i_2}, \ldots, {}^{k_r}x_{i_r}, {}^{k_{r+1}}x_{i_{r+1}}, \ldots],$$
$$r \leq h+1, \quad k_1 + k_2 + \ldots + k_r \equiv 0 \pmod{p},$$

with the same entry set. For each such commutator put

$${}^{k_{r+1}}y_1 = -[{}^{k_1}x_{i_1}, {}^{k_2}x_{i_2}, \ldots, {}^{k_r}x_{i_r}, {}^{k_{r+1}}x_{i_{r+1}}]$$

(it is clear that ${}^{k_{r+1}}y_1 \in L_{k_{r+1}}$) and denote

$$ {}^{k_{r+s}}y_s = {}^{k_{r+s}}x_{i_{r+s}} \text{ for } s \geq 2. $$

Note that ${}^{k_{r+1}}y_1 = [{}^{k_r}x_{i_{r+1}}, c_0]$, where

$$ c_0 = [{}^{k_1}x_{i_1}, {}^{k_2}x_{i_2}, \ldots, {}^{k_r}x_{i_r}] \in L_0. $$

We obtain a simple commutator of the form

$$ [{}^{k_{r+1}}y_1, {}^{k_{r+2}}y_2, \ldots] $$

whose weight is not less than $f - h$. For sufficiently large f we have $f - h \geq h + 2$, and we may then apply the same transformation to each of the resultant commutators, substituting its initial segment of weight $h + 1$ by its expression as a linear combination of simple commutators in the same elements ${}^{k_{r+1}}y_1, {}^{k_{r+2}}y_2, \ldots$ with initial segments from L_0. In each of these transformed commutators an element of the form ${}^{k_{r+1}}y_1 = [{}^{k_{r+1}}x_{i_{r+1}}, c_0]$ either occurs in the initial segment from L_0 – and this is a step towards the form (4.4.3) or does not occur there – and this serves for accumulation of occurrences of subcommutators from L_0 and may be regarded as a step towards the form (4.4.4). These transformations may be performed often enough, if f is large enough, since at each step, on replacing the variables, the weight diminishes at most by h.

These first steps may be illustrated by a picture ("l. c." being the abbreviation for "linear combination"):

$$ \begin{aligned} [x_1, x_2, \ldots, x_f] &= \text{l. c. } [[c_0, x], x, \ldots] \\ &\qquad\qquad \| \quad \| \\ &= \text{l. c. } \quad [y_1, \quad y_2, \ldots] \text{ of weight } \geq f - h \\ &= \text{l. c. } [[c_0(y), y], y, \ldots] \\ &\qquad\qquad \| \quad \| \\ &= \text{l. c. } \quad [z_1, \quad z_2, \ldots] \text{ of weight } \geq f - 2h \\ &= \ldots \end{aligned} $$

Now we start the full exposition of the proof of the proposition.

We consider commutators of the form

$$ [{}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_r}y_r], \tag{4.4.6} $$

where for each s the element ${}^{i_s}y_s$ has the from

$$ {}^{i_s}y_s = [{}^{i_s}x, c_1, c_2, \ldots, c_k], \quad k \geq 0 \tag{4.4.7} $$

where ${}^{i_s}x$ denotes one of the elements of X (the entry set of the commutator (4.4.5)), and each element $c_i \in L_0$ also has the form (4.4.6) with $r \leq h+1$ and $i_1 + i_2 + \ldots + i_r \equiv 0 \pmod{p}$.

We define the *height* of the commutator (4.4.6) inductively as a sum of heights of the elements ${}^{i_s}y_s (s = 1, 2, \ldots, r)$ the height of an element (4.4.7) being equal to the sum of k and the sum of the heights of the elements $c_1, c_2, \ldots, c_k$; for $k = 0$ the height of an element ${}^{i_s}y_s = {}^{i_s}x$ of the form (4.4.7) is 0 by definition.

We introduce an *HKK-transformation* of the commutator (4.4.6) which consists in

a) representing it as a linear combination of simple commutators in the same elements ${}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_r}y_r$ with initial segments from L_0 of weight $\leq h+1$, that is, commutators of the form

$$[c, {}^{k_{w+1}}y_{i_{w+1}}, {}^{k_{w+2}}y_{i_{w+2}}, \ldots, {}^{k_r}y_{i_r}], \tag{4.4.8}$$

where

$$\begin{gathered} c = [{}^{k_1}y_{i_1}, {}^{k_2}y_{i_2}, \ldots, {}^{k_w}y_{i_w}] \in L_0, \\ w \leq h+1, \quad k_1 + k_2 + \ldots + k_w \equiv 0 \pmod{p}, \end{gathered} \tag{4.4.9}$$

and, in light of this

b) changing notation

$${}^{k_{w+1}}z_1 = -[c, {}^{k_{w+1}}y_{i_{w+1}}], \quad {}^{k_{w+s}}z_s = {}^{k_{w+s}}y_{i_{w+s}} \text{ for } s \geq 2.$$

This transformation is possible by Corollary 4.3.10. We shall say that the resultant commutators of the form

$$[{}^{k_{w+1}}z_1, {}^{k_{w+2}}z_2, \ldots, {}^{k_r}z_{r-w}] \tag{4.4.10}$$

are obtained from (4.4.6) by means of HKK-transformation.

4.4.11 Lemma. *Every commutator of the form (4.4.10), obtained from the commutator (4.4.6) by means of HKK-transformation, is also a commutator of the form (4.4.6), and its height is precisely* 1 *greater.*

Proof. It is clear that the elements ${}^{k_{w+s}}z_s = {}^{k_{w+s}}y_{i_{w+s}}$ for $s \geq 2$ have the required form (4.4.7). For $s = 1$ we substitute into ${}^{k_{w+1}}z_1 = [{}^{k_{w+1}}y_{i_{w+1}}, c]$ the expression for ${}^{k_{w+1}}y_{i_{w+1}}$ given by (4.4.7):

$${}^{k_{w+1}}z_1 = [{}^{k_{w+1}}y_{i_{w+1}}, c] = [{}^{i_{w+1}}x, c_1, c_2, \ldots, c_k, c].$$

Here the c_i have the form (4.4.6) with $r \leq h+1$ and c is also of the form (4.4.6) with $r \leq h+1$ by (4.4.9). Hence ${}^{k_{m-1}}z_1$ has the form (4.4.7), and the whole commutator (4.4.10) has the form (4.4.6).

Evaluating the height of (4.4.10), we see that the only difference between it and the height of (4.4.6) is the additional occurrence of the element $c \in L_0$ in ${}^{k_{m-1}}z_1$.

The lemma is proved.

If HKK-transformation is applied sufficiently many times, it produces commutators satisfying the conclusion of Proposition 4.4.2.

4.4.12 Lemma. *If a commutator of the form (4.4.6) has weight $r \leq h+1$, and height $\geq f_1(m,n,p) = \sum_{i=1}^{m-1}(h(p)+1)^i \cdot n^i$, then it is equal to a linear combination of commutators, each containing either a subcommutator of the form (4.4.3), or a subcommutator of the form (4.4.4) from the conclusion of proposition 4.4.2.*

Proof. We construct a graph of occurrences of subcommutators $c \in L_0$ in the commutator (4.4.6). Its vertices will be partitioned into different levels, and each level will be partitioned into groups. The vertices of level 1 are the elements c_i from the expressions (4.4.7) for all elements ${}^{i_s}y_s$, occurring in the commutator (4.4.6) under consideration. These vertices c_i of level 1 are partitioned into groups with respect to elements ${}^{i_s}y_s$ – they form a group if they occur in the same expression (4.4.7). Every element c_i of level 1 also has the form (4.4.6) by definition and by formulae (4.4.7) it contains other elements from L_0 – the latter are called the vertices of level 2 connected with the given vertex of level 1. The vertices of level 2 are partitioned into groups in the same way, and so on.

It is obvious, that the total number of vertices of this graph is equal to the height of the original commutator (4.4.6). We note also that each vertex is connected with the vertices of the next level from not more than $h+1$ groups, and at level 1 there are also not more than $h+1$ groups, since the ambient commutator (4.4.6) has weight $r \leq h+1$ by the hypothesis of the lemma.

We shall show that either the number of levels is at least m, or at least one of the groups in some level contains at least n vertices. Indeed, if we suppose that this is not the case, then at level 1 there are less than $(h+1)n$ vertices, since there are at most $h+1$ groups at this level, and then at level 2 there are less than $(h+1)^2n^2$ vertices, because each vertex of level 1 is connected with vertices of level 2 from at most $h+1$ groups each containing less than n vertices, and so on. It is easy to establish by induction that at level i there are less than $(h+1)^i n^i$ vertices. The total number of vertices will be less than

$$\sum_{i=1}^{m-1}(h(p)+1)^i \cdot n^i = f_1(m,n,p).$$

which contradicts the hypothesis of the lemma.

But if some group at some level has at least n vertices, then there is a subcommutator of the form (4.4.4). And if the number of levels is at least m then there is a nested chain of subcommutators from L_0 of length m. Transforming them in turn starting from inside, by the Jacobi identity $[a,[b,c]] = [a,b,c] - [a,c,b]$, into linear combinations of simple commutators (in the elements of $X = \{{}^{i_s}x_s\}$) from L_0, we eventually reach a linear combination of commutators, each containing a subcommutator of the form (4.4.3).

The lemma is proved.

To prove Proposition 4.4.2 we apply HKK-transformation to the commutator (4.4.5), then to each commutator of the form (4.4.6) in the resultant linear combination which is equal to (4.4.5), and so on, $f_1(m,n,p)$ times. At every step the weight of each commutator of the form (4.4.10) is less than the weight of the commutator (4.4.6) from which it was obtained by means of HKK-transformation, by not more than $h+1$. Hence, if the initial weight is at least $f(m,n,p) = (h+1)\cdot f_1(m,n,p)+1$ then it is possible to apply HKK-transformation $f_1(m,n,p)$ times. If the resultant linear combination contains commutators of the form (4.4.6) of weight $> h+1$, then we proceed to apply HKK-transformation to them, until we get a linear combination of commutators of the form (4.4.6) of weight $r \leq h+1$, which is equal to the commutator (4.4.5).

Since the (4.4.5) is also a commutator of the form (4.4.6) of height 0, then after the transformations described above, it will be expressed by Lemma 4.4.11 as a linear combination of commutators of the form (4.4.6) of weight $\leq h+1$ and height $\geq f_1(m,n,p)$. Application of Lemma 4.4.12 completes the proof of the proposition.

We now turn to the proof of Theorem 4.4.1. Suppose that L is a Lie ring and φ is an automorphism of L of prime order p. We first consider the case where the number of fixed points is finite: $|C_L(\varphi)| = q$. Let $\tilde{L} = L \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$, where ω is a primitive p-th root of unity, and define the natural action of φ on $\tilde{L}$ by regarding $\mathbb{Z}[\omega]$ as a trivial $\mathbb{Z}\langle\varphi\rangle$-module. It is easy to see that $C_{\tilde{L}}(\varphi) = C_L(\varphi) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$, and hence $|C_{\tilde{L}}(\varphi)| = |C_L(\varphi)|^{p-1} = q^{p-1}$. If the assertion of Theorem 4.4.1 has been proved for $\tilde{L}$, then the intersection of its subring of (p, q^{p-1})-bounded index, which is nilpotent of p-bounded class, with L will clearly be the desired subring of L of (p,q)-bounded index, which is nilpotent of the same p-bounded class. We may therefore assume from the outset that the ground ring contains ω, that is, $\tilde{L} = L$.

At first we shall prove – and this will be the main part of the proof of the theorem – that the subring pL contains a subring of (p,q)-bounded index (in pL) which is nilpotent of p-bounded class.

As it was shown in § 4.1

$$pL \subseteq {}^0L + {}^1L + \ldots + {}^{p-1}L,$$

where ${}^iL$ are the φ-components of L. In each of the φ-components ${}^iL$ for $i \neq 0$ we shall construct generalized centralizers of different levels $s = 1, 2, \ldots, h+1$ (we recall that $h = h(p)$ is Higman's function). These are the additive subgroups ${}^iK(s) \leq {}^iL$ which contain each other and have (p, q)-bounded indices in ${}^iL$. Since $|{}^0L| = |C_L(\varphi)| = q$, the subring

$$pL \cap \langle {}^1K(h+1), {}^2K(h+1), \ldots, {}^{p-1}K(h+1)\rangle,$$

generated by the generalized centralizers of level $h+1$, has (p, q)-bounded index in pL. We shall prove that this subring is nilpotent of class $< f(h+1, h, p)$, where f is the function appearing in the statement of Proposition 4.4.2, and therefore will be the sought one (at this stage of the proof of Theorem 4.4.1, dealing with pL). We put $N = f(h+1, h, p)$ for brevity.

Definition. A *pattern* of a commutator in φ-homogeneous elements ${}^{i_s}x_s \in {}^{i_s}L$ is defined to be its bracket structure together with the arrangement of the upper indices in it. The commutator itself will be called the *value* of this pattern at the elements ${}^{i_s}x_s$.

For example,

$$[[{}^1a, {}^3b, {}^4c], [{}^2d, {}^5e]] \quad \text{and} \quad [[{}^1x, {}^3y, {}^4z], [{}^2u, {}^5v]]$$

are values of the same pattern

$$[[{}^1*, {}^3*, {}^4*], [{}^2*, {}^5*]].$$

We shall further need homomorphisms which resemble the mappings $\mu_a\colon x \to [x, a]$ into the set of commutators arising in the proof of Lemma 2.4.3.

For every ordered set (of arbitrary length k)

$$\bar{x} = ({}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_k}x_k)$$

of φ-homogeneous elements ${}^{i_s}x_s \in {}^{i_s}L$ we choose $j \leq p-1$, such that

$$j + i_1 + i_2 + \ldots + i_k \equiv 0 \pmod{p},$$

and define a homomorphism

$$\vartheta_{\bar{x}} : {}^{j}y \to [{}^{j}y, {}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_k}x_k], \quad {}^{j}y \in {}^{j}L,$$

from the additive subgroup ${}^{j}L$ into ${}^{0}L$. Since $|{}^{0}L| = q$, the following result holds.

4.4.13 Lemma. *For any $\bar{x}$ we have $|{}^{j}L : \operatorname{Ker}\vartheta_{\bar{x}}| \leq q$.*

The construction of generalized centralizers is achieved by induction, simultaneously fixing of certain φ-homogeneous elements, called *representatives* of different levels. First we put ${}^{i}K(1) = {}^{i}L$ for all $i \neq 0$.

Next, for each element $c \in {}^{0}L$ which may be represented as a value of some pattern of weight $\leq N$ at φ-homogeneous elements ${}^{i_s}x_s \in {}^{i_s}L$ for $i_s \neq 0$, we fix one such presentation for each possible pattern of weight $\leq N$. The fixed φ-homogeneous elements ${}^{i_s}x_s$ occurring in all of these presentations are called representatives of level 1, level being indicated by parenthesis: ${}^{i_s}x_s(1)$. It is clear that the total number of representatives of level 1 is (p, q)-bounded, since the number of all patterns of weight $\leq N = f(h+1, h, p)$ is evidently p-bounded and the number of elements $c \in {}^{0}L$ is at most q.

Now we construct the generalized centralizers ${}^{j}K(2)$ of level 2, putting for each $j = 1, 2, \ldots, p-1$

$${}^{j}K(2) = \bigcap_{\bar{x}} \operatorname{Ker}\vartheta_{\bar{x}},$$

where $\bar{x} = ({}^{i_1}x_1(1), {}^{i_2}x_2(1), \ldots, {}^{i_k}x_k(1))$ runs through all ordered sets of lengths $k \leq N$ of representatives of level 1 (and, of course, $j + i_1 + i_2 + \ldots + i_k \equiv 0 \pmod{p}$ in accordance with the definition of the homomorphisms $\vartheta_{\bar{x}}$). Since the number of these ordered sets is clearly (p, q)-bounded, and the index of each of the subgroups $\operatorname{Ker}\vartheta_{\bar{x}}$ in ${}^{j}L$ is at most q by Lemma 4.4.13, then the index of the subgroup ${}^{j}K(2)$ in ${}^{j}L$ is also (p, q)-bounded. (We recall that the index of an intersection of subgroups is not greater than the product of their indices.)

We note that the following centralizing property holds for the elements ${}^{j}y \in {}^{j}K(2)$ with respect to the representatives of level 1:

$$[{}^{j}y, {}^{i_1}x_1(1), {}^{i_2}x_2(1), \ldots, {}^{i_k}x_k(1)] = 0,$$

whenever $j + i_1 + i_2 + \ldots + i_k \equiv 0 \pmod{p}$ and $k \leq N$.

Now we proceed by induction. Suppose that for $t < h+1$ and for each $j = 1, 2, \ldots, p-1$ we have already constructed generalized centralizers of levels $\leq t$, that is additive subgroups of ${}^{j}L$

$${}^{j}K(1) \geq {}^{j}K(2) \geq \ldots \geq {}^{j}K(t)$$

of (p, q)-bounded index in ${}^{j}L$, and that we have fixed representatives ${}^{i_s}x_s(\varepsilon_s) \in {}^{i_s}L$ for $i_s \neq 0$ of levels $\varepsilon_s \leq t$, whose total number is (p, q)-bounded. Suppose also that for every $s \leq t$ the following centralizing property holds for the elements ${}^{j}y \in {}^{j}K(s)$ with respect to the representatives of levels $< s$:

$$\begin{gathered}[{}^{j}y, {}^{i_1}x_1(\varepsilon_1), {}^{i_2}x_2(\varepsilon_2), \ldots, {}^{i_k}x_k(\varepsilon_k)] = 0 \\ \text{whenever } {}^{j}y \in {}^{j}K(s), \quad k \leq N, \\ \varepsilon_j < s \ (j = 1, 2, \ldots, k) \quad \text{and} \quad j + i_1 + i_2 + \ldots + i_k \equiv 0 \pmod{p}.\end{gathered} \tag{4.4.14}$$

Now for $j = 1, 2, \ldots, p-1$ set

$${}^{j}K(t+1) = \bigcap_{\bar{x}} \operatorname{Ker} \vartheta_{\bar{x}},$$

where $\bar{x} = ({}^{i_1}x_1(\varepsilon_1), {}^{i_2}x_2(\varepsilon_2), \ldots, {}^{i_k}x_k(\varepsilon_k))$ runs through all ordered sets of lengths $k \leq N$ of representatives of levels $\varepsilon_u \leq t$ (and, of course, $j + i_1 + \ldots + i_k \equiv 0 \pmod{p}$ in accordance with the definition of the homomorphisms $\vartheta_{\bar{x}}$). The total number of all such ordered sets is obviously (p, q)-bounded, and the index of each of the subgroups $\operatorname{Ker} \vartheta_{\bar{x}}$ in ${}^{j}L$ is at most q by Lemma 4.4.13. Thus the index of the subgroup ${}^{j}K(t+1)$ in ${}^{j}L$ is also (p, q)-bounded. It is also clear that ${}^{j}K(t+1) \leq {}^{j}K(t)$.

It follows from the construction that the centralizing property (4.4.14) holds for the elements ${}^{j}y \in {}^{j}K(t+1)$ for $s = t+1$ with respect to the representatives of levels $\leq t$.

Next, for every element $c \in {}^{0}L$ which may be represented as a value of some pattern of weight $\leq N$ at φ-homogeneous elements ${}^{i_s}x_s \in {}^{i_s}K(t+1)$ for $i_s \neq 0$, we fix one such representation for each possible pattern of weight $\leq N$. The fixed φ-homogeneous elements ${}^{i_s}x_s$, occurring in all of these representations, are called the representatives of level $(t+1)$: ${}^{i_s}x_s(t+1)$. It is clear that the total number of representatives of level $(t+1)$ is also (p, q)-bounded.

We have finished the inductive definition of the generalized centralizers ${}^{j}K(1)$, ${}^{j}K(2), \ldots, {}^{j}K(h+1)$, $\quad j = 1, 2, \ldots, p-1$.

As was already noted at the beginning of the proof, the subring

$$\langle {}^{1}K(h+1), {}^{2}K(h+1), \ldots, {}^{p-1}K(h+1) \rangle,$$

generated by the generalized centralizers of level $h+1$ intersects pL in a subring of (p, q)-bounded index in pL. We shall prove that this subring is nilpotent of class $< N = f(h+1, h, p)$ where f is the function from Proposition 4.4.2. For this it is sufficient to show that every simple commutator of weight N of the form

$$[{}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_N}y_N], \quad {}^{i_s}y_s \in {}^{i_s}K(h+1), \tag{4.4.15}$$

in elements generating this subring is equal to 0 (see 3.1.2). In order to do this, in turn, it is sufficient by Proposition 4.4.2 to show that commutators of the form (4.4.3) and (4.4.4) from the conclusion of Proposition 4.4.2, applied to the commutator (4.4.15) with $m = h + 1$ and $n = h$, are all equal to 0.

We consider first the commutator

$$[{}^{j}y, c_1, c_2, \ldots, c_h] \tag{4.4.16}$$

of the form (4.4.4). For every $s = 1, 2, \ldots, h$ we replace c_s by its expression as a value of a pattern of weight $< N$ at the representatives of level s which were fixed above. This is possible, since c_s is the value of a pattern of weight $< N$ at elements of ${}^{i}K(h+1)$, $i \neq 0$, and since we have inclusions ${}^{i}K(s) \supseteq {}^{i}K(h+1)$ for all $s \leq h$, so that, by construction, there is also a representation of c_s as a value of that same pattern at representatives of level s.

Expanding the inner brackets using the formula $[a, [b, c]] = [a, b, c] - [a, c, b]$, we represent (4.4.16) as a linear combination of commutators of the form

$$[{}^{j}y, {}^{i_1}x(1), \ldots, {}^{i_k}x(1), {}^{i_{k+1}}x(2), \ldots, {}^{i_t}x(2), \ldots, {}^{i_{t+1}}x(h), \ldots, {}^{i_u}x(h)],$$
$${}^{j}y \in {}^{j}K(h+1). \tag{4.4.17}$$

Here, for simplicity, we have dropped the lower indices on the representatives since the only important thing here is that they are placed in order of increasing level.

The main idea in what follows is to represent each of the commutators (4.4.17) as a linear combination of simple commutators with initial segments of weight $h + 1$ consisting of representatives of ***different*** levels $1, 2, \ldots, h$ and the element ${}^{j}y \in {}^{j}K(h+1)$. (This will allow us via the Higman-Kreknin-Kostrikin Theorem to apply the centralizing property (4.4.14) as a result of which all these initial segments will turn out to be equal to 0.) To achieve this we begin by using the formula $[a, b, c] = [a, c, b] + [a, [b, c]]$ to transpose the first (from the left) in (4.4.17) representative of level 1 to the left to the second place directly after the element ${}^{j}y$, then to transpose the first in (4.4.17) of the representatives of level 2 to the left to the third place, and so on, aiming to arrive to a commutator with initial segment

$$[{}^{j}y, {}^{v_1}x(1), {}^{v_2}x(2), \ldots, {}^{v_h}x(h)].$$

Of course, in performing such transformations some additional summands will appear. Nevertheless, they also may all be transformed into linear combination of commutators of a similar form, and they will also be equal to 0 by some generalization of formulae (4.4.14).

Definition. A *quasirepresentative* of level s is a commutator in representatives which contains only one representative of maximal level s, the levels of the remaining representatives being smaller than s. The representatives themselves are also regarded as quasirepresentatives of the same level.

4.4.18 Lemma. *If* ${}^{i_1}\hat{x}_1(\varepsilon_1), {}^{i_2}\hat{x}_2(\varepsilon_2), \ldots, {}^{i_k}\hat{x}_k(\varepsilon_k)$ *are quasirepresentatives of levels* $\varepsilon_u < s$ $(u = 1, 2, \ldots, k)$ *and an element* ${}^{j}y$ *either belongs to* ${}^{j}K(s)$ *or is a quasirepresentative of level* s, *then*

$$[{}^{j}y, {}^{i_1}\hat{x}_1(\varepsilon_1), {}^{i_2}\hat{x}_2(\varepsilon_2), \ldots, {}^{i_k}\hat{x}_k(\varepsilon_k)] = 0, \tag{4.4.19}$$

whenever $j + i_1 + i_2 + \ldots + i_k \equiv 0 \pmod{p}$ *and* $k \leq N$.

Proof. We express each of the quasirepresentatives ${}^{i_u}\hat{x}_u(\varepsilon_u)$ as a commutator in the representatives of levels $< s$ and using the formula $[a, [b, c]] = [a, b, c] - [a, c, b]$ we expand the inner brackets in the commutator (4.4.19) as a commutator in these representatives and ${}^{j}y$.

If ${}^{j}y \in {}^{j}K(s)$, then (4.4.19) is equal to a linear combination of commutators of the form (4.4.14), which are equal to 0.

If ${}^{j}y$ is a quasirepresentative of level s, then the element ${}^{j}y$, as a commutator in representatives, is equal to a linear combination of simple commutators in representatives which start with the unique representative of maximal level s, which belongs to ${}^{j}K(s)$ by definition. Again the commutator (4.4.19) is equal to a linear combination of commutators of the form (4.4.14), which are all equal to 0.

The lemma is proved.

4.4.20 Lemma. *Any commutator of the form (4.4.17) is equal to a linear combination of commutators of the form*

$$[{}^{j}y, {}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_h}\hat{x}(h), {}^{v_{h+1}}\hat{x}(\varepsilon_{h+1}), \ldots], \tag{4.4.21}$$

in whose initial segments after ${}^{j}y \in {}^{j}K(h+1)$ *are situated quasirepresentatives, one from each level* $1, 2, \ldots, h$, *in order of increasing level.*

Proof. To the commutator (4.4.17) we apply a collecting process which will be defined now (this is not to be mixed up with the collecting process for groups from § 2.7). At each step the commutator (4.4.17) will be equal to a linear combination of commutators of the form

$$[{}^{j}y, {}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_s}\hat{x}(s), {}^{v_{s+1}}\hat{x}(\varepsilon_{s+1}), \ldots], \tag{4.4.22}$$

in ${}^{j}y$ and quasirepresentatives, each having the property that to the left of the first quasirepresentative (from the left) of any given level there lie only ${}^{j}y$ and

quasirepresentatives of lower levels. In particular, (4.4.17) is also a commutator of form (4.4.22). In such commutators (4.4.22) the collected part is the maximal initial segment of the form

$$[{}^{j}y, {}^{v_1}\hat{x}(1), \ldots, {}^{v_s}\hat{x}(s)], \quad s \geq 1,$$

which contains to the right of ${}^{j}y$ quasirepresentatives one of each level $1, 2, \ldots, s$ placed in order of increasing level.

The main step in the collecting process is to transpose in each of the commutators (4.4.22) the ***first*** quasirepresentative of level $s+1$ (from the non-collected part) – say, the element $\hat{x}(s+1)$ – to the left, to the end of the collected part, using the formula

$$[\ldots b, \hat{x}(s+1)\ldots] = [\ldots \hat{x}(s+1), b\ldots] + [\ldots [b, \hat{x}(s+1)]\ldots]$$

(where dots denote unaltered parts). After a finite number of such transpositions, (4.4.22) will be transformed into a commutator with the longer collected part in a sum with a linear combination of additional summands. However, all of these additional summands are also commutators of form (4.4.22), being ***shorter*** than that one to which this step of the collecting process is applied. Indeed, firstly, in the additional commutator of the form $[\ldots [b, \hat{x}(s+1)]\ldots]$ which arises, the subcommutator $[b, \hat{x}(s+1)]$ is a quasirepresentative of level $s+1$, since the level of the quasirepresentative b is less than $s+1$ by definition of (4.4.22); secondly there are no quasirepresentatives of levels $\geq s+1$ to the left from the quasirepresentative $[b, \hat{x}(s+1)]$, because there were none such to the left of $\hat{x}(s+1)$.

Note also that all commutators of the form (4.4.22) which appear, including the shorter ones, contain quasirepresentatives of all levels $1, 2, \ldots, h$. Of course, greater levels cannot appear so that this collecting process will finish after a finite number of steps, and (4.4.17) will be represented by the desired linear combination.

The lemma is proved.

We now show that any commutator of form (4.4.21) is 0. By the Higman-Kreknin-Kostrikin Theorem 4.3.10, the initial segment of (4.4.21) of weight $h+1$, namely,

$$[{}^{j}y, {}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_h}\hat{x}(h)]$$

is equal to a linear combination of simple commutators in the same elements ${}^{j}y, {}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_h}\hat{x}(h)$ with initial segments from ${}^{0}L$. If such an initial segment contains ${}^{j}y$ then on representing it as a linear combination of simple commutators with the same entry set beginning with ${}^{j}y$, we get 0 by Lemma 4.4.18, since ${}^{j}y \in {}^{j}K(h+1)$. If such an initial segment does not contain ${}^{j}y$ then it contains exactly one quasirepresentative ${}^{v_s}\hat{x}(s)$ of maximal level s. Representing

such an initial segment as a linear combination of simple commutators with the same entry set beginning with ${}^{v_s}\hat{x}(s)$, we also get 0 by Lemma 4.4.18.

Now by Lemma 4.4.20 any commutator of form (4.4.17), and hence any commutator of form (4.4.16), is equal to 0.

We now consider the commutator

$$[{}^{k_1}w_1, {}^{k_2}w_2, \ldots, {}^{k_t}w_t], \quad {}^{k_i}w_i \in X, \tag{4.4.23}$$

of the form (4.4.3) from the conclusion of Proposition 4.4.2, applied to (4.4.15) with $m = h + 1$ and $n = h$. This commutator has $h + 1$ distinct initial segments with upper index sums zero modulo p:

$$k_1 + k_2 + \ldots + k_{r_i} \equiv 0 \pmod{p}, \quad i = 1, 2, \ldots, h+1,$$
$$1 < r_1 < r_2 < \ldots < r_{h+1} = t.$$

The commutator (4.4.23) belongs to ${}^0L$ and is a commutator of elements from the generalized centralizers ${}^iK(h+1)$ of level $h+1$. It may therefore be represented as a value of the same pattern at representatives of level $h + 1$

$$[{}^{k_1}x(h+1), {}^{k_2}x(h+1), \ldots, {}^{k_t}x(h+1)]. \tag{4.4.24}$$

(Here, the lower indices are again dropped for simplicity.)

Next, the initial segment of length r_h of (4.4.24) also belongs to ${}^0L$ and is a commutator in elements from the generalized centralizers ${}^iK(h)$ of level h, since ${}^iK(h) \geq {}^iK(h+1)$. We replace it by its expression as a value of the same pattern at representatives of level h and so on. As a result, (4.4.23) will be equal to a commutator of the form

$$\left[{}^{k_1}x(1), \ldots, {}^{k_r}x(1), {}^{k_{r+1}}x(2), \ldots, {}^{k_s}x(2), \ldots, {}^{k_{u+1}}x(h+1), \ldots, {}^{k_t}x(h+1)\right]. \tag{4.4.25}$$

4.4.26 Lemma. *Every commutator of the form (4.4.25) is equal to a linear combination of commutators of the form*

$$[{}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_{h+1}}\hat{x}(h+1), {}^{v_{h+2}}\hat{x}(\varepsilon_{h+2}), \ldots], \tag{4.4.27}$$

whose initial segments contain one quasirepresentative of each of the levels $1, 2, \ldots, h + 1$ *located in order of increasing level.*

Proof. This is absolutely analogous to the proof of Lemma 4.4.20, the only difference being that the collected parts of the commutators, appearing in the course of

the collecting process, are initial segments of the form

$$[{}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_s}\hat{x}(s)], \quad s \geq 1.$$

Now in essentially the same way as in the case of (4.4.21), it is proved that any commutator of the form (4.4.27) is equal to 0. For this we apply the Higman-Kreknin-Kostrikin Theorem 4.3.10 to its initial segment of weight $h+1$ in order to represent it as a linear combination of simple commutators in the same elements ${}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_{h+1}}\hat{x}(h+1)$ with initial segments from ${}^0L$. Each of these latter initial segments contains exactly one quasirepresentative ${}^{v_s}\hat{x}(s)$ of maximal level s. Expressing such initial segments as linear combinations of simple commutators with the same entry set, beginning with ${}^{v_s}\hat{x}(s)$, we get 0 by Lemma 4.4.18.

Hence, by Lemma 4.4.26, commutators of the form (4.4.25), and therefore also commutators of the form (4.4.23), are equal to 0.

Therefore the subring

$$\langle {}^1K(h+1), {}^2K(h+1), \ldots, {}^{p-1}K(h+1)\rangle$$

generated by generalized centralizers of level $h+1$ is nilpotent of class $< N = f(h+1, h, p)$, where f is the function in the statement of Proposition 4.4.2.

We have proved that the subring

$$L_1 = pL \cap \langle {}^1K(h+1), {}^2K(h+1), \ldots, {}^{p-1}K(h+1)\rangle,$$

which has a (p, q)-bounded index in pL, is nilpotent of p-bounded class. We now finish the proof of Theorem 4.4.1 in the case of a finite number of fixed points.

Note that the subring L_1 is clearly φ-invariant since all of the additive subgroups $pL, {}^1K(h+1), {}^2K(h+1), \ldots, {}^{p-1}K(h+1)$ are φ-invariant.

We shall prove that the subring

$$L_2 = \frac{1}{p}L_1 = \{l \in L \mid pl \in L_1\}$$

has (p, q)-bounded index in L and that it also contains a subring of (p, q)-bounded index which is nilpotent of p-bounded class.

The assertion about the index $|L : L_2|$ is elementary: if $a_1, a_2, \ldots, a_r$ are the representatives of the cosets of L_1 in pL, then $a_1 = pb_1$, $a_2 = pb_2, \ldots, a_r = pb_r$ for some elements $b_1, b_2, \ldots, b_r \in L$. Now for any $l \in L$ we have: $pl \in a_i + L_1$ for some i and hence $pl = a_i + l_1 = pb_i + l_1$ for some $l_1 \in L_1$. This implies $p(l - b_i) \in L_1$, that is $l - b_i \in L_2$, so that $l \in b_i + L_2$. Hence,

$$(b_1 + L_2) \cup (b_2 + L_2) \cup \ldots \cup (b_r + L_2) = L$$

and the index of the additive subgroup L_2 in L is not greater, than $r = |pL : L_1|$.

It is clear that L_2 is also φ-invariant.

Suppose that the subring L_1 is nilpotent of class $g-1$ where $g = g(p)$, that is, $\gamma_g(L_1) = 0$. Then we have

$$0 = \gamma_g(L_1) \supseteq \gamma_g(pL_2) = p^g \gamma_g(L_2),$$

so that the additive subgroup $\gamma_g(L_2)$ is a p-group whose exponent divides p^g. This subgroup is clearly φ-invariant, and, since $|C_L(\varphi)| = q$, its rank is (p,q)-bounded by Corollary 1.7.4. Hence the subring $\gamma_g(L_2)$ is finite and its order is (p,q)-bounded. By Theorem 3.1.6 this implies that L_2 contains a subring of (p,q)-bounded index which is nilpotent of class $\leq g$. This subring is what is required, since its index in L is also (p,q)-bounded.

So, we have completed the proof of Theorem 4.4.1 in the case of a finite number of fixed points.

Now let L be a Lie algebra with an automorphism φ of prime order p, and let $\dim C_L(\varphi) = q < \infty$. If the characteristic of the ground field is equal to p, then by Corollary 1.7.5 the dimension of the whole algebra L is (p,q)-bounded.

For the case where the characteristic of the ground field is different from p, the proof may be obtained by simply repeating the above arguments in the case of a finite number of fixed points, replacing the words "order $|C_L(\varphi)|$" and "index" by "dimension $\dim C_L(\varphi)$" and "codimension", respectively (note, that here, under the hypothesis on the characteristic, we have $L = {}^0L \oplus {}^1L \oplus \ldots \oplus {}^{p-1}L$ after extending the ground field by a primitive p-th root of unity).

The theorem is proved.

§ 4.5 Comments

The contents of § 4.1–4.3 are to a large extent taken from Chapter VIII of [49], where the proof of the Higman-Kreknin-Kostrikin Theorem from the works of Kreknin and Kostrikin [82, 83] is given. We have introduced only a few modifications in order to emphasize the combinatorial nature of this theorem, because it is in that form that it is used in § 4.4. It is also more convenient to consider first of all a Lie ring which is a direct sum of its φ-components, and only afterwards to apply the result to the general case.

Regular automorphisms of prime order. The upper bound for the Higman's function $h(p)$ in Corollaries 4.3.8, 4.3.9, 4.3.10 is very far from Higman's conjecture that $h(p) = \frac{p^2-1}{4}$ for $p > 2$ and $h(2) = 1$. In [40] Higman has constructed

examples showing that $h(p) \geq \frac{p^2-1}{4}$ for $p > 2$. In the same paper Higman has shown that $h(5) = 6$. In addition, Scimemi has verified with the aid of a computer that $h(7) = 12$, and I. Hughes confirmed this later in [47].

It is quite straightforward to prove that $h(2) = 1$ and $h(3) = 2$. For $p = 2$ we only have to consider in a $\mathbb{Z}/2\mathbb{Z}$-graduated Lie ring $L = {}^0L + {}^1L$ a commutator of weight 2 in elements from ${}^1L$, which clearly belongs to ${}^0L$. For $p = 3$ one has to consider in a $\mathbb{Z}/3\mathbb{Z}$-graduated Lie ring $L = {}^0L + {}^1L + {}^2L$, commutators $[a, b, c]$ of weight 3 in elements a, b, c either from ${}^1L$ or from ${}^2L$. If $a, b, c \in {}^1L$ or $a, b, c \in {}^2L$, then $[a, b, c] \in {}^0L$. In all other cases one can assume without loss of generality that $a \in {}^1L$; if $b \in {}^2L$, then $[a, b] \in {}^0L$; and if $b \in {}^1L$ and $c \in {}^2L$, then by the Jacobi identity we have

$$[a, b, c] = [a, c, b] + [a, [b, c]] \in {}_{id}\langle {}^0L \rangle,$$

since both $[a, c]$ and $[b, c]$ are in ${}^0L$.

For applications of the Higman-Kreknin-Kostrikin Theorem to groups with regular or almost regular automorphisms of prime order see Chapter 5.

Regular automorphisms of non-prime order. Kreknin [84] proved that a ***finite-dimensional*** Lie algebra admitting a regular automorphism of infinite order, is also soluble. The proof is by reduction to the case where the regular automorphism has finite order.

Unfortunately, in the case of composite order, Kreknin's Theorem 4.3.1 on the solubility of Lie rings with a regular automorphism does not allow us to estimate the derived length of a nilpotent group with a regular automorphism, because there is no good correspondence between the derived length of a nilpotent group and that of its associated Lie ring. (However, Kreknin's Theorem has been recently successfully used in the "modular" case, where an automorphism of order p^k acts on a finite p-group – see Chapter 8, and in the theory of pro-p-groups of finite coclass – see the Comments in § 5.4.) Perhaps progress could be made if the following conjecture were to be proved: if a Lie ring admits a regular automorphism of order p^k (or of order $p_1^{k_1} \cdot p_2^{k_2} \cdot \ldots \cdot p_s^{k_s}$, $k = k_1 + k_2 + \ldots + k_s$), then it has a series of nested ideals whose length is bounded in terms of k and whose factor-rings are nilpotent of (p, k)-bounded class. This conjecture corresponds to some known results about finite groups with regular automorphisms – see § 5.4.

Up to now such a theorem is known to be true only in the simplest case where the regular automorphism has order 4. We consider the situation where $2L = L$, $\varphi \in \operatorname{Aut} L$, $|\varphi| = 4$ and $C_L(\varphi) = 0$. Then $\tilde{L} = L \otimes \mathbb{Z}[i] = {}^1L + {}^2L + {}^3L$, where ${}^iL$ are the φ-components. For the ideal $H = {}_{id}\langle -l + l^{\varphi^2} \mid l \in \tilde{L} \rangle$ we have, first of all $H = \langle {}^1L, {}^3L \rangle$ (that is, H is generated by ${}^1L$ and ${}^3L$ as a Lie ring). Secondly, the factor-ring $\tilde{L}/H$ is nilpotent of class 1 because ${}^2L$ admits the automorphism of

order 2 induced by φ, which acts like multiplication by -1 and by the hypothesis $2L = L$.

Simple combinatorics show that the Lie ring H is nilpotent of class ≤ 3. To prove this assertion it is sufficient to show that an arbitrary simple commutator of weight 4 in elements from ${}^1L$ or ${}^3L$ is 0. Note that $[a, b] = 0$, if $a \in {}^1L$, $b \in {}^3L$. We consider all possible patterns; without loss of generality we may assume that an element from ${}^1L$, if it occurs, is placed at the start:

$$[{}^1*, {}^1*, {}^1*, {}^1*] = 0;$$

$$[{}^1*, {}^1*, {}^1*, {}^3*] = [{}^1*, {}^1*, {}^3*, {}^1*] = [{}^1*, {}^3*, {}^1*, {}^1*] = 0;$$

$$[{}^1*, {}^1*, {}^3*, {}^3*] = [{}^1*, {}^3*, {}^1*, {}^3*] = [{}^1*, {}^3*, {}^3*, {}^1*] = 0;$$

$$[{}^1*, {}^3*, {}^3*, {}^3*] = 0: \quad [{}^3*, {}^3*, {}^3*, {}^3*] = 0.$$

Analogous calculations of a slightly more complicated nature show that also $[L^{(2)}, L] = 0$, so that, in particular, $\gamma_3(\gamma_2(L)) = 0$. This theorem is due to Kovács who proved in [79] the same (but more difficult) assertion for groups with a regular automorphism of order 4.

As we saw, the "combinatorial" value of Higman's function, satisfying Corollaries 4.3.9 and 4.3.10, is the same with the value fitting Corollary 4.3.8 on the nilpotency of Lie rings with a regular automorphism of prime order. It would be interesting to know whether a similar phenomenon occurs for Kreknin's functions in the statements of Theorems 4.3.1 and 4.3.2 – we had to double the function from 4.3.2 in order to prove Theorem 4.3.1.

Regular groups of automorphisms. In the theory of finite groups a lot of progress has been made in the study of groups with regular groups of automorphisms, that is groups G such that $C_G(A) = 1$ for some group of automorphisms $A \leq \operatorname{Aut} G$, which is not assumed to be cyclic, but has order coprime to the order of G. On the one hand, such a group must be soluble (this was verified by Clemens [15] modulo the classification of finite simple groups). On the other hand, for soluble groups with this property, the so-called nilpotent length is bounded (see § 5.4). However, for Lie rings there is a simple example showing that the existence of a regular group of automorphisms is not sufficient for the solubility of a Lie ring, even in the case where the regular group of automorphisms is a four group (that is, non-cyclic of order 4). Let L be a three-dimensional simple Lie algebra (over an arbitrary field of characteristic $\neq 2$) with basis e_1, e_2, e_3 and structural constants

$$[e_1, e_2] = e_3, \quad [e_2, e_3] = e_1, \quad [e_3, e_1] = e_2.$$

Linear transformations, which have matrices

$$\begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$$

relative to e_1, e_2, e_3, generate a subgroup $A \leq \operatorname{Aut} L$ such that $|A| = 4$ and $C_L(A) = 0$.

Almost regular automorphisms. If, for whatever reasons, it is known beforehand, that a Lie ring admitting an almost regular automorphism of prime order p, is soluble of derived length s, then it may turn out to be more valuable to estimate the index of the nilpotent subring in terms of s, p and q and its class in terms of p and s (where q is the number of fixed points) – this was done in [63]. Although this paper contains only a theorem on nilpotent groups with almost regular automorphisms of prime order which are soluble of derived length s, the analogous result on soluble Lie rings (without the additional nilpotency hypothesis) can be proved in exactly the same way.

Theorem 4.4.1 was proved in [68] under the additional assumption that $pL = L$; the extension to the general case was also obtained independently by Medvedev [111].

For almost regular automorphisms of prime order it would be interesting to find out the best (or better) estimates for the index and nilpotency class of the subring mentioned in Theorem 4.4.1. Of course, as we have already said, the exact values of Higman's function $h(p)$ are not yet known, but it is possible to look for the best (or better) estimates in the form of functions depending on $h(p)$ as we in fact did in the proof of Theorem 4.4.1. (Note that explicit expressions for these functions may be easily extracted from the proof of Theorem 4.4.1.) For example, may $h(p) + 1$ be a bound for the nilpotency class of a subring of (p, q)-bounded index? As it was shown by Hartley and Meixner in [35], this is true in the case $p = 2$. On the other hand it is apparently impossible to prove that a Lie ring, satisfying the conditions of Theorem 4.4.1, necessarily contains a subring of (p, q)-bounded index which is nilpotent of class $h(p)$.

As far as almost regular automorphisms of composite order are concerned, it is natural to conjecture that if a Lie ring admits an automorphism of order n, having a finite number q of fixed points, then it should contain a subring of (n, q)-bounded index, which is soluble of n-bounded derived length. However, up to now, this has not been proved even in the simplest case $n = 4$.

Chapter 5

Nilpotent groups admitting automorphisms of prime order with few fixed elements

The results of this chapter are based on the theorems of the previous one about automorphisms of Lie rings. First of all we prove Higman's Theorem bounding the nilpotency class of a nilpotent group which admits a regular automorphism of prime order.

We next show the rather unexpected fact that the Higman-Kreknin-Kostrikin Theorem is applicable to the "modular" case of finite p-groups with an automorphism of order p. We prove that if a finite p-group admits an automorphism of order p with exactly p^m fixed points then it contains a subgroup of (p, m)-bounded index which is nilpotent of class $h(p)$.

Then a theorem on nilpotent groups with an almost regular automorphism of prime order is proved: if a nilpotent group admits an automorphism of prime order p with exactly m fixed points then it contains a subgroup of (p, m)-bounded index which is nilpotent of p-bounded class. In the proof of this theorem considerable difficulties in going from groups to Lie rings and back have to be overcome. These difficulties arise because there is no good correspondence between subgroups of the group and subrings of the associated Lie ring.

We have included also the results of Makarenko, who refined the bounds for the nilpotency class of the subgroups in the theorems on almost regular automorphisms, and of Medvedev, who generalized the theorem on periodic nilpotent groups with an almost regular automorphism of prime order to the case of arbitrary nilpotent groups.

The latest results of Shalev and the author on automorphisms of order p^k acting on p-groups with few fixed points are contained in Chapter 8.

§ 5.1 Regular automorphisms of prime order

Using the Higman-Kreknin-Kostrikin Theorem one can obtain a bound for the nilpotency class of a nilpotent group which admits a regular automorphism of prime order p. In the case of a periodic group it suffices to consider its associated

Lie ring. By Corollary 1.7.3 the Sylow p-subgroup of such a group is trivial. Then, by Theorem 1.6.2, the induced automorphism of the associated Lie ring is also regular. Hence, by Corollary 4.3.8, this Lie ring is nilpotent of class $\leq h(p)$, and hence the group is also nilpotent of the same nilpotency class. But we cannot in general claim that the regularity of an automorphism of a group implies the regularity of the induced automorphism of the associated Lie ring. Nevertheless using another construction we are able to reduce the proof of the following theorem to the case of a regular automorphism of a Lie ring.

5.1.1 Theorem (Higman [40]). *If a nilpotent group G admits a regular automorphism φ of prime order p, then its nilpotency class is at most $h(p)$, where $h(p)$ is Higman's function bounding the nilpotency class of a Lie ring with a regular automorphism of prime order p.*

Proof. We fix notation as follows: c is the nilpotency class of the group G and $\gamma_i = \gamma_i(G)$ are the members of the lower central series of G.

We note first of all that G has no p-torsion. For if it had we would have a non-trivial subgroup $I_p(1)$, which, being a characteristic subgroup, is also φ-invariant. Since $I_p(1)$ is nilpotent, its centre $Z(I_p(1))$ is a non-trivial abelian φ-invariant p-subgroup. Then $C_{Z(I_p(1))}(\varphi) \neq 1$ by Corollary 1.7.3, which contradicts the fact that φ is regular.

By Theorem 2.6.2 the series of p-isolators

$$G \geq I_p(\gamma_2) \geq I_p(\gamma_3) \geq \ldots \geq I_p(\gamma_c) \geq I_p(\gamma_{c+1}) = I_p(1) = 1$$

of the members of the lower central series of G is a strongly central series of G. According to Theorem 3.2.6 this series gives rise to a Lie ring

$$L = \bigoplus_{i=1}^{c} I_p(\gamma_i)/I_p(\gamma_{i+1}),$$

where the nilpotency class of L coincides with the nilpotency class of G by Theorem 3.2.7. Therefore, if we prove that the induced automorphism φ of L is also regular, then an application of the Higman-Kreknin-Kostrikin Theorem will finish the proof.

It is clear that it is sufficient to show that φ induces regular automorphisms of all the factor-groups $G/I_p(\gamma_i)$. By an obvious induction on c, it is sufficient to show this for $G/I_p(\gamma_c)$. Assuming the contrary we have $C_{G/I_p(\gamma_c)}(\varphi) \neq 1$. Suppose that $gI_p(\gamma_c) \in C_{G/I_p(\gamma_c)}(\varphi)$ for $g \notin I_p(\gamma_c)$. Since $I_p(\gamma_c) \leq Z(G)$ (this is a particular case of the result that the series of p-isolators $\{I_p(\gamma_i)\}$ is strongly central in groups without p-torsion), the group $\langle g, I_p(\gamma_c)\rangle$ is commutative. Hence the elements $g, g^{\varphi}, g^{\varphi^2}, \ldots, g^{\varphi^{p-1}}$ commute since they lie in this group. Thus it is

clear that

$$g \cdot g^{\varphi} \cdot g^{\varphi^2} \cdot \ldots \cdot g^{\varphi^{p-1}} \in C_G(\varphi),$$

and so $g \cdot g^{\varphi} \cdot g^{\varphi^2} \cdot \ldots \cdot g^{\varphi^{p-1}} = 1$, since $C_G(\varphi) = 1$. But we have

$$g \cdot g^{\varphi} \cdot g^{\varphi^2} \cdot \ldots \cdot g^{\varphi^{p-1}} \equiv g^p \pmod{I_p(\gamma_c)},$$

so that we get $g^p \equiv 1 \pmod{I_p(\gamma_c)}$. This means, however, that $gI_p(\gamma_c)$ is a non-trivial element of order p in $G/I_p(\gamma_c)$, which contradicts the definition of $I_p(\gamma_c)$.

Thus, $C_{G/I_p(\gamma_i)}(\varphi) = 1$ for all i, and therefore $C_L(\varphi) = 0$. By Corollary 4.3.8, L is nilpotent of class $\leq h(p)$ and hence G is also nilpotent of class $\leq h(p)$.

The theorem is proved.

If we happen to know the derived length of a group in advance, then it may be better to use Theorem 4.2.1 instead of Corollary 4.3.8. By essentially repeating the proof of Theorem 5.1.1 we obtain the following result.

5.1.2 Theorem. *If a nilpotent group which is soluble of derived length s admits a regular automorphism of prime order p, then its nilpotency class is at most $\frac{(p-1)^s-1}{p-2}$.*

§ 5.2 Nilpotent p-groups with an automorphism of order p

Suppose that P is a nilpotent p-group, φ an automorphism of order p of P and $|C_P(\varphi)| = p^m$. We ask the question: how can one restrict the structure of G in terms of p and m?

At first glance the Higman-Kreknin-Kostrikin Theorem does not seem relevant to this situation. As we saw in the case of a regular automorphism, the orders of the group elements are coprime to p, and φ acts on abelian sections and on the associated Lie ring in a semi-simple, diagonalizable way, which seems quite impossible for the action of an automorphism of order p on a vector space over a field of characteristic p. However, as was first noticed by Alperin, the generalized combinatorial form of Corollary 4.3.10 allows us to apply the Higman-Kreknin-Kostrikin Theorem in the "modular" case. In fact, by Theorem 1.6.1 and Corollary 1.7.4, the ranks of the φ-invariant sections of P are bounded, and, modulo small subgroups or small indices, one has to consider sections of large exponent where the Higman-Kreknin-Kostrikin Theorem can work.

5.2.1 Theorem. *If a nilpotent p-group P admits an automorphism φ of prime order p with exactly p^m fixed points, then it has a subgroup of (p, m)-bounded index which is nilpotent of class $\leq h(p)$, where $h(p)$ is Higman's function.*

Proof. At first we shall prove the existence of a subgroup of (p, m)-bounded index with a slightly worse bound on the nilpotency class, namely, $h(p) + 1$. For what follows we put $h = h(p)$.

Consider the Lie ring $L = L(P) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$, where ω is a primitive p-th root of unity, and φ acts naturally on L. By Lemma 4.1.1 a) and Corollary 4.3.10 we have

$$\gamma_{h+1}(pL) \subseteq {}_{id}\langle C_L(\varphi)\rangle.$$

5.2.2 Lemma. *In the Lie ring L we have $p^m {}_{id}\langle C_L(\varphi)\rangle = 0$.*

Proof. We note first that it is easy to deduce from the definitions that

$$C_{L(P)}(\varphi) = \bigoplus_{i=1}^{\infty} C_{\gamma_i/\gamma_{i+1}}(\varphi) \quad \text{and} \quad C_L(\varphi) = C_{L(P)}(\varphi) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega],$$

where $\gamma_j = \gamma_j(P)$. By Theorem 1.6.1 we have

$$|C_{\gamma_i/\gamma_{i+1}}(\varphi)| \leq |C_P(\varphi)| = p^m,$$

so that, in additive notation, $p^m C_{\gamma_i/\gamma_{i+1}}(\varphi) = 0$ by Lagrange's Theorem. Hence $p^m C_L(\varphi) = 0$ whence also $p^m {}_{id}\langle C_L(\varphi)\rangle = 0$. (Indeed, every element of ${}_{id}\langle C_L(\varphi)\rangle$ is a linear combination of elements of the form $[c, a_1, \ldots, a_k]$, where $c \in C_L(\varphi)$. We have for such elements

$$p^m[c, a_1, \ldots, a_k] = [p^m c, a_1, \ldots, a_k] = [0, a_1, \ldots, a_k] = 0.)$$

The lemma is proved.

By this lemma we have

$$p^{m+h+1}\gamma_{h+1}(L) = p^m \gamma_{h+1}(pL) \leq p^m {}_{id}\langle C_L(\varphi)\rangle = 0.$$

Since $\gamma_{h+1}(L) = \bigoplus_{i>h} \gamma_i(L)/\gamma_{i+1}(L)$, this implies that all factor-groups $\gamma_i(P)/\gamma_{i+1}(P)$ for $i \geq h + 1$ have exponents dividing p^{m+h+1}. Furthermore, by Corollary 1.7.4 the ranks of all these factor-groups are at most pm (we recall that $|C_{\gamma_i/\gamma_{i+1}}(\varphi)| \leq |C_P(\varphi)| = p^m$). The restrictions on the rank and exponent taken together clearly restrict the orders of these abelian factor-groups:

$$|\gamma_i(P)/\gamma_{i+1}(P)| \leq p^{pm(m+h+1)} \text{ for all } i \geq h + 1.$$

It is clear that the same estimate is valid for any φ-invariant subgroup P_1:

$$|\gamma_i(P_1)/\gamma_{i+1}(P_1)| \leq p^{pm(m+h+1)} \text{ for all } i \geq h + 1. \tag{5.2.3}$$

On the other hand by P. Hall's Theorem 2.8.7 we have:

$$|\gamma_i(\gamma_r(P))/\gamma_{i+1}(\gamma_r(P))| \geq p^r \text{ as soon as } \gamma_{i+1}(\gamma_r(P)) \neq 1.$$

Compare this inequality with (5.2.3), setting $P_1 = \gamma_N(P)$. We see that when $N = pm(m + h + 1) + 1$, the only way to avoid a contradiction is to conclude that $\gamma_{h+2}(\gamma_N(P)) = 1$. Thus we have found a subgroup $\gamma_N(P)$ of P which is nilpotent of class $\leq h + 1$ and whose corresponding factor-group is nilpotent of class $\leq pm(m + h + 1)$.

However, a minor modification in the proof enables us to restrict the structure of P even more. This modification involves applying the above argument to the group $P\langle\varphi\rangle$ which admits the automorphism φ with exactly p^{m+1} fixed points. We obtain

$$\gamma_{h+2}(\gamma_M(P\langle\varphi\rangle)) = 1,$$

where $M = p(m + 1)(m + 1 + h + 1) + 1$.

The automorphism φ centralizes the factors of the lower central series of the group $P\langle\varphi\rangle$ so that

$$|\gamma_s(P\langle\varphi\rangle)/\gamma_{s+1}(P\langle\varphi\rangle)| \leq |C_{P\langle\varphi\rangle}(\varphi)| = p^{m+1}$$

(and the left-hand side is even $\leq p^m$ for $s \geq 2$, since $\gamma_2(P\langle\varphi\rangle) \leq P$). Hence $|P\langle\varphi\rangle : \gamma_M(P\langle\varphi\rangle)| \leq p^{m(M-1)+1}$ and $|P : \gamma_M(P\langle\varphi\rangle)| \leq p^{m(M-1)}$.

Thus $\gamma_M(P\langle\varphi\rangle)$ is the required subgroup of (p, m)-bounded index, which is nilpotent of class $\leq h + 1$.

We shall now show that P even has a subgroup of (p, m)-bounded index which is nilpotent of class $\leq h$. By what we have just proved we may assume that P is nilpotent of class $h + 1$. It was shown above that $p^{m+h+1}L(P) = 0$. Since $\gamma_{h+2}(P) = 1$, this means that $\gamma_{h+1}(P)^{p^{m-h+1}} = 1$. Since a homomorphism exists of the tensor product $\underbrace{P/P' \otimes \ldots \otimes P/P'}_{h+1}$ onto $\gamma_{h+1}(P) = \gamma_{h+1}(P)/\gamma_{h+2}(P)$, we have

$$[a_1{}^{p^r}, a_2{}^{p^r}, \ldots, a_{h+1}{}^{p^r}] = [a_1, a_2, \ldots, a_{h+1}]^{p^{r(h-1)}}$$

and for $r = r(p, m)$ large enough (for instance, $r = \left[\frac{m+h+1}{h+1}\right] + 1$) the right-hand side of this equation is equal to 1.

Furthermore, using the fact that $[\gamma_i, \gamma_j] \leq \gamma_{i+j}$ for all i, j we see that every commutator of weight $h + 1$ which involves at least one element from P' lies in γ_{h+2}, and hence equals 1.

Thus we have shown that every commutator of weight $h + 1$ in the generators of $P^{p^r}P'$ is equal to 1. Hence this subgroup is nilpotent of class $\leq h(p)$ (see Theorem 2.2.2), and, since both the rank and the exponent of the abelian factor-

group $P/(P^{p^r}P')$ are (p, m)-bounded, its order, that is $|P : P^{p^r}P'|$, is also (p, m)-bounded.

The theorem is proved.

Note that we may extract an explicit bound for the index of a nilpotent subgroup of class $\leq h(p) + 1$, namely

$$p^{mp(m+1)(m+1+h+1)};$$

and an explicit upper bound for the value $h(p)$ is contained in § 4.3. It is equally easy to obtain a bound for the index of a nilpotent subgroup of class $\leq h(p)$.

If we happen to know the derived length s of the group P, it may turn out to be better to estimate the index and the nilpotency class by functions which also depend on s. Such estimates are produced using Theorem 4.2.2 in essentially the same way as above.

5.2.4 Theorem. *If a soluble p-group P of derived length s admits an automorphism φ of prime order p with exactly p^m fixed points, then it has a subgroup of index*

$$\leq p^{p^m(m+1)\left(\frac{(p-1)^s-1}{p-2}+m+2\right)},$$

which is nilpotent of class $\leq \frac{(p-1)^s-1}{p-2} + 1$.

Proof. We note that a soluble p-group P is locally finite and locally nilpotent. And, since the ranks of the abelian sections of P are bounded, it is finitely generated and hence finite and nilpotent. It then suffices to repeat the proof of Theorem 5.2.1 with the obvious modifications – replacing certain functions and replacing the reference to Lemma 4.1.1 a) and Corollary 4.3.10 by a reference to Theorem 4.2.2. We leave this to the reader as an exercise.

We state two applications of Theorem 5.2.1. The first of them deals with finite p-groups of maximal class, that is, p-groups of order p^n and of nilpotency class $n-1$. Without plunging into the theory of such groups we merely point out that any p-group P of maximal class has an element s such that $|C_P(s)| = p^2$ (see [10] or [48, Chapter III]). It is clear that $s^p \in Z(P)$ and hence s induces an automorphism of P of order p by conjugation. We may therefore apply Theorem 5.2.1.

5.2.5 Corollary. *Every finite p-group of maximal class contains a subgroup of p-bounded index which is nilpotent of class $\leq h(p)$.*

In the next chapter we shall also show that in any p-group of maximal class the commutator subgroup (which has index p^2) is nilpotent of p-bounded class. How-

ever the papers of Shepherd [131] and Leedham-Green and McKay [93] contain stronger results on p-groups of maximal class.

5.2.6 Corollary. *Any locally finite p-group in which there is a centralizer of an element of prime order p which is finite of order p^m, is almost nilpotent and satisfies the conclusion of Theorem 5.2.1.*

Proof. The group theoretical property "containing a subgroup of index $\leq k$, which is nilpotent of class $\leq d$, for given k and d", may be expressed as a universal formula of the predicate calculus. All finitely generated subgroups of the group in the statement of the corollary which contain the given element x of order p with finite centralizer of order p^m, form a local covering. Each of these subgroups is a finite p-group which admits an automorphism of prime order, induced by x, with exactly p^m fixed points. Thus each subgroup in the local covering satisfies Theorem 5.2.1, and therefore the whole group also satisfies this property by Mal'cev's Local Theorem.

We note that, by a theorem of Blackburn [11], a group satisfying the hypothesis of Corollary 5.2.6, even has an abelian subgroup of finite index. The essence of Corollary 5.2.6 lies in the bounds obtained from Theorem 5.2.1, because there is no bound on the index of an abelian subgroup or even of a nilpotent subgroup of class c, where c is a constant not dependent either on p or m. This is shown by the following example.

5.2.7 Example. Let p be a prime greater than 2 and let ω be a primitive p-th root of unity over $\mathbb{Z}$. We denote by $\mathbb{Z}_{p^s}[\omega]$ the factor-ring of the ring $\mathbb{Z}[\omega]$ over the ideal $p^s\mathbb{Z}[\omega]$, and by U we denote the free $(p-1)$-generated $\mathbb{Z}_{p^s}[\omega]$-module. The group of matrices of the form

$$A = \left\{ \left[\begin{array}{ccccc} 1 & a_1 & a_2 & \dots & a_{p-2} \\ & 1 & a_1 & \ddots & \vdots \\ & & & \ddots & a_2 \\ & & \cdot & \cdot & a_1 \\ \mathrm{O} & & & & 1 \end{array} \right] \middle| \; a_i \in \mathbb{Z}_{p^s}[\omega] \right\}$$

is commutative and acts on U in the natural way. The element

$$\varphi = \left[\begin{array}{ccc} \omega & & \\ & \omega^2 & \mathrm{O} \\ \mathrm{O} & \ddots & \\ & & \omega^{p-1} \end{array} \right]$$

also acts on U and, by conjugation, on A. It may therefore be considered as an automorphism of the semidirect product $P = U \leftthreetimes A$.

It is easy to compute that for $k \not\equiv 0 \pmod{p}$ there are exactly p elements $b \in \mathbb{Z}_{p^s}[\omega]$ such that $(\omega^k - 1)b = 0$. From this, elementary calculations yield the order of the centralizer

$$|C_P(\varphi)| = |C_U(\varphi)| \cdot |C_A(\varphi)| = p^{p-1} \cdot p^{p-2} = p^{2p-3}.$$

This number does not depend on s. At the same time the minimal index of a subgroup of P, which is nilpotent of class exactly $p - 1$, tends to infinity with s.

§ 5.3 Nilpotent groups with an almost regular automorphism of prime order

In this section we prove a generalization of Higman's Theorem 5.1.1 on nilpotent group with regular automorphisms of prime order to the case where the number of fixed points is finite: from this "almost regularity" of the automorphism follows the "almost nilpotency" of the group – the existence of a subgroup which has index bounded in terms of the number of fixed points and the order of the automorphism, and which is nilpotent of class bounded in terms of the order of the automorphism only. This is first of all proved for a periodic nilpotent group with an almost regular automorphism of prime order p, where by Theorem 5.2.1 it suffices to consider the case of a p'-group. The result is then extended to the general case (this extension is due to Medvedev [111]). This extension, however, leads to worse bounds on the index and on the nilpotency class (bounds which were in any case far from best possible).

On the other hand Makarenko [103,104] showed that in this situation there is a (p, q)-bounded number s (where q is the number of fixed points), such that the subgroup generated by all s-th powers is nilpotent of class $\leq h(p)$ – and this bound for the nilpotency class cannot, of course, be improved.

5.3.1 Theorem. *If a periodic nilpotent group G admits an automorphism φ of prime order p having a finite number q of fixed points, then it has a subgroup of (p, q)-bounded index which is nilpotent of p-bounded class.*

This group-theoretic result, which is analogous to Theorem 4.4.1 on Lie rings, cannot, however, be derived from the latter as easily as in the case of a regular automorphism. Certainly, by Theorem 5.2.1, the group G may be assumed to be a p'-group, so that the induced automorphism φ of the associated Lie ring $L(G)$ has the same number q of fixed points. By Theorem 4.4.1 the Lie ring $L(G)$ contains

a subring of (p, q)-bounded index, which is nilpotent of p-bounded class. But this says nothing about the nilpotency class of any of the subgroups of G since there is no good correspondence between subgroups of G and subrings of the associated Lie ring (just like an upper bound on the derived length of the Lie ring $L(G)$ says nothing about the derived length of the group G).

We note also that we could not prove that under the hypothesis of Theorem 5.3.1 the group necessarily contains a subgroup of (p, q)-bounded index with a regular automorphism of order p; apparently, this seems to be wrong.

The proof of the theorem exploits the method of constructing generalized centralizers which was developed in the course of the proof of Theorem 4.4.1. In an analogous way we construct for G (via its associated Lie ring $L(G)$, extended by ω – a primitive p-th root of unity) representatives and augmented subgroups of increasing levels which have the centralizing property. In order to do this we need to translate properties of the elements of the Lie ring over $\mathbb{Z}[\omega]$ into group theoretic language. The main idea which enables us in the end to use the technique developed in the proof of Theorem 4.4.1, is to prove that G, rather than a subgroup of G, is itself nilpotent of p-bounded class. The advantage of such an approach lies in the fact that, in order to prove the nilpotency of the group itself, it suffices to prove that its associated Lie ring is nilpotent of the required class. However, the disadvantage of this approach lies in the fact that such an assertion is false!

This big "disadvantage" is overcome by induction on a complex parameter which controls the possibility of representing commutators in φ-homogeneous elements of the Lie ring as commutators in representatives of higher levels (in the language of the proof of Theorem 4.4.1). If this parameter is smaller for any subgroup of (p, q)-bounded index arising in the construction, then the induction hypothesis may be applied, and if it remains constant for long enough, then we can prove that $L(G)$ is nilpotent of p-bounded class, and, therefore, that G is nilpotent of the same class.

We move on now to the details.

Proof of Theorem 5.3.1. Since G is nilpotent it is the direct product of its Sylow subgroups; in particular, $G = G_p \times G_{p'}$, where G_p is a p-group, and $G_{p'}$ is a p'-group, both normal Hall subgroups being φ-invariant. Therefore,

$$C_G(\varphi) = C_{G_p}(\varphi) \times C_{G_{p'}}(\varphi).$$

We may apply Theorem 5.2.1 to the p-group G_p, so that it is sufficient to prove Theorem 5.3.1 in the case where G is a nilpotent p'-group which we shall from now on assume. We note in particular that for every $k \in \mathbb{N}$, $g \in G$ there exists a unique element $u \in G$, such that $u^{p^k} = g$ (in fact, $u = g^j$ for some $j \in \mathbb{N}$, since the mapping $g \to g^{p^k}$ is an automorphism of the cyclic group $\langle g \rangle$; uniqueness

follows, for instance, from 2.6.1 a)). The same is true for any abelian section of G, where we shall use additive notation to denote $u = (1/p^k)g$.

We recall that $h = h(p)$ denotes Higman's function.

We shall construct a decreasing series of subgroups $K(s)$, the generalized centralizers of levels $s \leq 2h + 1$, which have (p, q)-bounded indices. We simultaneously shall compute for them the value of a certain induction parameter. If, for any of the subgroups $K(r)$, this value is less than that for G itself, then, by the induction hypothesis, $K(r)$, and hence also G, contains the required subgroup of (p, q)-bounded index, which is nilpotent of p-bounded class. If, however, this parameter remains constant up to level $2h + 1$, then we shall prove that in this case the group itself is nilpotent of class $< f(h + 1, 2h, p)$, where f is the function appearing in the statement of Proposition 4.4.2. The same bound $f(h + 1, 2h, p)$ will also fit as a bound on the nilpotency class of a subgroup of (p, q)-bounded index in the conclusion of the theorem, because it may be incorporated in the induction hypothesis. For the rest of the proof we put $N = f(h + 1, 2h, p)$.

Let $L(G)$ be the associated Lie ring of G. Since G is a periodic p'-group, we have $|C_{L(G)}(\varphi)| = |C_G(\varphi)|$ by Theorem 1.6.2.

We put $L = L(G) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$, where ω is a primitive p-th root of unity, and we regard the Lie ring $L(G)$ as embedded in L by the rule $l \to l \otimes 1$.

Since G is a periodic p'-group, we have

$$pL = L = {}^0L \oplus {}^1L \oplus \ldots \oplus {}^{p-1}L,$$

where the ${}^iL$ are the φ-components of L in the sense of § 4.1. By contrast with the proof of Theorem 4.4.1 we shall only consider here commutators in elements of L which are homogeneous of weight 1, that is, elements of the form $\bar{x} \otimes r$, where $r \in \mathbb{Z}[\omega]$, and $\bar{x}$ is the image of $x \in G$ in $G/\gamma_2(G)$. We recall that $\bar{x} = {}^0x + {}^1x + \ldots + {}^{p-1}x$, where for every i

$${}^ix = \frac{1}{p} \sum_{s=0}^{p-1} \omega^{-is} \cdot \bar{x}^{\varphi^s} \in {}^iL$$

(see § 4.1). This notation will remain fixed for what follows and we shall say that these elements ${}^ix \in {}^iL$ are φ-components of the image $\bar{x}$ of $x \in G$ in $G/\gamma_2(G)$, which latter group is viewed as a homogeneous component of weight 1 of the Lie ring $L(G)$.

We shall construct the generalized centralizers $K(s)$, simultaneously fixing the representatives following the analogy of the proof of Theorem 4.4.1. The trouble is that here the generalized centralizers must be subgroups of G, while the centralizing property is naturally defined within the Lie ring $L = L(G) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$. This is why we have to translate this property into group theoretical terms and it is also why

along with representatives in the φ-components of the Lie ring we here also fix corresponding elements in G. This translation is done using the following lemma which gives a certain sufficient condition (there may very well be also a similar necessary condition, but we do not need it).

5.3.2 Lemma. *Suppose that the following congruence holds:*

$$j + i_1 + i_2 + \ldots + i_k \equiv 0 \pmod{p},$$

for some k, where $j, i_1, i_2, \ldots, i_k$ are residues modulo p. Then in order to have

$$[{}^j y, {}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_k}x_k] = 0,$$

in the Lie ring L for φ-homogeneous elements ${}^j y \in {}^j L$, ${}^{i_s}x_s \in {}^{i_s}L$, $s = 1, 2, \ldots, k$, which are homogeneous of weight 1, it is sufficient that, for every ordered set a_1, $a_2, \ldots, a_k$ of residues modulo p, the following congruence holds in G:

$$\prod_{l=0}^{p-1} \left(\left[y, x_1^{\varphi^{a_1}}, x_2^{\varphi^{a_2}}, \ldots, x_k^{\varphi^{a_k}} \right] \right)^{\varphi^l} \equiv 1 \pmod{\gamma_{k+2}(G)}.$$

(Here, of course, the φ-homogeneous elements ${}^j y, {}^{i_1}x_1, \ldots, {}^{i_k}x_k$ of L are constructed from the corresponding group elements $y, x_1, \ldots, x_k$ in the above-mentioned sense.)

Proof. We substitute the expressions of the ${}^j y, {}^{i_s}x_s$:

$$[{}^j y, {}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_k}x_k] =$$

$$= \left[\frac{1}{p} \sum_{s=0}^{p-1} \omega^{-js} \cdot \bar{y}^{\varphi^s}, \quad \frac{1}{p} \sum_{s=0}^{p-1} \omega^{-i_1 s} \cdot \bar{x}_1^{\varphi^s}, \quad \frac{1}{p} \sum_{s=0}^{p-1} \omega^{-i_2 s} \cdot \bar{x}_2^{\varphi^s}, \ldots, \frac{1}{p} \sum_{s=0}^{p-1} \omega^{-i_k s} \cdot \bar{x}_k^{\varphi^s} \right]$$

$$= \frac{1}{p^{k+1}} \sum_{l=0}^{p-1} \omega^l \sum_{\substack{-js_0 - i_1 s_1 - \ldots - i_k s_k \equiv l \pmod{p} \\ 0 \le s_i \le p-1}} \left[\bar{y}^{\varphi^{s_0}}, \bar{x}_1^{\varphi^{s_1}}, \bar{x}_2^{\varphi^{s_2}}, \ldots, \bar{x}_k^{\varphi^{s_k}} \right],$$

where the $\bar{y}$, $\bar{x}_i$ are the images of y, $x_i \in G$ in $G/\gamma_2(G)$, regarded as elements of L. Since $j + i_1 + \ldots + i_k \equiv 0 \pmod{p}$, the summation range in the inner sum on the right-hand side may be written in the form

$$i_1(s_0 - s_1) + i_2(s_0 - s_2) + \ldots + i_k(s_0 - s_k) \equiv l \pmod{p}.$$

Therefore, as it is easy to see, this inner sum is partitioned into several sums of the form

$$\sum_{t=0}^{p-1} \left(\left[\bar{y}, \bar{x}_1^{\varphi^{a_1}}, \bar{x}_2^{\varphi^{a_2}}, \ldots, \bar{x}_k^{\varphi^{a_k}} \right] \right)^{\varphi^t}.$$

By the definition of multiplication in the Lie ring $L(G)$ these sums are equal to 0 in L if and only if the corresponding congruences in the statement of the lemma hold.

The lemma is proved.

We now need homomorphisms, analogous to those used in the proof of Theorem 4.4.1, but this time defined for the group G.

For every ordered set $\bar{x} = (x_1, x_2, \ldots, x_k)$ of elements of G (of size k) and for every ordered set $\bar{a} = (a_1, a_2, \ldots, a_k)$ of residues modulo p, we define a homomorphism

$$\vartheta(\bar{x}, \bar{a})\colon\ y \to \prod_{t=0}^{p-1} \left(\left[y, x_1^{\varphi^{a_1}}, x_2^{\varphi^{a_2}}, \ldots, x_k^{\varphi^{a_k}} \right] \right)^{\varphi^t} \cdot \gamma_{k+2}(G)$$

of G into $\gamma_{k+1}(G)/\gamma_{k+2}(G)$ (the product on the right-hand side is a product of images of commutators of weight $k+1$).

5.3.3 Lemma. *For any $\bar{x}$ and $\bar{a}$ we have $|G : \operatorname{Ker} \vartheta(\bar{x}, \bar{a})| \leq q$.*

Proof. The image of an element y under the homomorphism $\vartheta(\bar{x}, \bar{a})$ is equal to the product of commuting elements of an orbit under the action of the automorphism φ on the abelian group $\gamma_{k+1}(G)/\gamma_{k+2}(G)$. Hence this image belongs to $C_{L(G)}(\varphi)$ and so

$$|G : \operatorname{Ker} \vartheta(\bar{x}, \bar{a})| \leq |C_{L(G)}(\varphi)| = |C_G(\varphi)| = q.$$

The lemma is proved.

Associated with the ordered set $\bar{x} = (x_1, x_2, \ldots, x_k)$ as above, we now define the subgroup

$$K(\bar{x}) = \bigcap_{s=0}^{p-1} \bigcap_{\bar{a}} (\operatorname{Ker} \vartheta(\bar{x}, \bar{a}))^{\varphi^s},$$

where $\bar{a}$ runs through all ordered sets $\bar{a} = (a_1, a_2, \ldots, a_k)$ of residues modulo p of size k. The total number of subgroups involved in this intersection is obviously (p, k)-bounded, and the index of each of them is at most q by Lemma 5.3.3. Hence, the index $|G : K(\bar{x})|$ is also (p, q, k)-bounded. (We recall that the index of

the intersection of subgroups is less than or equal to the product of their indices.) Moreover, $K(\bar{x})$ is φ-invariant by construction.

From the definition we see that $K(\bar{x})$ has the following centralizing property relative to $\bar{x} = (x_1, x_2, \ldots, x_k)$: for any $y \in K(\bar{x})$

$$[{}^{j}y, {}^{i_1}x_1, {}^{i_2}x_2, \ldots, {}^{i_k}x_k] = 0 \tag{5.3.4}$$

whenever $j + i_1 + i_2 + \ldots + i_k \equiv 0 \pmod{p}$.

We recall that the pattern of a commutator in φ-homogeneous elements ${}^{i_s}x_s \in {}^{i_s}L$ is its bracket structure together with the arrangement of upper indices in it, and that such a commutator is called the value of its pattern at the elements ${}^{i_s}x_s$. From now on we restrict our attention to commutators only in φ-homogeneous elements which are homogeneous of weight 1.

We turn now to the promised inductive construction of generalized centralizers of levels $i \leq 2h + 1$, that is, subgroups $K(i)$ of G of (p, q)-bounded index, and to the simultaneous specification of representatives. We put $K(1) = G$.

For every homogeneous element $c \in C_{L(G)}(\varphi)$ of weight $\leq N$ we express c (at level 1) as the values of all possible patterns $\mathbf{p}$ of weight $\leq N = f(h + 1, 2h, p)$ in φ-homogeneous elements which are homogeneous of weight 1 of the form ${}^{j}y(1) = r \cdot {}^{j}x(1)$, where $r \in \mathbb{Z}[\omega]$ and ${}^{j}x(1)$ is a φ-component of the image of $x(1) \in G$ in $G/\gamma_2(G)$ (of course, we only do this for such pairs $(c, \mathbf{p})$ for which such an expression exists). The level is indicated in parentheses and both ${}^{j}y(1) \in L$ and $x(1) \in G$ are called *representatives of level* 1. It is clear that the total number of representatives of level 1 is (p, q)-bounded since it is clear that the total number of patterns of weight $\leq N = f(h + 1, 2h, p)$ is p-bounded, and the number of homogeneous elements $c \in C_{L(G)}(\varphi)$ is at most q.

Now we define the subgroup

$$K(2) = \bigcap_{\bar{x}(1)} K(\bar{x}(1))$$

where $\bar{x}(1) = (x_1(1), x_2(1), \ldots, x_k(1))$ runs through all ordered sets of size k for all $k \leq N$, consisting of representatives of level 1. Note that $K(2)$ is φ-invariant since all $K(\bar{x})$ are φ-invariant. The index of $K(2)$ is (p, q)-bounded since the indices of the subgroups $K(\bar{x})$ and the number of them are (p, q)-bounded. The subgroup $K(2)$ is a generalized centralizer relative to the representatives of level 1 in the following sense: for every $z \in K(2)$

$$[{}^{j}z, {}^{i_1}y_1(1), {}^{i_2}y_2(1), \ldots, {}^{i_k}y_k(1)] = 0$$

whenever $j + i_1 + i_2 + \ldots + i_k \equiv 0 \pmod{p}$ and $k \leq N$. This follows from (5.3.4).

Now suppose that we have already constructed subgroups $K(1) \geq K(2) \geq \ldots \geq K(s)$ for some $s < 2h+1$ and, for every $i \leq s$, have fixed representatives of level i of the form ${}^{j}y(i) = r \cdot {}^{j}x(i)$, where $r \in \mathbb{Z}[\omega]$ and ${}^{j}x(i)$ is a φ-component of the image of $x(i) \in K(i)$ in $G/\gamma_2(G)$. We next define the subgroup

$$K(s+1) = \bigcap_{\bar{x}(s)} K(\bar{x}(s)),$$

where $\bar{x}(s) = (x_1(\varepsilon_1), x_2(\varepsilon_2), \ldots, x_k(\varepsilon_k))$ runs through all ordered sets of sizes $k \leq N$, consisting of representatives of levels $\varepsilon_i \leq s$. It is clear that $K(s+1) \leq K(s)$. Also $K(s+1)$ is φ-invariant, since the $K(\bar{x})$ are φ-invariant, and its index is (p,q)-bounded, since the indices and the number of the $K(\bar{x})$ are (p,q)-bounded. The subgroup $K(s+1)$ is a generalized centralizer relative to the representatives of level $\leq s$ in the following sense: for every $z \in K(s+1)$

$$[{}^{j}z, {}^{i_1}y_1(\varepsilon_1), {}^{i_2}y_2(\varepsilon_2), \ldots, {}^{i_k}y_k(\varepsilon_k)] = 0 \tag{5.3.5}$$

whenever $\varepsilon_i \leq s$ for all i, $j + i_1 + \ldots + i_k \equiv 0 \pmod{p}$ and $k \leq N$. This follows from (5.3.4).

For every homogeneous element $c \in C_{L(G)}(\varphi)$ of weight $\leq N$ we express c at level $s+1$ as the values of all possible patterns $\mathbf{p}$ of weight $\leq N = f(h+1, 2h, p)$ at φ-homogeneous elements which are homogeneous of weight 1 of the form ${}^{j}y(s+1) = r \cdot {}^{j}x(s+1)$, where $r \in \mathbb{Z}[\omega]$, and ${}^{j}x(s+1)$ is a φ-component of the image of $x(s+1) \in K(s+1)$ in $G/\gamma_2(G)$ (again, for such pairs $(c, \mathbf{p})$ for which such an expression exists). Both ${}^{j}y(s+1) \in L$ and $x(s+1) \in G$ are called *representatives of level* $s+1$. It is clear that the total number of representatives of level $s+1$ is (p,q)-bounded since it is clear that the total number of patterns of weight $\leq N = f(h+1, 2h, p)$ is p-bounded, and the number of homogeneous elements $c \in C_{L(G)}(\varphi)$ is at most q.

This completes the inductive definition of the generalized centralizers $K(i)$ and the representatives of levels $i \leq 2h+1$.

We now make some comments on the differences between this situation and that considered in the proof of Theorem 4.4.1 and on ways of overcoming the difficulties which arise. As we have already said, our objective is to prove that the group G itself is nilpotent of p-bounded class $< N$, or, equivalently, that the same is true for the Lie ring $L(G)$. In order to do this it is sufficient to show that every commutator of weight N in the φ-homogeneous elements of L which are homogeneous of weight 1 is equal to 0. To achieve this it is in turn enough, by Proposition 4.4.2, to consider the commutators mentioned in the statement of Proposition 4.4.2 which involve a large number of subcommutators from ${}^{0}L$. It would then be natural to try to apply to them the collecting process described at the end of the proof of Theorem 4.4.1. However, that process gave the desired

result there because it produced commutators with long enough initial segments filled by representatives (or quasirepresentatives) of different levels. In the proof of Theorem 4.4.1 the existence of the necessary reserve of representatives of different levels followed automatically from the fact that we were considering commutators in elements of maximal level and could replace subcommutators from ${}^0L$ by the values of the same patterns at elements of lower levels (because the subgroups of higher levels are contained in the subgroups of lower levels). But here we have to deal with commutators in arbitrary φ-homogeneous elements of L which are homogeneous of weight 1, and we can replace the subcommutators from ${}^0L$ by values of the same patterns at representatives of the 1-st level only.

This difficulty is overcome by introducing an induction parameter which controls the possibility of replacing commutators from ${}^0L$ by values of the same patterns at representatives of high levels. The proof of the theorem then proceeds by induction on this parameter. This gives a reduction to the case, where the substitutions mentioned above are possible. And in this case we only need repeat almost word by word, with little technical modification, the arguments from the end of the proof of Theorem 4.4.1.

Definition. The *induction parameter* is a triplet $(q, \bar{q}, t)$, where

$$q = |C_G(\varphi)|, \quad \bar{q} = (q_1, q_2, \ldots, q_N)$$
$$q_i = |C_{\gamma_i(G)/\gamma_{i+1}(G)}(\varphi)|, \quad i = 1, 2, \ldots, N, \quad t = |\mathbf{P}(\mathrm{G})|,$$

where $\mathbf{P}(G)$ is the number of pairs $(c, \mathbf{p})$, where c is an element of $C_{L(G)}(\varphi)$ which may be represented as a value of the pattern $\mathbf{p}$ of weight $\leq N$ at φ-homogeneous elements of L which are homogeneous of weight 1.

By $(q(H), \bar{q}(H), t(H))$ we shall mean a similarly constructed triplet, formed relative to the φ-invariant subgroup H (in particular, $(q, \bar{q}, t) = (q(G), \bar{q}(G), t(G))$), and $\mathbf{P}(H)$ we shall use to denote the corresponding set of pairs $(c, \mathbf{p})$ for the subgroup H.

We order the set of vectors $\bar{q} = (q_1, q_2, \ldots, q_N)$ in an inverse lexicographic way so that

$$\bar{q}_1 = (q_{11}, q_{12}, \ldots, q_{1N}) < \bar{q}_2 = (q_{21}, q_{22}, \ldots, q_{2N}) \Leftrightarrow$$
$$\Leftrightarrow \text{ for some } k \geq 1, \ q_{1i} = q_{2i} \text{ for all } i < k \text{ and } q_{1k} > q_{2k}.$$

We also order the set of triplets $(q, \bar{q}, t)$ lexicographically:

$$(q_1, \bar{q}_1, t_1) < (q_2, \bar{q}_2, t_2) \Leftrightarrow \text{ either } q_1 < q_2,$$
$$\text{or } q_1 = q_2 \text{ and } \bar{q}_1 < \bar{q}_2,$$
$$\text{or } q_1 = q_2, \ \bar{q}_1 = \bar{q}_2 \text{ and } t_1 < t_2.$$

5.3.6 Lemma. *For any φ-invariant subgroup H we have*

$$(q(H), \bar{q}(H), t(H)) \leq (q(G), \bar{q}(G), t(G)).$$

Proof. It is clear that $q(H) \leq q(G)$, since $C_H(\varphi) \subseteq C_G(\varphi)$. Now let $q(H) = q(G)$, that is, $C_G(\varphi) \subseteq H$; we want to prove that in this case $\bar{q}(H) \leq \bar{q}(G)$. Suppose that for some $k \geq 1$ we have $q_i(H) = q_i(G)$ for all $i < k$; we now have to show that $q_k(H) \geq q_k(G)$. Since $|C_G(\varphi)| = |C_H(\varphi)|$ and $q_i(H) = q_i(G)$ for all $i < k$, then $|C_{\gamma_k(H)}(\varphi)| = |C_{\gamma_k(G)}(\varphi)|$. Obviously, $C_{\gamma_k(H)}(\varphi) \subseteq C_{\gamma_k(G)}(\varphi)$ so that these two subgroups are identical; we set $C = C_{\gamma_k(H)}(\varphi) = C_{\gamma_k(G)}(\varphi)$ for short. By Theorem 1.6.2 we have

$$C_{\gamma_k(H)/\gamma_{k+1}(H)}(\varphi) = C \cdot \gamma_{k+1}(H)/\gamma_{k+1}(H) \cong C/C \cap \gamma_{k+1}(H),$$

and also

$$C_{\gamma_k(G)/\gamma_{k+1}(G)}(\varphi) = C \cdot \gamma_{k+1}(G)/\gamma_{k+1}(G) \cong C/C \cap \gamma_{k+1}(G).$$

The order of $C/C \cap \gamma_{k+1}(H)$ is clearly not less than the order of $C/C \cap \gamma_{k+1}(G)$. Hence, $q_k(H) \geq q_k(G)$, as required.

Finally, suppose that $\bar{q}(H) = \bar{q}(G)$; we shall prove in this case that $t(H) \leq t(G)$. In order to do this we define an injective mapping from the set of pairs $\mathbf{P}(H)$, defining the number $t(H)$, into the corresponding set of pairs $\mathbf{P}(G)$ as follows.

Suppose that the homogeneous element $\bar{c} \in C_{L(H)}(\varphi)$ of weight k, $k \leq N$, is the image in $\gamma_k(H)/\gamma_{k+1}(H)$ of $c \in C_{\gamma_k(H)}(\varphi) \subseteq C_{\gamma_k(G)}(\varphi)$, and also that $\hat{c} \in C_{L(G)}(\varphi)$ is the image of c in $\gamma_k(G)/\gamma_{k+1}(G)$. We have shown that if $\bar{q}(H) = \bar{q}(G)$ then $C_G(\varphi) \cap \gamma_i(H) = C_G(\varphi) \cap \gamma_i(G)$ for all i. Therefore, by Theorem 1.6.2, we have

$$\begin{aligned} C_{\gamma_k(H)/\gamma_{k+1}(H)}(\varphi) &\cong (C_G(\varphi) \cap \gamma_k(H))/(C_G(\varphi) \cap \gamma_{k+1}(H)) = \\ &= (C_G(\varphi) \cap \gamma_k(G))/(C_G(\varphi) \cap \gamma_{k+1}(G)) \cong C_{\gamma_k(G)/\gamma_{k+1}(G)}(\varphi). \end{aligned}$$

Hence, the mapping $\mu\colon \bar{c} \to \hat{c}$ is an isomorphism of $C_{\gamma_k(H)/\gamma_{k+1}(H)}(\varphi)$ with $C_{\gamma_k(G)/\gamma_{k+1}(G)}(\varphi)$.

Next, suppose that $\bar{c}$ is equal to the value of the pattern $\mathbf{p}$ of weight $k \leq N$ at φ-homogeneous elements which are homogeneous of weight 1 of the Lie ring $L(H) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$, that is, let $(c, \mathbf{p}) \in \mathbf{P}(H)$. We define the mapping

$$\nu\colon (\bar{c}, \mathbf{p}) \to (\hat{c}, \mathbf{p}).$$

The following lemma shows that if $\bar{q}(H) = \bar{q}(G)$ then $(\hat{c}, \mathbf{p})$ belongs to the set $\mathbf{P}(G)$.

5.3.7 Lemma. *Suppose that $\bar{q}(H) = \bar{q}(G)$ where H is a φ-invariant subgroup of G and that the homogeneous element $\bar{c} \in C_{L(H)}(\varphi)$ of weight k, $k \leq N$, which is the image in $\gamma_k(H)/\gamma_{k+1}(H)$ of the element $c \in C_G(\varphi)$, is equal to the value of the pattern* $\mathbf{p}$ *of weight k at the φ-homogeneous elements ${}^{j_i}y_i = r_i \cdot {}^{j_i}x_i$ which are homogeneous of weight* 1 *in the Lie ring $L(H) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$, where $r_i \in \mathbb{Z}[\omega]$ and the ${}^{j_i}x_i$ are the φ-components of the images in H/H' of the elements $x_i \in H$. Then the image $\hat{c} \in C_{L(G)}(\varphi)$ of c in $\gamma_k(G)/\gamma_{k+1}(G)$ is equal to the value of the same pattern* $\mathbf{p}$ *at the φ-homogeneous elements ${}^{j_i}\hat{y}_i = r_i \cdot {}^{j_i}\hat{x}_i$ which are homogeneous of weight* 1 *in the Lie ring $L(G) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$, the coefficients $r_i \in \mathbb{Z}[\omega]$ being as above, and ${}^{j_i}\hat{x}_i$ being the φ-components of the images of the x_i in G/G'.*

Proof. Since $\mathbb{Z}[\omega] = \mathbb{Z} \oplus \omega\mathbb{Z} \oplus \ldots \oplus \omega^{p-2}\mathbb{Z}$, then

$$L(H) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega] = L(H) \oplus \omega L(H) \oplus \ldots \oplus \omega^{p-2} L(H)$$

(see § 1.2). Hence two elements in $L(H) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ are equal if and only if the corresponding coefficients of $1, \omega, \omega^2, \ldots, \omega^{p-2}$ are equal: these coefficients being elements of $L(H)$. By the definition of the associated Lie ring $L(H)$, two of its elements are equal if and only if their homogeneous constituents are equal. And two homogeneous elements of $L(H)$ are equal if and only if the corresponding products of powers of group commutators are congruent modulo the corresponding members of the lower central series. Analogous assertions are also valid for the Lie ring $L(G) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$.

Therefore, since the φ-components ${}^{j_i}x_i$ are expressed by fixed formulae in the images in H/H' of the x_i, the fact that in $L(H) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ the element $\bar{c} \in C_{L(H)}(\varphi)$ is equal to the value of the pattern $\mathbf{p}$ mentioned in the statement of the lemma, is equivalent to a system of congruences in group commutators c_α of weight k in the elements x_i

$$\begin{cases} c \equiv \prod_\alpha c_\alpha^{s(0,\alpha)} \pmod{\gamma_{k+1}(H)}; \\ 1 \equiv \prod_\alpha c_\alpha^{s(1,\alpha)} \pmod{\gamma_{k+1}(H)}; \\ 1 \equiv \prod_\alpha c_\alpha^{s(2,\alpha)} \pmod{\gamma_{k+1}(H)}; \\ \ldots \\ \ldots \\ \ldots \\ 1 \equiv \prod_\alpha c_\alpha^{s(p-2,\alpha)} \pmod{\gamma_{k+1}(H)} \end{cases}$$

(here the i-th congruence expresses the fact that the coefficients of ω^{i-1} are equal). It is also clear that the integers $s(i, \alpha)$ may be taken as depending only on $\mathbf{p}$ and on the $r_i \in \mathbb{Z}[\omega]$. But the same congruences clearly also hold modulo $\gamma_{k+1}(G)$ (which contains $\gamma_{k+1}(H)$). And modulo $\gamma_{k+1}(G)$ this system of congruences is equivalent

to the fact that the element $\hat{c}$ is equal in the Lie ring $L(G) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ to the value of $\mathbf{p}$.

The lemma is proved.

We return now to the proof of Lemma 5.3.6. By Lemma 5.3.7, if $\bar{q}(H) = \bar{q}(G)$, then $(\hat{c}, \mathbf{p}) = \nu(\bar{c}, \mathbf{p})$ belongs to $\mathbf{P}(G)$, the set which defines the number $t(G)$. We have noted above that in this case the mapping $\mu\colon \bar{c} \to \hat{c}$ is an isomorphism of $C_{\gamma_k(H)/\gamma_{k+1}(H)}(\varphi)$ with $C_{\gamma_k(G)/\gamma_{k+1}(G)}(\varphi)$. Therefore ν clearly takes distinct pairs to distinct pairs. Therefore $t(H) \leq t(G)$ if $\bar{q}(H) = \bar{q}(G)$.

Lemma 5.3.6 is proved.

It is important in what follows to note that this argument yields that if $(q(H), \bar{q}(H), t(H)) = (q(G), \bar{q}(G), t(G))$ then ν is in fact a one-to-one correspondence between the set $\mathbf{P}(G)$ of all pairs $(\hat{c}, \mathbf{p})$, defining the number $t(G)$, and the set $\mathbf{P}(H)$ of all pairs $(\bar{c}, \mathbf{p})$, defining the number $t(H)$. In particular it is not hard to see that in the language of the generalized centralizers $K(i)$ and representatives of levels $i = 1, 2, \ldots, 2h+1$ we have

5.3.8 Lemma. *Suppose that, for all levels $i = 1, 2, \ldots, 2h+1$, we have*

$$(q(K(i)), \bar{q}(K(i)), t(K(i))) = (q(G), \bar{q}(G), t(G)).$$

Then for each $s = 1, 2, \ldots, 2h+1$ each homogeneous element of $C_{L(G)}(\varphi)$, which can be represented as the value of a pattern $\mathbf{p}$ at φ-homogeneous elements (of level 1) which are homogeneous of weight 1, can be also represented as the value of $\mathbf{p}$ at representatives of level s, that is, elements of the form ${}^{j}y(s) = r \cdot {}^{j}x(s)$, where $r \in \mathbb{Z}[\omega]$ and ${}^{i}x(s)$ is a φ-component of the image of $x(s) \in K(s)$ in $G/\gamma_2(G)$.

Now everything is ready for the completion of the proof of Theorem 5.3.1. We note that for a given value q of $q(G)$ the number of possible triplets $(q(G), \bar{q}(G), t(G))$ is clearly (p, q)-bounded. Thus for the proof of the theorem it is sufficient to show by induction on the parameter $(q(G), \bar{q}(G), t(G))$, that G contains a subgroup of (p, q)-bounded index which is nilpotent of class $< N = f(h+1, 2h, p)$. At the first induction step we have the case $q(G) = q = 1$, which means that φ is a regular automorphism. Then G is nilpotent of class $\leq h < N$ by Higman's Theorem 5.1.1.

If, for some $i = 1, 2, \ldots, 2h+1$, the induction parameter for $K(i)$ is smaller than that for G itself, that is, if

$$(q(K(i)), \bar{q}(K(i)), t(K(i))) < (q(G), \bar{q}(G), t(G))$$

then, by the induction hypothesis applied to the φ-invariant subgroup $K(i)$, the group G contains a nilpotent subgroup of class $< N$ which has (p, q)-bounded

index in $K(i)$. This subgroup is what we want, because the index of $K(i)$ in G is also (p, q)-bounded.

It therefore suffices to consider the case where, for each $i = 1, 2, \ldots, 2h + 1$,

$$(q(K(i)), \bar{q}(K(i)), t(K(i))) = (q(G), \bar{q}(G), t(G))$$

and this will be assumed in what follows. We shall prove that in this critical situation G itself is nilpotent of class $< N$. As is well known, in order to establish this, it is sufficient to show that the Lie ring $L = L(G) \otimes_{\mathbb{Z}} \mathbb{Z}[\omega]$ is nilpotent of class $< N$ (see Theorem 3.2.2 and § 1.3).

We note that we may also assume that $C_G(\varphi) \leq G'$. Otherwise $C_{[G,\varphi]}(\varphi) < C_G(\varphi)$, since, by Corollary 1.6.5, we have $[G, \varphi] \cap C_G(\varphi) \leq G'$, and the proof is completed by the induction hypothesis applied to the φ-invariant subgroup $[G, \varphi]$, whose index is at most q, because $G = [G, \varphi] \cdot C_G(\varphi)$ by Corollary 1.6.4. Hence we also have $C_L(\varphi) \leq \gamma_2(L)$, and L is generated by φ-components ${}^iL$ with $i \neq 0$. Since we need only verify the nilpotency identity on the generators of the Lie ring (Theorem 3.1.2), in order to prove that the Lie ring L is nilpotent of class $< N$, it is sufficient to show that

$$[{}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_N}y_N] = 0 \tag{5.3.9}$$

for any φ-homogeneous elements ${}^{i_s}y_s \in {}^{i_s}L$ with $i_s \neq 0$ which are homogeneous of weight 1.

In order to do this, according to Proposition 4.4.2, it is sufficient, in turn, to show that all commutators of the forms (4.4.3) and (4.4.4) from Proposition 4.4.2, applied to the commutator (5.3.9) with $m = h + 1$ and $n = 2h$, are 0 (we recall that $N = f(h + 1, 2h, p)$, where f is the function in Proposition 4.4.2).

The following argument essentially reproduces the end of the proof of Theorem 4.4.1 on replacing the centralizer property for representatives in the old sense by the centralizer property for representatives in the new sense. That this is possible is due to Lemma 5.3.8, which may be applied in the critical situation under consideration to guarantee the representability of commutators from ${}^0L$ in φ-homogeneous elements of level 1 which are homogeneous of weight 1 as values of the same patterns at representatives of higher levels. A small modification, involving enlarging the parameter n appearing in Proposition 4.4.2 from h to $2h$, is necessary only for commutator of the form (4.4.4).

First we consider the commutator

$$[{}^{k_1}w_1, {}^{k_2}w_2, \ldots, {}^{k_t}w_t], \quad {}^{k_i}w_i \in \{{}^{i_1}y_1, {}^{i_2}y_2, \ldots, {}^{i_N}y_N\}, \tag{5.3.10}$$

of the form (4.4.3) from Proposition 4.4.2 applied to commutator (5.3.9) with $m = h + 1$ and $n = 2h$. This commutator has $h + 1$ different initial segments with

upper index sums zero modulo p:

$$k_1 + k_2 + \ldots + k_{r_i} \equiv 0 \pmod{p}, \quad i = 1, 2, \ldots, h+1$$
$$1 < r_1 < r_2 < \ldots < r_{h+1} = t.$$

Since, for all $i = 1, 2, \ldots, 2h+1$, we have

$$(q(K(i)), \bar{q}(K(i)), t(K(i))) = (q(G), \bar{q}(G), t(G)),$$

then, by Lemma 5.3.8, the commutator (5.3.10) may be represented as a value of the same pattern at representatives of level $h+1$. Since $K(i) \geq K(i+1)$ for all i, we may then in turn represent the initial segment of weight r_h of the resultant commutator as a value of the same pattern at representatives of level h, then the initial segment of weight r_{h-1} of the resultant commutator as a value of the same pattern at representatives of level $h-1$, and so on (all these initial segments lie in ${}^0L$, as does the commutator (5.3.10) itself). As a result we obtain a commutator which is equal to (5.3.10) and has the form

$$\left[{}^{k_1}y(1), {}^{k_2}y(1), \ldots, {}^{k_r}y(1), {}^{k_{r+1}}y(2), \ldots, {}^{k_s}y(2), \ldots, {}^{k_{u-1}}y(h+1), \ldots, {}^{k_t}y(h+1)\right],$$

where the lower indices are omitted to simplify notation.

We may now apply to this commutator the arguments from §4.4 which were used to prove that (4.4.25) is equal to 0. The only modification needed is to replace the centralizer property (4.4.14) by (5.3.5), and the subgroups ${}^jK(s)$ – by the sets $\{{}^jy \mid {}^jy = r \cdot {}^jx, \ \ x \in K(s), \ \ r \in \mathbb{Z}[\omega]\}$.

We now consider the commutator

$$[{}^jy, c_1, c_2, \ldots, c_{2h}] \tag{5.3.11}$$

of the form (4.4.4). For each $s = 1, 2, \ldots, 2h$, we replace $c_s \in {}^0L$ in (5.3.11) by its expression as the value of the same pattern of weight $< N$ at the representatives of level s which were specified above. This is made possible by Lemma 5.3.8, since c_s is the value of some pattern of weight $< N$ at elements (of level 1) from ${}^iL, i \neq 0$. However, once we have expanded the inner brackets, we cannot directly apply the arguments which were used to prove that the commutator (4.4.17) is equal to 0, to every commutator of the form

$$\left[{}^jy, {}^{i_1}x(1), \ldots, {}^{i_k}x(1), {}^{i_{k+1}}x(2), \ldots, {}^{i_l}x(2), \ldots, {}^{i_{l+1}}x(2h), \ldots, {}^{i_u}x(2h)\right]$$

of the resultant linear combination. Indeed, the initial element ${}^jy$ here does not possess any *a priori* centralizer property, that is, it neither belongs to the analogue of a centralizer of high level, nor it is a representative of high level.

Nevertheless, the arguments which we applied to (4.4.17), allow us to represent the initial segment

$$\left[{}^{j}y, {}^{i_1}x(1), \ldots, {}^{i_k}x(1), {}^{i_{k+1}}x(2), \ldots, {}^{i_l}x(2), \ldots, {}^{i_{m-1}}x(h)\right]$$

as a linear combination of commutators of the form

$$\left[{}^{j}y, {}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_h}\hat{x}(h), {}^{v_{h+1}}\hat{x}(\varepsilon_{h+1}), \ldots\right],$$

where the initial segment of weight $h+1$, beginning with ${}^{j}y$, contains exactly one quasirepresentative of each of the levels $1, 2, \ldots, h$ (in the order of increasing level).

By the Higman-Kreknin-Kostrikin Theorem 4.3.10, the initial segment of weight $h+1$

$$\left[{}^{j}y, {}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_h}\hat{x}(h)\right]$$

of each of these commutators may be expressed as a linear combination of simple commutators in ${}^{j}y, {}^{v_1}\hat{x}(1), {}^{v_2}\hat{x}(2), \ldots, {}^{v_h}\hat{x}(h)$ with initial segments from ${}^{0}L$. If such an initial segment from ${}^{0}L$ does not contain ${}^{j}y$, then it is 0 for the same reasons as used for the commutator (4.4.27). If, however, such an initial segment from ${}^{0}L$ contains ${}^{j}y$, then we can replace it by a value of the same pattern at representatives of level 1. We now need only prove that

$$\begin{aligned}\Big[{}^{k_1}y(1), \ldots, {}^{k_r}y(1), {}^{k_{r+1}}y(\varepsilon_{r+1}), \ldots, {}^{k_s}y(\varepsilon_s), {}^{i_{a+1}}x(h+1), \ldots \\ \ldots, {}^{i_b}x(h+1), \ldots, {}^{i_{t+1}}x(2h), \ldots, {}^{i_u}x(2h)\Big]\end{aligned} \tag{5.3.12}$$

is 0. Note that this commutator involves only representatives, and all levels $\varepsilon_{r+1}, \ldots, \varepsilon_s$ are less than $h+1$. We apply essentially the same collecting process to it, as we did to (4.4.17). The only difference is that only the first quasirepresentatives from the left of levels $\geq h+1$ are transposed to the left, and the collected parts are the initial segments of the form

$$\left[{}^{k_1}y(1), {}^{v_1}\hat{x}(h+1), \ldots, {}^{v_s}\hat{x}(h+s)\right], \quad s \geq 0.$$

As a result the commutator (5.3.12) will be expressed as a linear combination of commutators of quasirepresentatives with initial segments of weight $h+1$ of the form

$$\left[{}^{k_1}y(1), {}^{v_1}\hat{x}(h+1), \ldots, {}^{v_h}\hat{x}(2h)\right].$$

Applying the Higman-Kreknin-Kostrikin Theorem to such an initial segment, we obtain a linear combination of commutators in quasirepresentatives of different

levels with initial segments from ${}^0L$, all of which are 0 by the same arguments as used in the case of (4.4.21).

The theorem is proved.

Under the hypothesis of Theorem 5.3.1 Makarenko recently obtained a somewhat different conclusion with best-possible bound for the nilpotency class. (She also extended this result to the case where the group is not necessarily periodic, but we prefer to deal first with the periodic case and to refer to the proof later, after the extension of Theorem 5.3.1 will be obtained.)

5.3.13 Corollary (Makarenko [103]). *Under the hypothesis of Theorem 5.3.1 there is a (p, q)-bounded number $s = s(p, q)$ such that the subgroup $G^s = \langle g^s \mid g \in G \rangle$ is nilpotent of class $\leq h(p)$.*

Proof. We note first that $G^{mn} \leq (G^m)^n$ for any $m, n \in \mathbb{N}$.

By Theorem 5.3.1, G has a subgroup G_1 of (p, q)-bounded index which is nilpotent of p-bounded class. By Poincaré's Theorem, G contains a normal subgroup G_2 of (p, q)-bounded index $k \leq (|G : G_1|)!$. By Lagrange's Theorem we have $G^k \leq G_2$, so that, since G^k is φ-invariant, in order to prove Corollary 5.3.13 we may assume from the outset that G is nilpotent of p-bounded class $c = c(p)$.

By an obvious induction on c it is sufficient to show that, if $c > h(p)$, then for some (p, q)-bounded natural number $r = r(p, q)$, the subgroup $G^r = \langle g^r \mid g \in G \rangle$ is nilpotent of class $\leq c - 1$.

By 4.1.1 a) and the Higman-Kreknin-Kostrikin Theorem 4.3.10, applied to the associated Lie ring $L(G)$ of G, we have

$$p^{h+1}\gamma_{h+1}(L(G)) = \gamma_{h+1}(pL(G)) \subseteq {}_{id}\langle C_{L(G)}(\varphi)\rangle.$$

As in the proof of Theorem 5.2.1 we obtain

$$qp^{h+1}\gamma_{h+1}(L(G)) \subseteq q_{id}\langle C_{L(G)}(\varphi)\rangle = 0.$$

In terms of the group G, this means that for $i \geq h + 1$ the exponent of $\gamma_i(G)/\gamma_{i+1}(G)$ divides qp^{h+1}. In particular, $\gamma_c(G)^{qp^{h+1}} = 1$, if $c > h$.

Since there is a homomorphism of $\underbrace{G/G' \otimes \ldots \otimes G/G'}_{c}$ onto $\gamma_c(G)$, we have

$$\left[g_1^{qp^{h+1}}, g_2^{qp^{h+1}}, \ldots, g_c^{qp^{h+1}}\right] = \left[g_1, g_2, \ldots, g_c\right]^{(qp^{h+1})^c} = 1$$

for any $g_i \in G$. Hence the subgroup $G^{qp^{h+1}}$, generated by the elements $g^{qp^{h+1}}$, is nilpotent of class $\leq c - 1$.

The corollary is proved.

An extension of Theorem 5.3.1 to arbitrary nilpotent groups admitting an almost regular automorphism of prime order, was obtained recently by Medvedev [111]. The following result is formally more general, but its proof makes use of Theorem 5.3.1 and consists in reduction (which is, though, by no means trivial) to the periodic case.

5.3.14 Theorem. *If a nilpotent group G admits an automorphism φ of prime order p having a finite number q of fixed points, then G has a subgroup of (p, q)-bounded index which is nilpotent of p-bounded class.*

Proof. The property of a group to have a nilpotent subgroup of given nilpotency class, whose index is bounded by a given number, may be written as a quasiuniversal formula in the predicate calculus. It therefore suffices to prove the theorem (with appropriate bounds for the index and the nilpotency class of a subgroup) for finitely generated groups. Indeed, the finitely generated subgroups $\{H_\alpha\}$ form a local covering of G. For each of these subgroups H_α, the subgroup $\langle {H_\alpha}^{\langle\varphi\rangle}\rangle$ is also finitely generated and φ-invariant. If the theorem holds for each $\langle {H_\alpha}^{\langle\varphi\rangle}\rangle$, then, by Mal'cev's Local Theorem, it also holds for G itself (see § 1.1).

We therefore suppose that G is a finitely generated nilpotent group with an automorphism φ of prime order p having a finite number $q = |C_G(\varphi)|$ of fixed points. We denote by $T = I(1)$ its periodic part, which is a finite subgroup by Proposition 2.5.10. Since T is the direct product $T = T_p \times T_{p'}$ of its Hall p- and p'-subgroups, by Remak's Theorem, G embeds in the direct product of G/T_p and $G/T_{p'}$. Hence, in order to prove Theorem 5.3.14, we may assume that T is either a p-group or a p'-group.

The following two lemmas hold in either case.

5.3.15 Lemma. *Under the hypothesis of Theorem 5.3.14 the automorphism φ induces a regular automorphism of the factor-group G/T.*

Proof. Assume that this is not so and let aT be a non-trivial element of $C_{G/T}(\varphi)$ that is, $a \notin T$ and $a^\varphi = at$ for some $t \in T$. We define the subgroup

$$H = \langle a, a^\varphi, \ldots, a^{\varphi^{p-1}} \rangle = \langle a, t, tt^\varphi, \ldots, tt^\varphi \cdot \ldots \cdot t^{\varphi^{p-1}} \rangle.$$

For every $i \in \mathbb{N}$ the factor-group $\gamma_i(H)/\gamma_{i+1}(H)$ is generated by the images of the commutators of weight i in the elements $a, t, tt^\varphi, \ldots, tt^\varphi \cdot \ldots \cdot t^{\varphi^{p-1}}$. Since there is a homomorphism of $\underbrace{G/G' \otimes \ldots \otimes G/G'}_{i}$ onto $\gamma_i(G)/\gamma_{i+1}(G)$, the images of these commutators, which for $i \geq 2$ are dependent on at least one of the elements

of finite order $t, tt^{\varphi}, \ldots, tt^{\varphi} \cdot \ldots \cdot t^{\varphi^{p-1}}$, also have finite order for $i \geq 2$. Hence, the factors of the lower central series of the group H, from the second one on, are finite since they are both finitely generated and have finite exponent.

Therefore, the commutator subgroup of H is finite. Then, by Theorem 2.4.4, the centre of H has finite index in H. Therefore, $a^m \in Z(G)$ for some $m \in \mathbb{Z}$. Since H is φ-invariant and the image of a^m also belongs to $C_{G/T}(\varphi)$, then $(a^m)^{\varphi} = a^m t_1$, where t_1 is an element of H of finite order n. We then have

$$(a^{mn})^{\varphi} = ((a^m)^{\varphi})^n = (a^m t_1)^n = a^{mn} t_1^n = a^{mn},$$

that is, a^{mn} belongs to $C_G(\varphi)$. But a^{mn} has infinite order, since $a \notin T$. This contradicts the finiteness of $C_G(\varphi)$.

The lemma is proved.

(Note that if we had at our disposal the useful Hall-Petresco formula, then we could have easily obtained that there is a natural m, large enough (relative to the nilpotency class of G and the order of t), such that

$$(a^m)^{\varphi} = (a^{\varphi})^m = (at)^m =$$
$$= a^m \cdot t^m \cdot [a, t]^{f_1(m)} \cdot [a, t, t]^{f_2(m)} \cdot [a, t, a]^{f_3(m)} \cdot \ldots = a^m.)$$

Applying Theorem 5.1.1 we obtain

5.3.16 Corollary. *The factor-group G/T is nilpotent of class $\leq h(p)$.*

We also have

5.3.17 Corollary. *If B is a φ-invariant normal periodic subgroup of G then $|C_{G/B}(\varphi)| \leq |C_G(\varphi)|$.*

Proof. The periodic part of G/B is clearly the image of the periodic part T of G. By Lemma 5.3.15 φ is regular on G/T so that $C_{G/B}(\varphi) = C_{T/B}(\varphi)$. Since T is finite, we have

$$|C_{T/B}(\varphi)| \leq |C_T(\varphi)| = |C_G(\varphi)|$$

by Theorem 1.6.1.

We now prove

5.3.18 Lemma. *Let N be a normal φ-invariant subgroup of the group G and let gN be an element of the centralizer $C_{G/N}(\varphi)$. Then for every natural k there exists*

$n \in N$ such that the image of $g^{p^k} n$ in the factor-group $G/[N, \underbrace{G, \ldots, G}_{k}]$ belongs to the centralizer $C_{G/[N, \underbrace{G, \ldots, G}_{k}]}(\varphi)$.

Proof. Induction on k. The case $k = 0$ is covered by the hypothesis. For $k > 0$ by the induction hypothesis there is an $n' \in N$ such that the image of $g_1 = g^{p^{k-1}} n'$ belongs to $C_{G/[N, \underbrace{G, \ldots, G}_{k-1}]}(\varphi)$, that is, $g_1^{\varphi} = g_1 n_1$, where $n_1 \in [N, \underbrace{G, \ldots, G}_{k-1}]$.

Since g_1 and n_1 commute modulo $[N, \underbrace{G, \ldots, G}_{k}]$, then modulo $[N, \underbrace{G, \ldots, G}_{k}]$, the elements $g_1, g_1^{\varphi}, g_1^{\varphi^2}, \ldots, g_1^{\varphi^{p-1}}$ also commute. The image of their product in $G/[N, \underbrace{G, \ldots, G}_{k}]$ therefore belongs to $C_{G/[N, \underbrace{G, \ldots, G}_{k}]}(\varphi)$. But we have:

$$g_1 \cdot g_1^{\varphi} \cdot g_1^{\varphi^2} \cdot \ldots \cdot g_1^{\varphi^{p-1}} \equiv g_1^{p} = (g^{p^{k-1}} n')^p \equiv g^{p^k} \pmod{N},$$

which proves our assertion.

The lemma is proved.

We now turn to the proof of Theorem 5.3.14. We consider first of all the case where the periodic part T of G is a finite p'-group. Then the group G, being finitely generated and nilpotent, is a residually finite p'-group (see, for instance, [123, Theorem 9.38]). Hence, there exists a normal subgroup N such that $N \cap T = 1$ and G/N is a finite p'-group. Replacing, if necessary, N by $\bigcap_{i=0}^{p-1} N^{\varphi^i}$, we may assume that N is φ-invariant (note that $G \Big/ \bigcap_{i=0}^{p-1} N^{\varphi^i}$ is also a finite p'-group, since by Remak's Theorem it embeds in the direct product of the finite p'-groups G/N^{φ^i}).

We observe that

$$[N, \underbrace{G, G, \ldots, G}_{h}] \leq \gamma_{h+1}(G) \cap N \leq T \cap N = 1.$$

This means that $N \leq \zeta_h(G)$.

We prove now that $C_{G/N}(\varphi) = C_G(\varphi)N/N$. Indeed, if $gN \in C_{G/N}(\varphi)$, then by Lemma 5.3.18, there is an element $n \in N$ such that

$$g^{p^h} n \in C_{G/[N, \underbrace{G, \ldots, G}_{h}]}(\varphi) = C_G(\varphi).$$

Hence, the image of g^{p^h} in G/N belongs to the image of $C_G(\varphi)$. Since G/N is a finite p'-group, the image of g also belongs to the image of $C_G(\varphi)$, as required.

Thus, in particular,

$$|C_{G/N}(\varphi)| = |C_G(\varphi)N/N| = |C_G(\varphi)/C_G(\varphi) \cap N| = |C_G(\varphi)| = q,$$

since $C_G(\varphi) \cap N \leq T \cap N = 1$. By Theorem 5.3.1, G/N contains a subgroup G_1/N of (p, q)-bounded index which is nilpotent of some p-bounded class $g = g(p)$. Then its full inverse image – the subgroup G_1 – is nilpotent of class $\leq g + h$, since $N \leq \zeta_h(G)$, as shown above. The subgroup G_1 is the required subgroup of (p, q)-bounded index which is nilpotent of p-bounded class.

We now consider the case where the periodic part T of G is a finite p-group.

By Corollary 1.7.4 the ranks of all of the φ-invariant abelian sections of T are (p, q)-bounded.

By Corollary 5.3.13 there is a (p, q)-bounded natural number $r = r(p, q)$, such that T^{p^r} is nilpotent of class $\leq h$.

We note that the order of T/T^{p^r} is bounded in terms of p and q. Indeed, the derived length of T/T^{p^r} is (p, q)-bounded, since this is true for T itself. Also the ranks of all factors of the derived series of T/T^{p^r} are (p, q)-bounded. This, together with the fact that their exponents are (p, q)-bounded (they divide p^r), yields the desired bound for the order.

We put $G_1 = C_G(T/T^{p^r})$, that is $G_1 = \{g \in G \mid [T, g] \leq T^{p^r}\}$. The subgroup G_1 is characteristic in G, just as T^{p^r} is, and its index is (p, q)-bounded, since G/G_1 embeds in the automorphism group of T/T^{p^r} which is of (p, q)-bounded order. Therefore it suffices to prove that the theorem holds for G_1 (which is, of course, φ-invariant).

The factor-group G_1/T^{p^r} is nilpotent of class $\leq h + 1$, since $\gamma_{h+1}(G_1) \leq T$ by Corollary 5.3.16 and since $[T, G_1] \leq T^{p^r}$ by definition. The group T^{p^r}, being nilpotent of class h, is soluble of derived length $\leq 1 + \log_2 h$ and so it has a characteristic series of p-bounded length having abelian factors. The theorem now follows from the following proposition by an obvious induction on the length of such a series (since the order of the centralizer of φ does not increase in the factor-groups over subgroups contained in T).

5.3.19 Proposition. *Let H be a φ-invariant subgroup of G_1 which contains a normal φ-invariant finite abelian p-subgroup A such that H/A is nilpotent of class s. Then H contains a φ-invariant subgroup of (s, p, q)-bounded index which is nilpotent of class $\leq s + h + 1$.*

Proof. We shall use here the well-known fact that the ranks of all abelian sections of a Sylow p-subgroup of the automorphism group of a finite abelian p-group are bounded in terms of the rank of the group (see [23] or [115]). The rank of

A is (p, q)-bounded by Corollary 1.7.4 and therefore the ranks of all abelian sections of $H/C_H(A)$ are also (p, q)-bounded, since the latter group embeds in the automorphism group of A and is a p-group since H is nilpotent.

Our aim is to estimate the order of $C_{H/C_H(A)}(\varphi)$. By Lemma 5.3.18, for any element $gC_H(A) \in C_{H/C_H(A)}(\varphi)$ there exists $c \in C_H(A)$ such that the image of $g^{p'}c$ in $H/[C_H(A), \underbrace{H, \ldots, H}_{s}]$ belongs to $C_{H/[C_H(A), \underbrace{H, \ldots, H}_{s}]}(\varphi)$. Since $[C_H(A), \underbrace{H, \ldots, H}_{s}] \leq \gamma_{s+1}(H) \leq A$, the image of $g^{p'}c$ in H/A also belongs to $C_{H/A}(\varphi)$. The order of $C_{H/A}(\varphi)$ does not exceed $q = |C_G(\varphi)|$ by Corollary 5.3.17. Hence, $(g^{p'}c)^q \equiv 1 \pmod{A}$ and therefore $g^{p'q} \in C_H(A)$, that is, the order of an arbitrary element $gC_H(A) \in C_{H/C_H(A)}(\varphi)$ divides $p^s q$.

Thus the exponent of $C_{H/C_H(A)}(\varphi)$ is (s, p, q)-bounded. Together with the bound for the nilpotency class (which is at most s, since H/A is nilpotent of class s) and the bound for the ranks of the factors of the lower central series (obtained above for all abelian sections of $H/C_H(A)$), this yields a bound for $|C_{H/C_H(A)}(\varphi)|$ in terms of s, p and q.

The natural semidirect product $A \leftthreetimes H/C_H(A)$ admits an automorphism φ of prime order p having an (s, p, q)-bounded number of fixed points equal to $|C_A(\varphi)| \cdot |C_{H/C_H(A)}(\varphi)|$. By Theorem 5.2.1, $A \leftthreetimes H/C_H(A)$ contains a subgroup B of (s, p, q)-bounded index which is nilpotent of class $\leq h$. We denote by A_1 the intersection of B with A, by $\bar{B}$ its image in $(A \leftthreetimes H/C_H(A))/A \cong H/C_H(A)$ and by H_1 the full inverse image of $\bar{B} \leq H/C_H(A)$ in H. Note that the index of H_1 in H is (s, p, q)-bounded.

Since the index $|A : A_1|$ is (s, p, q)-bounded and the rank of the abelian subgroup A is (p, q)-bounded, there is an (s, p, q)-bounded number $t = t(s, p, q)$, such that $A^{p^t} \leq A_1$. It is clear that the order of A/A^{p^t} is also (s, p, q)-bounded. (We replace A_1 by A^{p^t} since A^{p^t} is characteristic in A and is therefore φ-invariant and normal in H.)

Now we put

$$H_2 = H_1 \cap C_H(A/A^{p^t}),$$

where $C_H(A/A^{p^t}) = \{h \in H \mid [A, h] \leq A^{p^t}\}$. The factor-group $H/C_H(A/A^{p^t})$ embeds in the automorphism group of A/A^{p^t} and its order is therefore (s, p, q)-bounded. So, the index of H_2 in H is also (s, p, q)-bounded.

We shall prove that H_2 is nilpotent of class $\leq s + h + 1$, and is therefore what we require. We have

$$\gamma_{s+h+2}(H_2) \leq [A, \underbrace{H_2, H_2, \ldots, H_2}_{h+1}] \leq$$

$$\leq [A^{p'}, \underbrace{H_2, H_2, \ldots, H_2}_{h}] \leq [A_1, \underbrace{H_1, H_1, \ldots, H_1}_{h}].$$

By the definition of the semidirect product $A \leftthreetimes H/C_H(A)$, the subgroup $[A_1, \underbrace{H_1, \ldots, H_1}_{h}]$ coincides with $[A_1, \underbrace{B, \ldots, B}_{h}]$, which is trivial, since B is nilpotent of class $\leq h$.

The theorem is proved.

5.3.20 Corollary (Makarenko [104]). *Under the hypothesis of Theorem 5.3.14 there is a (p, q)-bounded number $s = s(p, q)$ such that the subgroup $G^s = \langle g^s \mid g \in G \rangle$ is nilpotent of class $\leq h(p)$.*

Proof. As in the proof of Corollary 5.3.13, an application of Theorem 5.3.14 enables us to assume from the outset that G is nilpotent of p-bounded class $c = c(p)$. It will then suffice to repeat the proof of Corollary 5.3.13 provided that we show that the additive group of the ideal ${}_{id}\langle C_{L(G)}(\varphi)\rangle$ has (p, q)-bounded exponent. This in turn follows from the fact that the additive group of $C_{L(G)}(\varphi)$ has (p, q)-bounded exponent.

It is clearly sufficient to prove that $C_{G/\gamma_i(G)}(\varphi)$ has (p, q)-bounded exponent for all i. Suppose that $g\gamma_i(G)$ is an element of $C_{G/\gamma_i(G)}(\varphi)$. Then by Lemma 5.3.18 there exists $n \in \gamma_i(G)$ such that the image of $g^{p^{c-i+1}}n$ in $G/[\gamma_i(G), \underbrace{G, \ldots, G}_{c-i+1}] = G$ belongs to $C_G(\varphi)$. (Note that here $[\gamma_i(G), \underbrace{G, \ldots, G}_{c-i+1}] \leq \gamma_{i+c-i+1}(G) = \gamma_{c+1}(G) = 1$). Hence, by Lagrange's Theorem, $\left(g^{p^{c-i+1}}n\right)^q = 1$ and therefore $g^{qp^{c-i+1}} \in \gamma_i(G)$. This means that the exponent of $C_{G/\gamma_i(G)}(\varphi)$ divides qp^c and is therefore (p, q)-bounded, since $c = c(p)$ is (p, q)-bounded.

The corollary is proved.

§ 5.4 Comments

Regular automorphisms. Unlike Lie rings, where the regularity of an automorphism implies solubility, and even nilpotency if the order of the regular automorphism is prime, in the case of groups one must impose some restrictions on the group. The example of a free group $F = \langle x_0, x_1, \ldots, x_{p-1}\rangle$ and the automorphism φ cyclically permuting the free generators ($x_i^{\varphi} = x_{i+1}$, where $i + 1$ is the residue modulo p), shows that in general the regularity of the automorphism (it is easy to

see that $C_F(\varphi) = 1$) does not imply the solubility of the group, even in the case where $|\varphi| = 2$. There exist also infinite soluble groups admitting regular automorphisms of prime order (and even of order 2) which are not nilpotent (and whose derived length is not bounded).

On the other hand, by Thompson's Theorem [140], any finite group admitting a regular automorphism of prime order, is nilpotent. This allows us to replace "nilpotent" in the hypothesis of Theorem 5.1.1 by "locally finite".

For estimates of Higman's function, see § 4.5.

As we said above in § 4.5, in the case of composite order, Kreknin's Theorem 4.3.1 does not imply any bound on the derived length of an (even *a priori* nilpotent) group with a regular automorphism. Such a bound is known only in the case where the automorphism is of order 4; more precisely, Kovács proved in [79] that if a locally finite or a locally nilpotent group G admits a regular automorphism of order 4, then $G'' \leq Z(G)$. We prove here a somewhat weaker result.

If G is a nilpotent periodic group with a regular automorphism of order 4, then the factor-group $G/[G, \varphi^2]$ is abelian, because φ induces on it a regular automorphism of order 2. In the associated Lie ring $L = L([G, \varphi^2])$ we have $C_L(\varphi) = 0$ and $L = 2L$. Therefore, $\tilde{L} = \langle {}^1L, {}^3L \rangle$ and ${}^0L = 0$ for $\tilde{L} = L \otimes_{\mathbb{Z}} \mathbb{Z}[i]$, where $i = \sqrt{-1}$. As was shown in § 4.5, this implies that $\tilde{L}$ is nilpotent of class ≤ 3, and hence the group $[G, \varphi^2]$ is also nilpotent of class ≤ 3. Thus G has a normal series $G \geq [G, \varphi^2] \geq 1$ of length 2 whose factors are nilpotent of classes 1 and 3.

On the other hand, under certain conditions, there exists a better correspondence between a nilpotent group and a Lie ring, which preserves the derived length. For example, the Baker-Hausdorff formula provides the so-called Mal'cev correspondence (a "category isomorphism") between radicable torsion-free nilpotent groups and nilpotent $\mathbb{Q}$-Lie algebras (see Chapter 8).

Another example is the recent work of Shalev and Zel'manov [130] where they gave a new, much shorter, proof of the solubility of pro-p-groups of finite coclass using Kreknin's Theorem. (Earlier Donkin [19] proved that p-adic analytic pro-p-groups of finite coclass are soluble for $p > 3$ and Leedham-Green [91] showed that every pro-p-groups of finite coclass is p-adic analytic. The work of Donkin applies the classification of simple p-adic Lie algebras, as well as Iwahori and Matsumoto's arithmetic theory of p-adic Chevalley groups.)

Kreknin' Theorem turned out to be applicable also in the "modular" case of an automorphism of order p^k acting on a finite p-group with few fixed points – see Chapter 8.

The example of a Lie ring L with a regular non-cyclic group of automorphisms A of order 4 from § 4.5 may be transformed into a family of examples of nilpotent groups, admitting A as a regular group of automorphisms, where derived lengths cannot be bounded by any constant ("function", depending on A only). To do this, we introduce a prime factor p into the structural constants of the Lie ring. Namely, we define a Lie ring R over $\mathbb{Z}$, whose additive group is generated by

elements e_1, e_2, e_3, with structural constants

$$[e_1, e_2] = pe_3;\ [e_2, e_3] = pe_1;\ [e_3, e_1] = pe_2.$$

It is easy to see that the factor-ring R/p^sR is nilpotent of class s, and its derived length is equal to $[\log_2 s] + 1$.

If the additive group of a nilpotent Lie ring is a p-group and its nilpotency class s is less than p, then the Baker-Hausdorff formula enables us, using the Lie ring operations, to define a group operation on the same set R/p^sR. The resultant group P is also nilpotent, its derived length is the same as before and all automorphisms of the Lie ring acting as before on this set turn out to be automorphisms of P. (This was noticed by Lazard [89]; this transition from Lie rings to groups is reversible and is also a so-called categorical isomorphism. Lazard's theorem generalizes the analogous Mal'cev correspondence [106] for nilpotent torsion-free radicable groups, see Chapter 8.)

It is easy to verify that the linear transformations defined on the generators e_1, e_2, e_3 by matrices

$$\begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix} \text{ and } \begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix}$$

generate a regular non-cyclic group A of order 4 of automorphisms of the Lie ring R. Hence $A \leq \operatorname{Aut} P$ and $C_P(A) = 1$, but the derived length of the group P tends to infinity with s and $p\ (s < p)$.

This easy way of using the Baker-Hausdorff formula has, however, made the derived length depend on p, but it seems that some more complicated construction may give examples which are not thus dependent.

In view of the example above the following result of Shumiatskii [134] seems very natural. He proves that if a finite soluble group of derived length k admits a regular group of automorphisms of exponent 2 and order 2^n, then it has a normal series of length n whose factors are nilpotent of (k, n)-bounded classes; in the case $n = 2$, he obtains a k-bounded estimate of the nilpotency class of the commutator subgroup. It would be interesting to generalize this theorem, for instance, to the case, where a finite soluble group of derived length k admits a regular group of automorphisms of order p^n or where a finite soluble p-group of derived length k admits a group of automorphisms of order p^n with a given number p^m of fixed points.

In this book we have been interested in automorphisms of nilpotent groups. We should mention that within the wider context of the theory of finite groups in general, a lot of effort has been invested in trying to prove the solubility of an arbitrary finite group G admitting a regular group of automorphisms A of coprime

order, that is, where $C_G(A) = 1$ and $(|G|, |A|) = 1$. This condition on orders is essential here, since the group of inner automorphisms of any finite group with trivial centre acts fixed-point-free. This conjecture has been proved modulo the classification of finite simple groups (K.-H. Clemens [15]). There is also a number of partial results (for example, when A is elementary abelian or cyclic of order a product of two primes, etc.), which do not use the classification theorem, but which, of course, are based on the techniques and results developed within classification theory (for instance, signalizer functors).

Within the theory of finite soluble groups much effort has been devoted to attempting to settle the following conjecture: if A is a regular group of automorphisms of a finite soluble group G where the order of A is coprime to that of G (that is $C_G(A) = 1$ and $(|G|, |A|) = 1$), then the nilpotent length of G is at most $n(A)$, where $n(A)$ denotes the number of prime factors in the decomposition of $|A|$ into a product of (not necessarily distinct) primes. The nilpotent length $h(G)$ of a soluble group G is the length of its shortest normal series with nilpotent factors. Thompson [141] proved that if A is a soluble group of automorphisms of a soluble group G and $(|G|, |A|) = 1$, then $h(G) \leq 5^{n(A)} \cdot h(C_G(A))$; Kurzweil [87] achieved a linear bound under the same hypothesis: $h(G) \leq 4n(A) + h(C_G(A))$. Further efforts were applied to obtaining estimates closer to the best possible one $n(A)$ (in the case of a soluble regular group of automorphisms A). At present this conjecture has been proved for a large class of groups of automorphisms, including cyclic ones (Shult [133], Gross [26], Berger [8, 9]); we also cite one of the results of Turull [142]: $h(G) \leq 2n(A) + h(C_G(A))$.

If one drops here the coprime condition on the orders, then, for every non-nilpotent finite group A there exist finite soluble groups of unbounded nilpotent lengths which admit A as a regular group of automorphisms. Such examples were constructed by Bell and Hartley [7].

If, however, A is nilpotent, then the coprime condition on the orders can be omitted at the expense of enlarging the bound for the nilpotent length of G: Dade [17] proved that if a soluble group H contains a nilpotent subgroup N which coincides with its normalizer, then the nilpotent length of H is bounded by some (exponential) function of $n(N)$; it is clear that if $C_G(A) = 1$ for a group of automorphisms A of a group G, then the normalizer of A in the natural semidirect product $G \leftthreetimes A$ coincides with A.

The method used for estimating the nilpotent length is known as the so-called Hall-Higman Theorems, originating from the works of P. Hall and G. Higman [32] which contains the reduction of the Restricted Burnside Problem for soluble groups to groups of prime-power exponents p^k, and of W. Feit and J.G. Thompson [20], proving the solubility of finite groups of odd order. The essence of this technique lies in considering the action of a group on its commutative invariant sections, which action is similar to that of a matrix group of linear transformations. This facilitates the use of results from representation and character theory. For instance,

one of the key steps in the proof of the Hall-Higman Theorem involves the action of an automorphism on the enveloping algebra of a representation of a finite group, which is a full matrix algebra.

Almost regular automorphisms: "the modular case". Alperin was the first in [2] to apply the theorem on a regular automorphism φ of prime order p of a Lie ring L, or rather its combinatorial form $\gamma_{h(p)+1}(pL) \subseteq {}_{id}(C_L(\varphi))$, to bound the derived length of a group satisfying the hypothesis of Theorem 5.2.1. For $p = 2$, Theorem 5.2.1 was proved by Hartley and Meixner in [35] (in fact, this paper deals with the more general case of a periodic group with an involution having a finite centralizer, where the basic fact is Shunkov's theorem [135] stating that such a group is locally finite and almost soluble).

In [59] Theorem 5.2.1 was proved giving the bound $h(p) + 1$ for the nilpotency class of a subgroup. This bound was improved to the apparently best possible value $h(p)$ by Makarenko in her recent work [103].

Within the theory of finite p-groups of maximal class, founded by Blackburn's work [10], Shepherd [131] and Leedham-Green and McKay [93] have proved that if $|C_P(\varphi)| = p$, where φ is an automorphism of prime order p of a finite p-group P, then P contains a subgroup of p-bounded index which is nilpotent of class ≤ 2. (They have also shown that a finite p-group of maximal class necessarily contains a subgroup of index p which is nilpotent of p-bounded class. In Chapter 6 we shall show that the commutator subgroup of a finite p-group of maximal class which has index p^2, is nilpotent of p-bounded class.)

We remark that the answer to a question raised by Leedham-Green and McKay in [93] on the existence of finite p-groups of maximal class having arbitrarily large derived lengths (of course, with increasing p), was given by Panfyorov in [118]. Namely, let $e_1, e_2, \ldots, e_n$ be a basis of a Lie algebra over $GF(p)$ with structural constants

$$[e_i, e_j] = \begin{cases} (i-j)e_{i+j}, & \text{if } i+j \leq n \\ 0, & \text{if } i+j > n \end{cases}.$$

It is easy to verify that these structural constants really define a Lie algebra which is nilpotent of class $n-1$ and whose derived length is $\approx \log_2 n$. The order of this Lie algebra is p^n. When $n - 1 < p$, by Baker-Hausdorff formula, this Lie algebra may be transformed into a finite p-group of the same order with the same nilpotency class and derived length (see above).

In further articles Donkin, Leedham-Green, Mann, McKay, Neubüser, Newman, Plesken, Shalev and Zel'manov have developed a theory of finite p-groups of given coclass. A finite p-group is said to have coclass d, if its order is p^n for some $n > d$ and its nilpotency class is $n - d$. It turns out that every finite p-group of coclass d contains a subgroup of (p, d)-bounded index which is nilpotent of class ≤ 2. These results prompt the following interesting questions. Does there

exist a function $f(m)$ such that any finite p-group admitting an automorphism of order p with exactly p^m fixed points, contains a subgroup of (p, m)-bounded index which is nilpotent of class $\leq f(m)$? Does there exist a function $g(m)$ such that any finite p-group admitting an automorphism of order p^k with exactly p^m fixed points, contains a subgroup of (p, k, m)-bounded index which is soluble of derived length $\leq g(m)$?

Chapter 8 contains the latest achievments in the "modular" case where a nilpotent p-group P admits an automorphism of order p^k with exactly p^m fixed points. It is proved there that such a group P is almost soluble with a strong bound, in terms of p and k only, on the derived length of a subgroup of bounded index. The proof is based on Kreknin's Theorem on Lie rings from Chapter 4. The proof uses group-theoretic corollary to Kreknin's Theorem, obtained with the help of the Mal'cev correspondence given by the Baker-Hausdorff formula, and some techniques from the theory of powerful p-groups, especially, from Shalev's work [129], where a weak bound, in terms of p, k and m, for the derived length of P was obtained.

We note that in the case $|\varphi| = 4$ one can expect a stronger conclusion since there is a stronger theorem for regular automorphisms of order 4 (see § 4.5).

Almost regular automorphisms of coprime order. If the derived length s of a group with an automorphism of prime order p with exactly q fixed points is known in advance, then it may turn out to be better to estimate the index of a nilpotent subgroup in terms of s, p and q and its class in terms of p and s, using a theorem proved in [63].

Theorem 5.3.1 was proved for metabelian groups by Meixner in [113] and in the case of $p = 2$ by Hartley and Meixner in [35]. (They also proved some other general facts on periodic groups with an involution with finite centralizer using Shunkov's Theorem [135].)

Hartley and Meixner [36] proved that if φ is an automorphism of prime order p of a finite soluble group G and $|C_G(\varphi)| = q$, then G contains a nilpotent subgroup of (p, q)-bounded index. Together with Fong's work [22], where it is proved modulo the classification of finite simple groups that an arbitrary finite group with an element of prime order p which has centralizer of order q, has a soluble subgroup of (p, q)-bounded index, this gives the following result: a locally finite group with an element of prime order p having a finite centralizer of order q, has a locally nilpotent subgroup of (p, q)-bounded index. Combining this with Theorem 5.3.1 we get (modulo the classification of finite simple groups) the following corollary.

5.4.1 Corollary. *If a locally finite group contains an element of prime order p having centralizer of finite order q, then it has a nilpotent subgroup of (p, q)-bounded index whose nilpotency class is p-bounded.*

The structure of finite soluble groups with almost regular groups of automorphisms (modulo the structure of nilpotent groups) was recently significantly clarified. Hart-

ley and Turau [37] proved that if a finite soluble group admits an automorphism of prime-power order p^k (which is not necessarily coprime to the order of the group) with exactly q fixed points then it has a subgroup of (q, p, k)-bounded index whose nilpotent length is at most k (this is a very natural generalization of results on regular automorphisms). Earlier Meixner [114] proved this theorem under the additional assumption that p is coprime to the order of the group. Under some additional assumptions a more general theorem was proved by Turull [143]. Finally, in a recent paper [34], Hartley and Isaacs proved that if A is a soluble group of automorphisms of a finite soluble group G with $|C_G(A)| = q$ and $(|G|, |A|) = 1$ then G contains a subgroup of $(q, |A|)$-bounded index whose nilpotent length is at most $2n(A) + 1$ (where $|A|$ is the product of $n(A)$ not necessarily distinct primes). It is also shown there that if A is a finite p-group, the coprime condition on the orders can be dropped at the expense of increasing by 1 the nilpotent length in the conclusion.

Here, estimation of nilpotent length is also done by using the above-mentioned theorems of Hall-Higman type.

It may be conjectured that if a nilpotent group admits an automorphism of prime-power order p^k having exactly q fixed points, then it has a subgroup of (p, k, q)-bounded index whose derived length is (p, k)-bounded. We recall, however, that up to now even the case of a regular automorphism, when $q = 1$, remains open. This conjecture is not yet proved either in the case of $p^k = 4$, where the regular case is known.

We finally point out one more application. P.V. Shumiatskii uses our Theorem 5.3.1 in his preprint *On locally finite groups admitting an automorphism of prime order whose centralizer is Černikov*, Thechnion, Haifa, 1992, to prove that a locally finite q-group admitting an automorphism of prime order p whose centralizer is a Černikov group, is soluble. Together with earlier results of B. Hartley (1982, 1988) and V. Turau (1985), this yields that a locally finite group having an element of prime order with Černikov centralizer has a soluble subgroup of finite index. (A group is said to be *Černikov* if it satisfies the minimum condition on subgroups and contains an abelian subgroup of finite index.)

Chapter 6

Nilpotency in varieties of groups with operators

In this chapter we prove two theorems of a rather general nature, bounding the nilpotency classes of nilpotent groups in certain varieties of groups with operators (for groups with operators see § 1.9). Let Ω be a group and let $\{v_\alpha\}$ be a family of Ω-identities defining a variety of operator groups $\mathfrak{M}$. We denote by $\{\bar{v}_\alpha\}$ the family of (ordinary) group identities obtained from $\{v_\alpha\}$ by replacing all operators from Ω by 1 and by $\overline{\mathfrak{M}}$ the variety of groups defined by the identities $\{\bar{v}_\alpha\}$.

We first of all suppose that there is a constant c bounding the nilpotency class of any nilpotent group in $\overline{\mathfrak{M}}$. The first result (Theorem 6.2.1) then states that, if for an Ω-group $G \in \mathfrak{M}$ the semidirect product $G \leftthreetimes \Omega$ is nilpotent, then the nilpotency class of G is also bounded by c.

Although the requirement that $G \leftthreetimes \Omega$ is nilpotent seems rather strong, it may be satisfied automatically – for instance in the case where both G and Ω are finite p-groups.

The author's result in [56] on the nilpotency of soluble groups with an automorphism φ of prime order p satisfying the $\langle\varphi\rangle$-identity

$$x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1$$

(that is, with a *splitting* automorphism of prime order p) is a prototype for the first theorem. If we put $\varphi = 1$ we get the identity $x^p = 1$. Now soluble groups of prime exponent p and of derived length s are nilpotent of class $\leq \frac{(p-1)^s-1}{p-2}$ by Proposition 3.4.4. Hence, Theorem 6.2.1 will yield that a soluble p-group of derived length s with a splitting automorphism of prime order p is nilpotent of class $\leq \frac{(p-1)^s-1}{p-2}$. This specific result is related to the structure theory of finite p-groups admitting a partition and will be applied in Chapter 7. We shall deduce it in this chapter as a corollary of a more general theorem. Its consequence is that the nilpotency class of the commutator subgroup of a finite p-group of maximal class is bounded in terms of p only.

The second result (Theorem 6.3.1) gives a positive solution to the Restricted Burnside Problem for a variety of operator groups $\mathfrak{M}$, provided that this problem has a positive solution for the corresponding ordinary variety $\overline{\mathfrak{M}}$. More exactly, suppose that the Restricted Burnside Problem has a positive solution for the variety

$\overline{\overline{\mathfrak{M}}}$ in the sense that locally nilpotent groups from $\overline{\overline{\mathfrak{M}}}$ constitute a subvariety and, moreover, that the associated Lie ring of a free group in $\overline{\overline{\mathfrak{M}}}$ satisfies a system of ***multilinear*** identities which define a locally nilpotent variety of Lie rings with a function $f(d)$ bounding the nilpotency class of a d-generator ring in this variety. We shall prove that if, for an Ω-group $G \in \mathfrak{M}$, the semidirect product $G \leftthreetimes \Omega$ is locally nilpotent, then G belongs to a locally nilpotent variety in which the nilpotency class of a d-generator group is bounded by $f\left(d\frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}\right)$.

Again we note that the strong condition that $G \leftthreetimes \Omega$ is locally nilpotent is automatically satisfied if both G and Ω are locally finite p-groups.

Instead of a condition on the identities of the associated Lie ring we could give an analogous condition on the group identities $\{\bar{v}_\alpha\}$, but such a condition would be stronger. An example given at the end of this chapter shows that the word "multilinear" in this condition is essential.

It is not yet clear whether the condition that the group of operators Ω is finite, is essential in the second theorem and whether one can choose the function to be independent of $|\Omega|$.

The main result of [61], which bounds the nilpotency class of a d-generator nilpotent group with a splitting automorphism of prime order p (that is, an analogue of the Restricted Burnside Problem for the variety of groups with operators which consists of all groups with a splitting automorphism of prime order p was solved in the positive in [61]) is a prototype for the second theorem. This theorem from [61] forms the basis for a structural theory of finite p-groups admitting a partition, including the positive solution of the Hughes problem for almost all finite p-groups. This theory is expounded in Chapter 7, where we also give the original proof of the theorem from [61] which yields some additional information and illustrates some aspects of Lie ring technique. This original proof uses a rather complicated transition to Lie rings, where Kostrikin's Theorem 1.3.1 is applied. Now the theorem from [61] may be obtained as a corollary to a more general result from this chapter since the identity $x^p = 1$ (obtained from the $\langle\varphi\rangle$-identity $x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1$) implies that the associated Lie ring satisfies the multilinear identities

$$px = 0 \quad \text{and} \quad \sum_{\pi \in \mathbb{S}_{p-1}} [x_0, x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(p-1)}] = 0$$

(equivalent to identities $px = 0$ and $[x, \underbrace{y, y, \ldots, y}_{p-1}] = 0$) from which the identities yielding local nilpotency follow by Kostrikin's Theorem 1.3.1.

The main idea of the proofs of both the theorems in this chapter is an iterative procedure expressing commutators in terms of commutators of greater weight. The argument is typical for the theory of nilpotent groups; see, for example, the proof

of the following elementary fact: if $\gamma_k(G) = \gamma_{k+1}(G)$ in a *nilpotent* group G then $\gamma_k(G) = 1$. Proof: from the hypothesis we obtain by substitution

$$\gamma_k(G) = \gamma_{k+1}(G) = [\gamma_k(G), G] = [\gamma_{k+1}(G), (G)] = \\ = [\gamma_k(G), G, G] = [\gamma_{k+1}(G), G, G] = \dots$$

and so on, so that $\gamma_k(G) = \gamma_{k+s}(G)$ for all $s \geq 1$; but $\gamma_N(G) = 1$ for some N, since G is nilpotent. Hence $\gamma_k(G) = 1$.

The proof of the second theorem is more difficult because of the need to control the number of generators involved in the iterations. For this purpose we use the power structure as well as the commutator structure in order to express commutators, not only in terms of commutators of greater weight, but also in terms of increasing powers of commutators of the same weight.

One of the main tools in the proofs is Higman's Lemma (see § 1.10).

§ 6.1 Preliminary lemmas

We fix notation: Ω will denote the group of operators and $V = \{v_\alpha\}$ a set of Ω-words $v_\alpha = v_\alpha(x_1, x_2, \dots, x_{n(\alpha)})$, where $x_1, x_2, \dots$ are free generators of a free Ω-group F. The group F, as an abstract group, contains a free subgroup $\bar{F}$, freely generated by the elements $x_1, x_2, \dots$. It is clear that $F = \langle \bar{F}^\Omega \rangle = [\bar{F}, \Omega] \leftthreetimes \bar{F}$.

Definition. The *projection* of an Ω-word $v = v(x_1, x_2, \dots, x_n)$ – an element of F – is the word $\bar{v} \in \bar{F}$, obtained from v by replacing all operators from Ω by the trivial one, that is, if

$$v = x_{i_1}^{\omega_1} \cdot x_{i_2}^{\omega_2} \cdot \dots \cdot x_{i_s}^{\omega_s}, \quad \omega_i \in \Omega,$$

then

$$\bar{v} = x_{i_1} \cdot x_{i_2} \cdot \dots \cdot x_{i_s}.$$

For example the projection of the $\langle \varphi \rangle$-word $x \cdot x^\varphi \cdot x^{\varphi^2} \cdot \dots \cdot x^{\varphi^{p-1}}$ is the word x^p.

We clearly have

6.1.1 Lemma. *The projection $\vartheta\colon v \to \bar{v}$ is a homomorphism of the group F onto $\bar{F}$, which is the identity on $\bar{F}$. Its kernel is $[\bar{F}, \Omega]$. Also $\vartheta^2 = \vartheta$.*

We make several remarks on verbal Ω-subgroups of F.

6.1.2 Lemma. *Let V be a set of Ω-words and $V(F)$ the corresponding verbal Ω-subgroup of F. Then $\overline{V(F)} = \bar{V}(\bar{F})$, where $\overline{V(F)}$ is the projection of $V(F)$ and $\bar{V}(\bar{F})$ is the verbal subgroup of $\bar{F}$ defined by the projection $\bar{V}$.*

Proof. Immediate from the definitions.

6.1.3 Lemma. *If V is an Ω-subgroup of F then $\bar{V}[F, \Omega] = V[F, \Omega]$.*

Proof. By Lemma 6.1.1 we have $\vartheta(\bar{V}) = \vartheta^2(V) = \vartheta(V)$, that is, the projections of V and $\bar{V}$ are equal. Hence their full inverse images are also equal so that $\bar{V}[F, \Omega] = V[F, \Omega]$, since $\operatorname{Ker}\vartheta = [F, \Omega]$ by Lemma 6.1.1.

6.1.4 Lemma. *The subgroups $[F, \Omega]$ and $\gamma_n(F\Omega) \cap F$, $n = 1, 2, \ldots$, are verbal Ω-subgroups of F.*

Proof. If σ is an endomorphism of F then the mapping which coincides with σ on F and is the identity on Ω may be extended to an endomorphism $\hat{\sigma}$ of the semidirect product $F \leftthreetimes \Omega$. Indeed, for any $f_1, f_2 \in F$, $\omega_1, \omega_2 \in \Omega$ we have

$$\begin{aligned}
\hat{\sigma}((f_1\omega_1)(f_2\omega_2)) = \hat{\sigma}(f_1\omega_1 f_2\omega_1^{-1}\omega_1\omega_2)) = \hat{\sigma}(f_1 \cdot f_2^{\omega_1^{-1}} \cdot \omega_1\omega_2)) = \\
= \sigma(f_1 \cdot f_2^{\omega_1^{-1}}) \cdot (\omega_1\omega_2) = \sigma(f_1) \cdot \sigma(f_2^{\omega_1^{-1}}) \cdot (\omega_1\omega_2) = \\
= \sigma(f_1) \cdot (\sigma(f_2))^{\omega_1^{-1}} \cdot (\omega_1\omega_2) = \sigma(f_1) \cdot \omega_1 \cdot \sigma(f_2) \cdot \omega_1^{-1} \cdot \omega_1\omega_2 = \\
= \sigma(f_1) \cdot \omega_1 \cdot \sigma(f_2) \cdot \omega_2 = \hat{\sigma}(f_1 \cdot \omega_1) \cdot \hat{\sigma}(f_2 \cdot \omega_2),
\end{aligned}$$

because $\sigma(f_2^{\omega_1^{-1}}) = (\sigma(f_2))^{\omega_1^{-1}}$, since the action of a homomorphism of an Ω-group commutes by definition with the action of operators from Ω (see § 1.9).

Now we have

$$\sigma([F, \Omega]) = \hat{\sigma}([F, \Omega]) = [\hat{\sigma}(F), \hat{\sigma}(\Omega)] = [\sigma(F), \Omega] \leq [F, \Omega],$$

and also

$$\sigma(\gamma_n(F\Omega) \cap F) = \hat{\sigma}(\gamma_n(F\Omega) \cap F) \leq \hat{\sigma}(\gamma_n(F\Omega)) \cap \hat{\sigma}(F) \leq \gamma_n(F\Omega) \cap F.$$

Hence, the subgroups $[F, \Omega]$ and $\gamma_n(F\Omega) \cap F$ of the free Ω-group F are invariant under every endomorphism of F, that is, they are verbal.

The lemma is proved.

The following arguments are variations on the theme of Higman's Lemma and collecting process. We recall that the normal closure of an element a in a group H is denoted by $\langle a^H \rangle$.

6.1.5 Lemma. *For every m the image of the subgroup*

$$\gamma_m(F) \cap [F, \Omega] \cap \bigcap_{i=1}^{n} \langle x_i^{F\Omega} \rangle$$

in $F/\gamma_{m+1}(F)$ is generated by the images of simple commutators of weight m in the elements x_i and $[x_i, \omega]$, $\omega \in \Omega$, which have the form

$$\big[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots\big], \tag{6.1.6}$$

where $\{i_1, i_2, \ldots\} \supseteq \{1, 2, \ldots, n\}$ (that is, the commutator (6.1.6) depends on all of the elements $x_1, x_2, \ldots, x_n$), the dots in simple subcommutators $[x_{i_s}, \ldots]$ denote occurrences of elements $\omega \in \Omega$ (which may not be there, in which case, by definition $[x_{i_s}, \ldots] = x_{i_s}$), and the commutator (6.1.6) has at least one such occurrence.

(In particular, since (6.1.6) has weight $\geq n$, we obtain automatically that $m \geq n$. That is,

$$[F, \Omega] \cap \bigcap_{i=1}^{n} \langle x_i^{F\Omega} \rangle \leq \gamma_n(F).)$$

Proof. The group F is generated by the elements x_i and $x_i^{\omega} = x_i \cdot [x_i, \omega]$, $\omega \in \Omega$, and hence it is also generated by the elements x_i and $[x_i, \omega]$, $\omega \in \Omega$. Therefore $\gamma_m(F)/\gamma_{m+1}(F)$ is generated by images of simple commutators of weight m in the generators x_i and $[x_i, \omega]$, $\omega \in \Omega$, which have the form (6.1.6) (see Proposition 2.1.5 c)). Hence, an arbitrary element

$$g \in \gamma_m(F) \cap [F, \Omega] \cap \bigcap_{i=1}^{n} \langle x_i^{F\Omega} \rangle$$

modulo $\gamma_{m+1}(F)$ is a product

$$g \equiv c_1^{k_1} \cdot c_2^{k_2} \cdot \ldots \cdot c_s^{k_s} \pmod{\gamma_{m+1}(F)} \tag{6.1.7}$$

of the powers of commutators c_i of the form (6.1.6). An argument analogous to the proof of Lemma 1.10.1, shows that each commutator c_i in (6.1.7) may be taken to involve all the $x_1, x_2, \ldots, x_n$. Indeed, for each $j \in \{1, 2, \ldots, n\}$ the homomorphism τ_j extending the mapping

$$x_j \to 1; \quad x_k \to x_k \text{ for } k \neq j$$

takes g to 1 because $g \in \langle x_j^{F\Omega} \rangle$ and takes to 1 all commutators c_i which involve x_j. Applying τ_j to (6.1.7) we obtain that the product of all powers of the commutators

occurring in (6.1.7) which do not involve x_j is congruent to 1 modulo $\gamma_{m+1}(F)$ and that this product may be omitted (here matters are simplified by the fact that the commutators c_i of weight m commute modulo $\gamma_{m+1}(F)$).

In analogous way it is easy to see that in (6.1.7) every commutator c_i, regarded as a commutator in the x_i and $\omega \in \Omega$, may be assumed to involve at least one element $\omega \in \Omega$. To show this we apply the projection homomorphism ϑ to (6.1.7). Since $g \in [F, \Omega]$, we have $\vartheta(g) = 1$. It is also clear that ϑ takes to 1 all commutators c_i which involve at least one element $\omega \in \Omega$. Applying ϑ to (6.1.7) we obtain that the product of all occurring in (6.1.7) powers of commutators which do not involve any $\omega \in \Omega$ is congruent to 1 modulo $\gamma_{m+1}(F)$ and this product may be also omitted.

The lemma is proved.

Definitions. We shall always denote by $[x_{i_s}, \ldots]$ simple commutators of the form

$$[x_{i_s}, \ldots] = [x_{i_s}, \omega_1, \omega_2, \ldots, \omega_t], \quad \omega_i \in \Omega, \tag{6.1.8}$$

where the dots denote occurrences of the elements of Ω (which may not be there, in which case, by definition $[x_{i_s}, \ldots] = x_{i_s}$).

For convenience we introduce the following definition: for an arbitrary commutator in the generators x_i of F and the elements $\omega \in \Omega$ the total number of occurrences of the x_i is called its *X-weight* and the total number of occurrences of the $\omega \in \Omega$ is called its *Ω-weight*. Of course the sum of the X-weight and the Ω-weight is equal to the weight of the commutator.

We now prove

6.1.9 Proposition. *For every m the image of $[F, \Omega] \cap \bigcap_{i=1}^{n} \langle x_i^{F\Omega} \rangle$ in $F/\gamma_{m+1}(F)$ is generated by commutators in the elements x_i and $\omega \in \Omega$, having the form*

$$\left[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots\right], \tag{6.1.10}$$

where $\{i_1, i_2, \ldots\} \supseteq \{1, 2, \ldots, n\}$ (that is, the commutator (6.1.10) involves all of the $x_1, x_2, \ldots, x_n$), all the simple subcommutators $[x_{i_s}, \ldots]$ are of the form (6.1.8) and the Ω-weight of (6.1.10) is at least 1.

Proof. We proceed by induction on m. For $m = n$ the proposition immediately follows from Lemma 6.1.5 and from the fact that $[F, \Omega] \cap \bigcap_{i=1}^{n} \langle x_i^{F\Omega} \rangle \leq \gamma_n(F)$. For $m > n + 1$ by the induction hypothesis an arbitrary element $g \in [F, \Omega] \cap \bigcap_{i=1}^{n} \langle x_i^{F\Omega} \rangle$

modulo $\gamma_m(F)$ is a product $\varkappa$ of powers of commutators of the form (6.1.10). Then $g\varkappa^{-1} \in \gamma_m(F)$ and also $g\varkappa^{-1} \in [F, \Omega] \cap \bigcap_{i=1}^{n} \langle x_i^{F\Omega} \rangle \leq \gamma_n(F)$, because all commutators (6.1.10) clearly belong to $[F, \Omega] \cap \bigcap_{i=1}^{n} \langle x_i^{F\Omega} \rangle \leq \gamma_n(F)$. Hence, by Lemma 6.1.5, the element $g\varkappa^{-1}$ modulo $\gamma_{m+1}(F)$ is congruent to a product $\varkappa'$ of powers of commutators of the form (6.1.10). As a result $g \equiv \varkappa \cdot \varkappa' \pmod{\gamma_{m+1}(F)}$.

The proposition is proved.

We record one more result of a technical nature.

6.1.11 Lemma. *In an arbitrary nilpotent group the following identity holds for any elements $a_1, a_2, \ldots, a_k$ and for any $m \in \mathbb{Z}$:*

$$\begin{aligned} &[a_1, a_2, \ldots, a_{s-1}, a_s^m, a_{s+1}, \ldots, a_k] = \\ &= [a_1, a_2, \ldots, a_{s-1}, a_s, a_{s+1}, \ldots, a_k]^m \cdot \varkappa \end{aligned} \tag{6.1.12}$$

where $\varkappa$ is a product of powers of commutators of weight $\geq k+1$ each of which involves all of the $a_1, a_2, \ldots, a_k$.

Proof. This is obtained by repeated applications of standard commutator identities from 2.1.1 and 2.7.5.

§ 6.2 A Nilpotency Theorem

In this section we prove the first of our results bounding the nilpotency class. We note that this theorem places no restrictions on the order of the operator group and so cannot be obtained as a corollary to our second theorem on locally nilpotent groups, since there the group of operators is finite.

We continue to use the notation from the previous section.

6.2.1 Theorem. *Let Ω be a group, $V = \{v_\alpha\}$ a set of Ω-words and let $\mathfrak{M}$ be the variety of Ω-groups defined by the identities V. Suppose that there is a constant c bounding the nilpotency class of any nilpotent group in the variety of groups $\overline{\mathfrak{M}}$ defined by the set of projections $\bar{V} = \{\bar{v}_\alpha\}$ of V. Then, if for an Ω-group $G \in \mathfrak{M}$ the semidirect product $G \leftthreetimes \Omega$ is nilpotent, the nilpotency class of G is at most c.*

Proof. The free Ω-group in $\mathfrak{M}$ is the factor-group $F/V(F)$ of the free Ω-group F by the verbal subgroup $V(F)$. Hence the set of Ω-identities $V(F)$ defines the same variety $\mathfrak{M}$. At the same time the free group in $\overline{\mathfrak{M}}$ is the factor-group $\bar{F}/\bar{V}(\bar{F})$ of the

free group $\bar{F}$ over the verbal subgroup $\bar{V}(\bar{F})$. By Lemma 6.1.2 we have $\overline{V(F)} = \bar{V}(\bar{F})$, so that replacing the set of Ω-identities V by $V(F)$ in the hypothesis of the theorem does not change the varieties $\mathfrak{M}$ and $\overline{\mathfrak{M}}$. Therefore in order to prove the theorem one can take V to be a verbal Ω-subgroup of F and we assume this in what follows.

It is sufficient to prove that all countably generated Ω-subgroups of G are nilpotent of class c. We may therefore assume that G is itself countably generated. Therefore there is a natural homomorphism of F onto G which extends the mapping of the free generators x_i to the generators of G. This homomorphism extends to a homomorphism of the semidirect product $F \leftthreetimes \Omega$ onto $G \leftthreetimes \Omega$ which is the identity on Ω. If m is the nilpotency class of $G \leftthreetimes \Omega$, then the kernel of this homomorphism contains $\gamma_{m+1}(F\Omega)$; it also contains the verbal subgroup V, since $G \in \mathfrak{M}$. It is therefore sufficient to prove that the nilpotency class of $F/V(F \cap \gamma_{m+1}(F\Omega))$ is bounded by c. Since the subgroups V and $F \cap \gamma_{m+1}(F\Omega)$ are verbal (the latter, by Lemma 6.1.4), it is sufficient to show that

$$[x_1, x_2, \ldots, x_{c+1}] \in V(F \cap \gamma_{m+1}(F\Omega))$$

where $x_1, x_2, \ldots, x_{c+1}$ are free generators of F.

The hypothesis of the theorem implies that the following holds in $\bar{F}$:

$$[x_1, x_2, \ldots, x_{c+1}] \in \bar{V}\gamma_{m+1}(\bar{F}).$$

Since $\bar{V}[F, \Omega] = V[F, \Omega]$ by Lemma 6.1.3, we have

$$[x_1, x_2, \ldots, x_{c+1}] \in V[F, \Omega]\gamma_{m+1}(F)$$

so that

$$[x_1, x_2, \ldots, x_{c+1}] \in V[F, \Omega](F \cap \gamma_{m+1}(F\Omega)).$$

To this we now apply Corollary 1.10.6 with

$$D_J = \bigcap_{i=1}^{c+1} \langle x_i^{F\Omega} \rangle, \quad M = [F, \Omega], \quad N = V(F \cap \gamma_{m+1}(F\Omega)),$$

recalling that $[F, \Omega]$ and $V(F \cap \gamma_{m+1}(F\Omega))$ are verbal subgroups by Lemma 6.1.4. We obtain

$$[x_1, x_2, \ldots, x_{c+1}] \equiv \varkappa \pmod{V(F \cap \gamma_{m+1}(F\Omega))},$$

where $\varkappa \in [F, \Omega] \cap \bigcap_{i=1}^{c+1} \langle x_i^{F\Omega} \rangle$. By Proposition 6.1.9, the element $\varkappa$ may be taken to be a product of powers of commutators of the form (6.1.10), that is,

$$[x_1, x_2, \ldots, x_{c+1}] \equiv c_1^{k_1} \cdot \ldots \cdot c_s^{k_s} \pmod{V(F \cap \gamma_{m+1}(F\Omega))}, \tag{6.2.2}$$

where each of the elements c_i is a commutator in the x_i and $\omega \in \Omega$, having the form

$$\big[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots\big], \tag{6.2.3}$$

where $\{i_1, i_2, \ldots\} \supseteq \{1, 2, \ldots, c+1\}$, the simple subcommutators $[x_{i_s}, \ldots]$ are of the form (6.1.8) and the Ω-weight of (6.2.3) is at least 1.

Because of the hypothesis $\{i_1, i_2, \ldots\} \supseteq \{1, 2, \ldots, c+1\}$, the commutator (6.2.3), as a simple commutator in the simple subcommutators $[x_{i_j}, \ldots]$, has weight $\geq c+1$. We express its initial segment of weight $c+1$ with the help of (6.2.2). We first apply to (6.2.2) the endomorphism of the Ω-group F extending the mapping

$$x_j \to [x_{i_j}, \ldots] \text{ for } j = 1, 2, \ldots, c+1, \quad x_j \to x_j \text{ for } j > c+1$$

(this homomorphism may be assumed to have been extended to a homomorphism of the semidirect product $F\Omega$ which is the identity on Ω). The image of the left-hand side of (6.2.2) is the following initial segment of (6.2.3)

$$\big[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots, [x_{i_{c-1}}, \ldots]\big],$$

and the images of the c_i on the right-hand side of (6.2.2) are clearly commutators in the elements $[x_{i_j}, \ldots]$ and $\omega \in \Omega$, having the form

$$\big[[[x_{m_1}, \ldots], ***], [[x_{m_2}, \ldots], ***], \ldots\big], \tag{6.2.4}$$

where $\{m_1, m_2, \ldots\} \supseteq \{i_1, i_2, \ldots, i_{c+1}\}$, the stars in the simple subcommutators $[[x_{i_s}, \ldots], ***]$ denote occurrences of elements $\omega \in \Omega$ (which may not be present, in which case, by definition $[[x_{i_s}, \ldots], ***] = [x_{i_s}, \ldots]$) and (6.2.4) has at least one such occurrence. We substitute the expressions obtained for the initial segment c_i of the form (6.2.3) into the original congruence (6.2.2) and transform the resultant right-hand side by applying 2.1.1 and 2.7.5 to get

$$[x_1, x_2, \ldots, x_{c+1}] \equiv c_1^{k_1} \cdot \ldots \cdot c_s^{k_s} \pmod{V(F \cap \gamma_{m+1}(F\Omega))}, \tag{6.2.5}$$

where (after changing notation), again, every element c_i is a commutator in the elements x_i and $\omega \in \Omega$, having the form

$$\big[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots\big], \tag{6.2.6}$$

where $\{i_1, i_2, \ldots\} \supseteq \{1, 2, \ldots, c+1\}$ and the simple subcommutators $[x_{i_s}, \ldots]$ are of the form (6.1.8) – but now the Ω-weight of (6.2.6) is at least 2, that is, (6.2.6) has at least ***two*** occurrences of elements $\omega \in \Omega$. This is because at least one occurrence is guaranteed by the dots in the simple subcommutators in (6.2.3), and at least one more by the stars in the simple subcommutators in (6.2.4).

Applying the same substitutions to (6.2.5) we obtain a congruence of the same form, where each commutator c_i has at least three occurrences of elements $\omega \in \Omega$ and so on, increasing the Ω-weight. But if the Ω-weight is large enough then all the c_i have weight large enough to be contained in $\gamma_{m+1}(F\Omega)$, and hence also in $F \cap \gamma_{m+1}(F\Omega)$, as they belong to F. Hence,

$$[x_1, x_2, \ldots, x_{c+1}] \equiv 1 \pmod{V(F \cap \gamma_{m+1}(F\Omega))},$$

as required.

The theorem is proved.

§ 6.3 A Local Nilpotency Theorem

We continue using the same notation.

6.3.1 Theorem. *Let Ω be a finite group, $V = \{v_\alpha\}$ a set of Ω-words, $\mathfrak{M}$ the variety of Ω-groups defined by identities V and $\overline{\mathfrak{M}}$ the variety of groups defined by the set of projections $\bar{V} = \{\bar{v}_\alpha\}$ of identities V. Suppose that the associated Lie ring $L(\bar{F}/\bar{V}(\bar{F}))$ of the free countably-generated group $\bar{F}/\bar{V}(\bar{F})$ of the variety $\overline{\mathfrak{M}}$ satisfies a system of multilinear identities which defines a locally nilpotent variety of Lie rings with a function $f(d)$ bounding the nilpotency class of a d-generator ring. If for an Ω-group $G \in \mathfrak{M}$ the semidirect product $G \leftthreetimes \Omega$ is locally nilpotent, then G belongs to a locally nilpotent variety in which the nilpotency class of a d-generator group is bounded by $f\left(d \cdot \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}\right)$.*

Proof. As we saw in the proof of Theorem 6.2.1, replacing the set of Ω-identities V by $V(F)$ in the hypothesis does not change the varieties $\mathfrak{M}$ and $\overline{\mathfrak{M}}$. Therefore, to prove the theorem, we may take V to be a verbal Ω-subgroup of F.

It is sufficient to obtain the required estimate $f\left(d \cdot \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}\right)$ for the nilpotency class of each d-generator subgroup H of G for all natural d. Replacing H by $\langle H^\Omega \rangle$ we may consider it to be Ω-invariant (and d-generator as an Ω-group). It is therefore sufficient to obtain the required estimate for the nilpotency class of G in the case where G is a d-generator group, and we shall assume that this is the

case in what follows. Then, since Ω is finite and $G \leftthreetimes \Omega$ is locally nilpotent, this semidirect product is, in fact, nilpotent of class m, say.

Since finitely generated nilpotent groups are residually finite (see, for instance, [123]), G has a family $\{N_\alpha\}$ of normal subgroups of finite index with trivial intersection. Replacing N_α by $\bigcap_{\omega \in \Omega} N_\alpha^\omega$ which also has finite index because Ω is finite, we may assume the N_α to be Ω-invariant. Each factor-group G/N_α is also a d-generator Ω-group in $\mathfrak{M}$ with the property that the semidirect product $(G/N_\alpha) \leftthreetimes \Omega$ is nilpotent. It is clear that it is sufficient to prove the required estimate for the nilpotency class of each of these groups. We may therefore assume that G is finite. Each of its Sylow p-subgroups is also a d-generator Ω-group, because it is the factor-group by the Hall p'-subgroup. Thus, G may be taken to be a finite p-group for some prime p which we fix for the rest of the section.

Since by Theorem 2.2.2 the identity of nilpotency of given class may be verified on the generators of the group, it is sufficient to show that any commutator of weight $f\left(d \cdot \frac{|\Omega|^{|\Omega'|}-1}{|\Omega|-1}\right)+1$ in the generators of G is equal to 1.

Here, instead of working in G, it is convenient to switch to working in a free group. Let F be a free Ω-group with a sufficiently large number of free generators $x_1, x_2, \ldots$ (or countably many of them) – a reserve of generators will be needed for technical reasons in the proof. There is a homomorphism of an Ω-group F onto G extending the mapping of $x_1, x_2, \ldots, x_d$ onto a set of generators for G. It may be also extended to a homomorphism of the semidirect product $F\Omega$ onto $G\Omega$ which is the identity on Ω. If m is the nilpotency class of $G\Omega$ then the kernel of this homomorphism contains $\gamma_{m+1}(F\Omega)$ and if p^n is the exponent of the finite p-group G then the kernel also contains the subgroup F^{p^n}; since $G \in \mathfrak{M}$, the kernel also contains V. It is therefore sufficient to prove that any commutator of weight $f\left(d \cdot \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}\right)+1$ in $x_1, x_2, \ldots, x_d$ is contained in $V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$.

The following proposition is the first step towards an iteration process of expressing any such commutator in terms of commutators of greater weight or in increasing powers of commutators.

6.3.2 Proposition. *For any k any commutator g of weight $f(k)+1$ in the free generators $x_1, x_2, \ldots, x_k$ of F is congruent modulo $V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$ to a product*

$$g \equiv c_1^{k_1} \cdot \ldots \cdot c_s^{k_s} \pmod{V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))}, \tag{6.3.3}$$

of powers of commutators c_i in the $x_1, x_2, \ldots, x_k$ and elements from Ω where each c_i has the form

$$\big[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots\big], \tag{6.3.4}$$

where $\{i_1, i_2, \ldots\} \subseteq \{1, 2, \ldots, k\}$, the simple subcommutators $[x_{i_s}, \ldots]$ have the form (6.1.8) and every commutator c_i either

a) *has at least* $f(k)+2$ *occurrences of the* $x_1, x_2, \ldots, x_k$ *(that is, its X-weight is greater than that of g), or*

b) *has at least* $f(k)+1$ *occurrences of the* $x_1, x_2, \ldots, x_k$ *(that is, its X-weight is at least* $f(k)+1$*) and has at least one occurrence of elements* $\omega \in \Omega$ *(that is, its* Ω*-weight is at least* 1*).*

Proof. Let $\bar{g}$ denote the Lie ring commutator with the same bracket structure as that of g in generators $\bar{x}_i$ of the associated Lie ring $L(\bar{F})$ of the free (ordinary) group $\bar{F}$, where $\bar{x}_i$ denotes the image of x_i in $\bar{F}/\bar{F}'$ considered as an element of $L(\bar{F})$ which is homogeneous of weight 1 (that is, $\bar{g}$ is obtained from g by replacing the $x_1, x_2, \ldots, x_k$ by $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$, respectively). By hypothesis $\bar{g}$ is contained in the verbal ideal of $L(\bar{F})$ generated by certain multilinear Lie polynomials $u_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{w_i})$ (homogeneous of degree w_i).

It is well-known that the additive group of this ideal is generated by the values $u_i(v_{i1}, v_{i2}, \ldots, v_{iw_i})$ of the polynomials $u_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{w_i})$ at arbitrary elements v_{ij} of $L(\bar{F})$. Since the u_i are multilinear, it is sufficient to take for the v_{ij} commutators in the $\bar{x}_i$ which generate the additive group $L(\bar{F})$. The Lie ring $L(\bar{F})$ is free and, hence, multihomogeneous. Therefore we have

$$\bar{g} = \sum_i \alpha_i u_i(v_{i1}, v_{i2}, \ldots, v_{iw_i}), \quad \alpha_i \in \mathbb{Z}, \tag{6.3.5}$$

where, for every i, the sum of the weights of the commutators $v_{i1}, v_{i2}, \ldots, v_{iw_i}$ in each variable $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ is equal to the weight of $\bar{g}$ in this variable; in particular, the sum of the weights of $v_{i1}, v_{i2}, \ldots, v_{iw_i}$ is $f(k)+1$. (In order to simplify the notation here the u_i are not supposed to be necessarily different for different i.)

Let us translate this statement into the language of the group $\bar{F}$. We denote by $\tilde{v}_{ij}$ group commutators in generators x_s with the same bracket structure as that of v_{ij} in generators $\bar{x}_s$. Let $\tilde{u}_i$ denote the product of powers of group commutators corresponding to u_i as a linear combination of Lie ring commutators (the exponents of the group commutators are equal to the coefficients of the corresponding Lie ring commutators), the order of the factors in this product here being irrelevant. Then, by definition of the associated Lie ring $L(\bar{F})$, (6.3.5) is equivalent to the congruence

$$g \equiv \prod_i \tilde{u}_i(\tilde{v}_{i1}, \tilde{v}_{i2}, \ldots, \tilde{v}_{iw_i})^{\alpha_i} \pmod{\gamma_{f(k)+2}(\bar{F})} \tag{6.3.6}$$

of products of commutators of weight $f(k)+1$ modulo $\gamma_{f(k)+2}(\bar{F})$ (see § 3.2). Here also, for every i, the sum of the weights of the $\tilde{v}_{i1}, \tilde{v}_{i2}, \ldots, \tilde{v}_{iw_i}$ in each variable $x_1, x_2, \ldots, x_k$ is equal to the weight of g in this variable.

For every polynomial u_i the fact that the associated Lie ring $L(\bar{F}/\bar{V})$ of the free group $\bar{F}/\bar{V}$ of the variety $\overline{\mathfrak{M}}$ satisfies the identity $u_i = 0$ is, in turn, equivalent to

$$\tilde{u}_i(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{w_i}) \equiv 1 \pmod{\gamma_{w_i+1}(\bar{F}/\bar{V})},$$

where $\hat{x}_i$ are the images of elements x_i in $\bar{F}/\bar{V}$. Lifting up to $\bar{F}$ this may be written as

$$\tilde{u}_i(x_1, x_2, \ldots, x_{w_i}) \in \bar{V} \cdot \gamma_{w_i+1}(\bar{F}).$$

To this we apply Corollary 1.10.6 with

$$D_J = \bigcap_{j=1}^{w_i} \langle x_j^{\bar{F}} \rangle, \quad M = \gamma_{w_i+1}(\bar{F}), \quad N = \bar{V}$$

where it is clear that $\bar{V}$ and $\gamma_{w_i+1}(\bar{F})$ are verbal subgroups of $\bar{F}$ and that $\tilde{u}_i(x_1, x_2, \ldots, x_{w_i})$ belongs to $D_J = \bigcap_{j=1}^{w_i} \langle x_j^{\bar{F}} \rangle$ because $u_i(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_{w_i})$ is multilinear. We obtain

$$\tilde{u}_i(x_1, x_2, \ldots, x_{w_i}) \equiv \varkappa_{i1} \pmod{\bar{V}},$$

where $\varkappa_{i1} \in \gamma_{w_i+1}(\bar{F}) \cap \bigcap_{j=1}^{w_i} \langle x_j^{\bar{F}} \rangle$. By Lemma 1.10.1

$$\varkappa_{i1} = d_1^{l_1} \cdot d_2^{l_2} \cdot \ldots \cdot d_s^{l_s}, \quad l_i \in \mathbb{Z},$$

where the d_i are commutators of weight $\geq w_i + 1$ in the x_j and each d_i involves all the elements $x_1, x_2, \ldots, x_{w_i}$.

We rewrite this congruence as

$$\varkappa_{i1}^{-1} \cdot \tilde{u}_i(x_1, x_2, \ldots, x_{w_i}) \in \bar{V},$$

and apply Lemma 6.1.3 to obtain

$$\varkappa_{i1}^{-1} \cdot \tilde{u}_i(x_1, x_2, \ldots, x_{w_i}) \in [F, \Omega] \cdot V$$

whence

$$\varkappa_{i1}^{-1} \cdot \tilde{u}_i(x_1, x_2, \ldots, x_{w_i}) \in [F, \Omega] \cdot V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega)).$$

To this we apply Corollary 1.10.6 with

$$D_J = \bigcap_{j=1}^{w_i} \langle x_j^{F\Omega} \rangle, \quad M = [F, \Omega], \quad N = V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$$

recalling that $[F, \Omega]$ and $F \cap \gamma_{m+1}(F\Omega)$ are verbal subgroups by Lemma 6.1.4 and both $\varkappa_{i1}^{-1}$ and $\tilde{u}_i(x_1, x_2, \ldots, x_{w_i})$ belong to $D_J = \bigcap_{j=1}^{w_i} \langle x_j^{F\Omega} \rangle$. We get

$$\varkappa_{i1}^{-1} \cdot \tilde{u}_i(x_1, x_2, \ldots, x_{w_i}) \equiv \varkappa_{i2} \pmod{V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))}$$

where $\varkappa_{i2} \in [F, \Omega] \cap \bigcap_{j=1}^{w_i} \langle x_j^{F\Omega} \rangle$. By Proposition 6.1.9, the element $\varkappa_{i2}$ may be taken to be a product of powers of commutators of the form (6.1.10), that is,

$$\varkappa_{i2} \equiv e_1^{m_1} \cdot \ldots \cdot e_s^{m_s} \pmod{V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))},$$

where each e_i is a commutator in $x_1, x_2, \ldots, x_{w_i}$ and $\omega \in \Omega$, of the form

$$\big[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots\big], \tag{6.3.7}$$

where $\{i_1, i_2, \ldots\} \supseteq \{1, 2, \ldots, w_i\}$, the simple subcommutators $[x_{i_s}, \ldots]$ are of the form (6.1.8) and the Ω-weight of (6.3.7) is at least 1.

As a result we obtain

$$\tilde{u}_i(x_1, x_2, \ldots, x_{w_i}) \equiv \varkappa_{i1} \cdot \varkappa_{i2} \pmod{V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))}.$$

We apply to this congruence the endomorphism of F which extends the mapping

$$\vartheta_i \colon x_1 \to \tilde{v}_{i1}, \ x_2 \to \tilde{v}_{i2}, \ldots, x_{w_i} \to \tilde{v}_{iw_i},$$

where $\tilde{v}_{i1}, \tilde{v}_{i2}, \ldots, \tilde{v}_{iw_i}$ are subcommutators from the right-hand side of (6.3.6).

The nature of $\varkappa_{i1}$ implies that its image under this endomorphism is equal to a product of powers of commutators $\vartheta_i(d_j)$ which have weight $\geq f(k) + 2$ as commutators in $x_1, x_2, \ldots, x_k$ – the weight $f(k) + 1$ is gained already from the single occurrences of the elements $\tilde{v}_{i1}, \tilde{v}_{i2}, \ldots, \tilde{v}_{iw_i}$ in $\vartheta_i(d_j)$, and, since the weight of d_j is at least $w_i + 1$, there must be at least one further occurrence. By the formulae from 2.1.1 and 2.7.5, every commutator $\vartheta_i(d_j)$ may be represented in the form of a product of powers of simple commutators of weight $\geq f(k) + 2$ in the x_i – these commutators have the form (6.3.4) and satisfy the property a).

From the description of the element $\varkappa_{i2}$ given above we see that its image under ϑ_i is equal to a product of powers of commutators $\vartheta_i(e_j)$ which, as commutators in $x_1, x_2, \ldots, x_k$ and $\omega \in \Omega$, have X-weight at least $f(k)+1$, gained from the occurrences of all the elements $\tilde{v}_{i1}, \tilde{v}_{i2}, \ldots, \tilde{v}_{iw_i}$, and have Ω-weight at least 1 because the Ω-weight of commutators (6.3.7) is at least 1. By Proposition 6.1.9 each commutator $\vartheta_i(e_j)$ may be represented as a product of powers of commutators of the form (6.1.10) of X-weight at least $f(k)+1$ and of Ω-weight at least 1 which are also commutators of the form (6.3.4) satisfying property b).

It remains to replace in (6.3.6) the elements $\tilde{u}_i(\tilde{v}_{i1}, \tilde{v}_{i2}, \ldots, \tilde{v}_{iw_i})$ by the expressions we have obtained and to rewrite the resulting congruence modulo $V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$:

$$g \equiv \varkappa_0 \cdot \prod_i (\vartheta_i(\varkappa_{i1}) \cdot \vartheta_i(\varkappa_{i2}))^{\alpha_i} \pmod{V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))},$$

where $\varkappa_0$ is an element of $\gamma_{f(k)+2}(\tilde{F})$ which may be represented modulo $F \cap \gamma_{m+1}(F\Omega)$ in the form of a product of powers of simple commutators of weight $\geq f(k)+2$ in the x_i which are commutators of the form (6.3.4) satisfying property a). In light of the collecting process (identities 2.7.5) it is clear that powers $(\vartheta_i(\varkappa_{i1}) \cdot \vartheta_i(\varkappa_{i2}))^{\alpha_i}$ of products of powers of commutators of the form (6.3.4) satisfying either property a) or b) are also products of powers of commutators of the form (6.3.4) satisfying either property a) or b).

The proposition is proved.

Now we are going to apply consequences of the congruences (6.3.3) to commutators appearing on the right-hand sides of them and then to commutators on the right-hand sides of the resultant congruences and so on. In the proof of Theorem 6.2.1 similar iterations allowed us to express the ambient commutator modulo $V \cdot (F \cap \gamma_{m+1}(F\Omega))$ in terms of commutators of large weight belonging to $F \cap \gamma_{m+1}(F\Omega)$. Here, however, there is an obstacle: the subcommutators $[x_{i_s}, \ldots]$ appearing on the right-hand side of (6.3.3) may be different elements even for the same x_{i_s} since different elements of Ω are involved (when the weight increases unboundedly). But to apply the congruences (6.3.3) we require definite dependence of the weight of a commutator on the number of variables occurring in it. This dependence is specified by the function $f(d)$ from the hypothesis of the theorem.

The growth of the number of variables occurring in the iteration process will be bounded by means of introducing p^k-th powers of commutators into that process; their exponents will increase, so that these powers will eventually belong to F^{p^n}.

Consider the factor-group $\bar{X} = X/(X'X^p)$, where $X = \langle x^\Omega \rangle \cong \langle x_i^\Omega \rangle$ for any free generator x_i of F. The group $\bar{X}$ is generated by $|\Omega|$ elements, the images of the elements x^ω, $\omega \in \Omega$, is commutative and has exponent p. Its order is thus at most $p^{|\Omega|}$. The group Ω acts on $\bar{X}$ in a natural way. The semidirect product $\bar{X} \leftthreetimes \Omega$ is

nilpotent by the hypothesis of the theorem: this implies that in the series

$$\bar{X} \geq [\bar{X}, \Omega] \geq [\bar{X}, \Omega, \Omega] \geq \ldots$$

all inclusions are strict until the series terminates at the identity subgroup (see Theorem 2.2.3 a)). Hence, the length of the series is at most $|\Omega|$.

So, for any elements $\omega_1, \omega_2, \ldots, \omega_t \in \Omega$ and for $t \geq |\Omega|$, the commutator $[x, \omega_1, \omega_2, \ldots, \omega_t]$ is contained in $X'X^p$. This statement may be elaborated in the following way.

6.3.8 Lemma. *Let x be one of the generators x_i of F. For any elements $\omega_1, \omega_2, \ldots, \omega_t \in \Omega$ and for $t \geq |\Omega|$ we have*

$$[x, \omega_1, \omega_2, \ldots, \omega_t] \equiv d_1^{l_1} \cdot d_2^{l_2}, \cdot \ldots \cdot d_s^{l_s} \pmod{F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))} \tag{6.3.9}$$

where in the

$$d_i = [[x, \ldots], [x, \ldots], \ldots], \tag{6.3.10}$$

all subcommutators $[x, \ldots]$ have the form (6.1.8) and have weight $\leq |\Omega|$ and for every factor $d_i^{l_i}$ in (6.3.9) either

1) *d_i has X-weight at least 2, or*
2) *l_i is a multiple of p.*

Proof. We have already noted that $[x, \omega_1, \omega_2, \ldots, \omega_t]$, for $t \geq |\Omega|$, lies in $X'X^p$. The subgroup X is generated by the elements $x^\omega = x[x, \omega]$, $\omega \in \Omega$. The elements x, $[x, \omega]$ commute modulo X' so that their p-th powers generate $X'X^p$ modulo X'.

The group X is generated by elements x, $[x, \omega]$, where $\omega \in \Omega$, each of which has X-weight ≥ 1. Via formulae 2.1.1 and 2.7.5, the commutator subgroup X' is generated modulo $F \cap \gamma_{m+1}(F\Omega)$ by commutators in x, $[x, \omega]$, each of which has X-weight at least two. By Proposition 6.1.9 each of these commutators may be expressed as a product of powers of commutators d_i of the form (6.3.10) having X-weight ≥ 2. If any of the d_i again has subcommutators of the form $[x, \omega_1, \omega_2, \ldots, \omega_t]$ with $t \geq |\Omega|$, then we replace them in an analogous way by products of p-th powers of x and $[x, \omega]$ and powers of commutators of X-weight ≥ 2. The resultant expression is then transformed into a product of powers of commutators of the form (6.3.10). Every such substitution replaces the factor $d_i^{l_i}$ in (6.3.9) by a product of factors of a similar form, but where the sum

$$(X\text{-weight}) + (\text{power of } p, \text{ dividing the exponent})$$

is greater. It is also clear that all resultant commutators of the form (6.3.10) satisfy at least one of the properties 1) or 2) of the lemma.

After a finite number of such steps all newly appearing factors will belong to $F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$ and they may be omitted modulo $F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$. This process will terminate at the congruence (6.3.9), in which, in commutators of the form (6.3.10), all subcommutators $[x, \ldots]$ have weight at most $|\Omega|$.

The lemma is proved.

We now complete the proof of the theorem. Let g be an arbitrary commutator of weight $f\left(d \cdot \frac{|\Omega|^{\Omega}-1}{|\Omega|-1}\right)+1$ in the d elements $x_1, x_2, \ldots, x_d$. We first of all use Proposition 6.3.2 to express g modulo $VF^{p^n}(F \cap \gamma_{m+1}(F\Omega))$ in the form of a product

$$g \equiv c_1^{k_1} \cdot \ldots \cdot c_s^{k_s} \pmod{V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))}, \tag{6.3.11}$$

of powers of commutators c_i of the form (6.3.4) satisfying one of the properties a), b).

Using Lemma 6.3.8 we replace every occurrence of simple subcommutators of the form $[x, \omega_1, \omega_2, \ldots, \omega_t]$, for $t \geq |\Omega|$, in commutators c_i from (6.3.11), by its expression as a product of powers of commutators of the form (6.3.10) satisfying one of the properties 1), 2). Each of the resultant commutators c_i may be expressed by Lemma 6.1.11 and by formulae from 2.1.1 and 2.7.5 as a product of powers of commutators of the form (6.3.4). This process gives either commutators of greater X-weight or p-th powers of commutators of the same X-weight. (Here formula 6.1.12 and formulae from 2.1.1 and 2.7.5 are applied to the c_i regarded as commutators in variables $[x_{i_s}, \ldots]$ which are not being changed.)

Taking into account the properties of the original congruence (6.3.11) we obtain that g may be expressed modulo $V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$ as a product of the form (6.3.11), where all c_i of the form (6.3.4) appearing now only contain subcommutators $[x_{i_s}, \ldots]$ of weight $\leq |\Omega|$, and for each factor $c_i^{k_i}$, either

a) the X-weight of c_i is greater than that of g,

or

b) the X-weight of c_i is equal to that of g and the Ω-weight of c_i is at least 1,

or

c) the X-weight of c_i is equal to that of g and k_i is a multiple of p.

Now we can start an iteration process aimed at the desired congruence (6.3.11) with trivial right-hand side. We shall describe its first step in some detail.

The weight of any c_i, as a simple commutator in subcommutators $[x_{i_s}, \ldots]$, is clearly equal to the X-weight of c_i. Since in each case a), b), c) this weight is at least $f\left(d \cdot \frac{|\Omega|^{\Omega}-1}{|\Omega|-1}\right)+1$, the commutator c_i has an initial segment of weight

$f+1 = f\left(d \cdot \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}\right)+1$ of the form

$$\left[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots, [x_{i_{f+1}}, \ldots]\right]. \tag{6.3.12}$$

Note that, since the subcommutators $[x_{i_s}, \ldots]$ have length (weight) at most $|\Omega|$ and $x_{i_s} \in \{x_1, x_2, \ldots, x_d\}$, at most

$$d(1 + |\Omega| + |\Omega|^2 + \ldots + |\Omega|^{|\Omega|-1}) = d \cdot \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}$$

of them may be distinct. Hence, each c_i, as a simple commutator in subcommutators $[x_{i_s}, \ldots]$, involves at most $d \cdot \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}$ variables.

In particular, there are at most $d \cdot \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}$ distinct elements among the $[x_{i_1}, \ldots]$, $[x_{i_2}, \ldots]$, $\ldots, [x_{i_{f+1}}, \ldots]$. Let us denote them by $y_1, y_2, \ldots, y_r$, where $r \leq d \times \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}$. We may therefore apply a consequence of a congruence of type (6.3.3) to such an initial segment. Namely, let τ be the endomorphism of F extending the mapping

$$x_1 \to y_1, \ x_2 \to y_2, \ldots, x_r \to y_r; \quad x_{r+s} \to x_{r+s}, \ s \geq 1.$$

By Proposition 6.3.2 the simple commutator in $x_1, x_2, \ldots, x_r$, which is obtained from (6.3.12) by replacing all elements $y_1, y_2, \ldots, y_r$ by $x_1, x_2, \ldots, x_r$, respectively, is congruent, modulo $V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$, to a product of commutators of the form (6.3.4). The image of the left-hand side of this congruence under τ is equal to (6.3.12). In the right-hand side of the resulting congruence

$$\begin{aligned} &\left[[x_{i_1}, \ldots], [x_{i_2}, \ldots], \ldots, [x_{i_{f+1}}, \ldots]\right] \equiv \\ &\equiv d_1^{l_1} \cdot \ldots \cdot d_s^{l_s} \pmod{V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))}, \end{aligned} \tag{6.3.13}$$

the images d_i of commutators of the form (6.3.4) are commutators in elements of the form $[y_{i_s}, \ldots]$ satisfying Proposition 6.3.2, the letter x being replaced by y.

Obviously, subcommutators of the form $[y_{i_s}, \ldots]$ are also commutators of the form $[x_{i_s}, \ldots]$. If among them there are "long" commutators $[x_{i_s}, \omega_1, \omega_2, \ldots, \omega_t]$ with $t \geq |\Omega|$, then by Lemma 6.3.8 we replace all their occurrences by products of powers of commutators of the form (6.3.10) satisfying the conclusion of this lemma. We then express the transformed commutators as products of powers of commutators of the form (6.3.4) by applying Lemma 6.1.11 and formulae from 2.1.1 and 2.7.5, regarding the transformed commutators as commutators in variables $[x_{i_s}, \ldots]$ which are not being changed. As a result we obtain an expression for (6.3.12) in the form (6.3.13), where all commutators d_i, however, now contain only subcommutators $[x_{i_s}, \ldots]$ of weight $\leq |\Omega|$ and, for each factor $d_i^{l_i}$, either

a′) d_i has X-weight $\geq f+2$,

or

b′) d_i has X-weight $\geq f+1$ and its Ω-weight is greater than the Ω-weight of (6.3.12),

or

c′) d_i has X-weight $\geq f+1$ and l_i is a multiple of p.

We express the initial segments of weight $f\left(d\cdot\frac{|\Omega|^{s}-1}{|\Omega|-1}\right)+1$ of all commutators c_i from (6.3.11) in this way and replace these initial segments in the same congruence (6.3.11) by their expressions. After transformation using Lemma 6.1.11 and formulae from 2.1.1 and 2.7.5, we obtain a new congruence of the form (6.3.11), where, however, every factor $c_i^{k_i}$ has a "doubled quality" because of different combinations of properties a), b), c) and a′), b′), c′). That is, either

– its X-weight is $\geq f+3$,

or

– its X-weight is $\geq f+2$ and its Ω-weight is ≥ 1,

or

– its X-weight is $\geq f+1$ and its Ω-weight is ≥ 2,

or

– its X-weight is $\geq f+2$ and k_i is divisible by p,

or

– its X-weight is $\geq f+1$, its Ω-weight is ≥ 1 and k_i is divisible by p,

or

– its X-weight is $\geq f+1$ and the exponent k_i is divisible by p^2.

Note that in the right-hand side of the resultant congruence all commutators c_i of the form (6.3.4) contain only simple subcommutators $[x_{i_s},\ldots]$ of weight $\leq|\Omega|$.

The latter remark implies that every commutator c_i in the new congruence of the form (6.3.11) also involves at most $d\cdot\frac{|\Omega|^{s}-1}{|\Omega|-1}$ different subcommutators $[x_{i_s},\ldots]$ of weight $\leq|\Omega|$. Each of the c_i has X-weight $\geq f+1$ and so again contains an initial segment of weight $f+1$ of the form (6.3.12). Again, to these initial segments we can apply appropriate consequences of congruences (6.3.3) from Proposition 6.3.2. Subsequent application of Lemma 6.3.8, formulae (6.1.12) and formulae from 2.1.1 and 2.7.5 will express them as products of powers of commutators of the form (6.3.4) in subcommutators $[x_{i_s},\ldots]$ of weight $\leq|\Omega|$ satisfying a′), b′), c′). Inserting these expressions into the c_i in the congruence of type (6.3.11) under consideration produces a new congruence of the form (6.3.11), but where, for each factor $c_i^{k_i}$, either the X-weight of c_i, or the Ω-weight of c_i, or the power of p dividing k_i is greater. At the same time the number of variables occurring in the congruence (subcommutators $[x_{i_s},\ldots]$ of weight $\leq|\Omega|$) will still be at most $d\cdot\frac{|\Omega|^{s}-1}{|\Omega|-1}$.

More precisely, we define the *height* of a factor $c_i^{k_i}$ in the product (6.3.11) to be the pair

$$\langle(X\text{-weight of } c_i)+(\text{power of } p,\text{ dividing } k_i);\ \ \Omega\text{-weight of } c_i\rangle.$$

We order such pairs lexicographically.

In fact, we have shown that subsequent application of Proposition 6.3.2 and Lemma 6.3.8 to the appropriate initial segment of c_i and application of formulae (6.1.12) and formulae from 2.1.1 and 2.7.5 allow us to express the factor $c_i^{k_i}$ in the form of a product of factors of the same form each of which has ***greater height.***

It is clear that multiple application of such iterations produces new congruences of the form (6.3.11) in which the heights of all factors increase unboundedly, the weights of subcommutators $[x_{i_s}, \ldots]$ remaining at most $|\Omega|$. Because of the bound $|\Omega|$ on the weights of the subcommutators $[x_{i_s}, \ldots]$, the Ω-weight of c_i is bounded by a function of its X-weight. Hence, if the height of a factor $c_i^{k_i}$ is large enough, then the sum

$$(X\text{-weight of } c_i) + (\text{power of } p \text{ dividing } k_i)$$

is also large enough. When this sum is greater than $m + n$, the factor $c_i^{k_i}$ belongs to $F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))$ and may be omitted in (6.3.11). As a result, we obtain

$$g \equiv 1 \pmod{V \cdot F^{p^n} \cdot (F \cap \gamma_{m+1}(F\Omega))},$$

as required.

The theorem is proved.

§ 6.4 Corollaries

First we derive consequences of the Nilpotency Theorem 6.2.1.

6.4.1 Corollary. *Suppose that P is a soluble locally finite p-group of derived length s which admits a finite p-group Ω as a group of operators. Suppose also that there exist elements $\omega_1, \omega_2, \ldots, \omega_p \in \Omega$ such that*

$$x^{\omega_1} \cdot x^{\omega_2} \cdot \ldots \cdot x^{\omega_p} = 1$$

for all $x \in P$. Then P is nilpotent and its nilpotency class is at most $\frac{(p-1)^s - 1}{p-2}$.

Proof. In order to prove that a group is nilpotent of class k it is sufficient to show that the identity

$$[x_1, x_2, \ldots, x_{k+1}] = 1$$

holds in it. We may therefore assume that P is finitely generated and, hence, finite. Then $P \leftthreetimes \Omega$ is also a finite p-group and thus nilpotent. Hence, we can apply Theorem 6.2.1, which yields the required bound on the nilpotency class of P. This

is because the projection of the given Ω-identity is the identity $x^p = 1$ and, by Proposition 3.4.4, soluble groups of derived length s and of prime exponent p are nilpotent of class $\leq \frac{(p-1)^s-1}{p-2}$.

6.4.2 Corollary. *If a soluble group G of derived length s admits a splitting automorphism φ of prime order p, that is,*

$$x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1$$

for all $x \in G$, then G is nilpotent and its nilpotency class is at most $\max\left\{\frac{(p-1)^s-1}{p-2}, h(p)\right\}$, *where $h(p)$ is Higman's function.*

Proof. Induction on the derived length s. For $s = 1$ the statement is trivial. Suppose that $s > 1$. The condition is obviously inherited by factor-groups modulo φ-invariant subgroups. Hence, by the induction hypothesis, $G/G^{(s-1)}$ is nilpotent. Therefore, G is abelian-by-nilpotent and hence residually finite by P. Hall's theorem [31]. If $\{N_\alpha\}$ is a family of normal subgroups of finite index in G with trivial intersection then $\left\{\bigcap_{i=1}^{p} N_\alpha^{\varphi^i}\right\}$ is a family of φ-invariant normal subgroups of finite index in G with trivial intersection. It is clearly sufficient to prove that every factor-group $G\Big/\bigcap_{i=1}^{p} N_\alpha^{\varphi^i}$ is nilpotent of appropriate nilpotency class. We may therefore assume that G is finite.

By Kegel's theorem [54] a finite soluble group with a splitting automorphism of prime order is nilpotent. Hence G is the direct product of its Sylow q-subgroups all of which are φ-invariant. It is easy to see that a splitting automorphism of order p is regular on p'-groups: if $x^{\varphi} = x$ then $1 = x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = x^p$, whence $x = 1$ if $(|x|, p) = 1$. Hence, all Sylow q-subgroups of G for $q \neq p$ are nilpotent of class at most $\min\left\{\frac{(p-1)^s-1}{p-2}, h(p)\right\}$ by Theorems 5.1.1 and 5.1.2.

We can apply Corollary 6.4.1 to bound the nilpotency class of the Sylow p-subgroup of G by $\frac{(p-1)^s-1}{p-2}$.

The corollary is proved.

Recall that a finite p-group is called a *p-group of maximal class* if, for some $n \in \mathbb{N}$, its order is p^n and its nilpotency class is $n - 1$. Such a group necessarily contains an element $a \in P \setminus P'$ with $|C_P(a)| = p^2$; in particular, $a^p \in Z(P)$ (see [10] or [48, Ch. III]).

6.4.3 Corollary. *The nilpotency class of the commutator subgroup of any finite p-group of maximal class is bounded in terms of p only.*

Proof. Let $|P| = p^n$ and let $a \in P \setminus P'$ be such that $|C_P(a)| = p^2$. By the well-known formula (Lemma 2.4.3) we obtain that

$$|\{[g, a] \mid g \in P\}| = |P : C_P(a)| = p^{n-2}.$$

Since $\{[g, a] \mid g \in P\} \subseteq P'$ and $|P : P'| \geq p^2$ (for any group P), these sets of equal orders must coincide: $\{[g, a] \mid g \in P\} = P'$.

Hence, for any $h \in P'$, there is an element $g \in P$, such that $h = [g, a] = g^{-1}g^a$, and therefore

$$h \cdot h^a \cdot h^{a^2} \cdot \ldots \cdot h^{a^{p-1}} =$$
$$= (g^{-1}g^a) \cdot (g^{-1}g^a)^a \cdot (g^{-1}g^a)^{a^2} \cdot \ldots \cdot (g^{-1}g^a)^{a^{p-1}} = g^{-1}g^{a^p} = 1.$$

Hence, a induces a splitting automorphism of prime order p of P' by conjugation.

By Corollary 5.2.5 the derived length of P is bounded in terms of p. Hence, by Corollary 6.4.2 the nilpotency class of P' is also bounded in terms of p.

The corollary is proved.

We now turn to consequences of the Local Nilpotency Theorem 6.3.1.

6.4.4 Corollary. *Suppose that P is a locally finite p-group which admits a finite p-group Ω as a group of operators. Suppose further that there exist $\omega_1, \omega_2, \ldots, \omega_p \in \Omega$ such that*

$$x^{\omega_1} \cdot x^{\omega_2} \cdot \ldots \cdot x^{\omega_p} = 1$$

for all $x \in P$. Then P belongs to a locally nilpotent variety in which the nilpotency class of a d-generator group is bounded by $k\left(d \cdot \frac{|\Omega|^{|\Omega|}-1}{|\Omega|-1}\right)$, where $k(d)$ is the function bounding the nilpotency class of a d-generated $(p-1)$-Engel Lie algebra of characteristic p.

(The function $k(d)$ exists by Kostrikin's Theorem 1.3.1.)

Proof. The projection of the given Ω-identity is the identity $x^p = 1$. The associated Lie ring of a group of prime exponent p satisfies the multilinear identities

$$px = 0 \quad \text{and} \quad \sum_{\pi \in \mathbb{S}_{p-1}} [x_0, x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(p-1)}] = 0$$

by the Magnus-Sanov Theorem 3.3.2. These identities imply the identities of a locally nilpotent variety by Kostrikin's Theorem 1.3.1. The semidirect product $P \leftthreetimes \Omega$ is a locally finite p-group and, hence, is locally nilpotent. Therefore, we can apply Theorem 6.3.1.

6.4.5 Corollary. *If a d-generator nilpotent p-group G admits a splitting automorphism φ of prime order p, that is,*

$$x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1$$

for all $x \in G$, then the nilpotency class of G does not exceed $k\left(d \cdot \frac{p^p-1}{p-1}\right)$ where $k(d)$ is the function bounding the nilpotency class of a d-generated $(p-1)$-Engel Lie algebra of characteristic p.

§ 6.5 Comments

Note that in Corollary 6.4.2 we obtain a slightly better bound for the nilpotency class than that obtained in the original work [56] – there it was $\frac{p^5-1}{p-1}$.

The author would be interested to learn about any other varieties where there is a positive solution to the Restricted Burnside Problem in the "multilinear" sense of the hypothesis of Theorem 6.3.1.

Unfortunately, the identities of associated Lie rings of groups of composite orders p^k are not multilinear. Therefore, although Zel'manov [158-160] has obtained the positive solution to the Restricted Burnside Problem for such groups, we cannot apply Theorem 6.3.1 to the corresponding operator groups. We show by means of an example at the end of this section that it is in fact impossible to bound the nilpotency class of these operator groups.

Corollary 6.4.5 is the main result of [61], it is the basis for a structural theory of finite p-groups admitting a partition, including the positive solution of the Hughes problem for almost all finite p-groups. This theory is described in Chapter 7, where we also give the original proof of the theorem from [61], which yields some additional information and illustrates some aspects of Lie ring technique. This original proof uses a rather complicated transition to Lie rings, where Kostrikin's Theorem 1.3.1 is applied. Now the main theorem of [61] may be obtained as a corollary to the more general Theorem 6.3.1 (and, perhaps, with better bounds for the nilpotency classes) – but, of course, the proof also uses Kostrikin's Theorem 1.3.1.

The main step of the original proof consists in bounding the order of $C_G(\varphi)$ in terms of d and p (where G is a d-generator p-group with a splitting automorphism φ of order p). This also enables us to obtain another result for a group satisfying the hypothesis of Corollary 6.4.5. Note that a bound of the same kind follows also *a posteriori* from Corollary 6.4.5. Namely, since the nilpotency class and the number of generators of the group G under consideration are bounded in terms of d and p, it follows, via Corollary 2.5.6 and Proposition 2.5.7, that every subgroup of G may also be generated by a (d, p)-bounded number of elements. Since $C_G(\varphi)$

is nilpotent of the same (d, p)-bounded class and all of the factor-groups of its lower central series have exponent p, as does $C_G(\varphi)$ itself, the order of $C_G(\varphi)$ is also bounded in terms of d and p.

We shall see in Chapter 7 how Corollary 6.4.5 is connected with the theory of finite p-groups admitting a partition which are exactly those p-groups which are different from their Hughes subgroups. It would be interesting to find analogous applications of the more general Theorem 6.3.1 possibly by defining some subgroup generalizing the Hughes subgroup.

Another application of the result of Corollary 6.4.5 is the remark on periodic compact groups in § 7.5, where it is proved that if such a group contains an open subset of elements of prime order p then it has a subgroup of finite index contained in a locally nilpotent variety. This implies also that the group is locally finite and has bounded exponent. Zel'manov, using the techniques of his positive solution to the Restricted Burnside Problem for groups of exponent p^k, proved in [161] that any periodic compact group is locally finite. It remains an open problem whether every such group has bounded exponent.

Periodic compact groups may be characterized as periodic profinite groups. It is sufficient to consider pro-p-groups; the reduction of the above-mentioned problem to this case is due to Wilson [154].

It is not difficult to show that any pro-p-group has an open subset consisting of elements of equal order p^k. Such a subset can be linked with an automorphism φ of order p^k of some open subgroup H for which

$$x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p^k-1}} = 1$$

for all $x \in H$. One can conjecture that a result analogous to Corollary 6.4.5 must be valid for groups admitting such an automorphism. As we have already remarked, although the projection of the given $\langle\varphi\rangle$-identity is the identity $x^{p^k} = 1$, we cannot combine Theorem 6.3.1 with Zel'manov's solution of the Restricted Burnside Problem for groups of exponent p^k, because the corresponding Lie ring identities are not multilinear. Moreover, even for $|\varphi| = 4$, it is impossible to bound the nilpotency class of finitely generated groups with such an automorphism – see the Example below. One can only suppose that if a nilpotent group H admits such an automorphism then it belongs to a locally soluble variety – this may turn out to be enough to prove the boundedness of the exponent of a periodic compact group.

6.5.1 Example. Let $\langle\varphi\rangle$ be an automorphism of order 2 of the wreath product $G = \langle b\rangle \wr \langle a\rangle$ of cyclic groups $\langle b\rangle$ and $\langle a\rangle$ of orders 2 and 2^n, respectively, acting on it in the following way: if

$$\langle b\rangle \wr \langle a\rangle = \langle a, b_1, b_2, \ldots, b_{2^n} \mid b_1^2 = b_2^2 = \ldots = b_{2^n}^2 = 1,\ b_i b_j = b_j b_i,$$
$$b_i^a = b_{i+1},\ \text{where } i+1 \text{ is a residue modulo } 2^n\rangle,$$

then $a^{\varphi} = a^{-1}$ and $b_i^{\varphi} = b_{-i}$, where $-i$ is a residue modulo 2^n. Then G, as a $\langle\varphi\rangle$-group, satisfies the $\langle\varphi\rangle$-identity

$$x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot x^{\varphi^3} = 1. \tag{6.5.2}$$

Indeed, $x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot x^{\varphi^3} = (x \cdot x^{\varphi})^2$, because $\varphi^2 = 1$, and, modulo the basis $\langle b_1, b_2, \ldots, b_{2^n}\rangle$ of the wreath product, the product $x \cdot x^{\varphi}$ is trivial, since $a^{\varphi} = a^{-1}$. Hence, $x \cdot x^{\varphi} \in \langle b_1, b_2, \ldots, b_{2^n}\rangle$ and therefore $x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot x^{\varphi^3} = (x \cdot x^{\varphi})^2 = 1$, since $\langle b_1, b_2, \ldots, b_{2^n}\rangle$ has exponent 2.

The projection of the $\langle\varphi\rangle$-identity (6.5.2) is the identity $x^4 = 1$ and groups of exponent 4 constitute a locally nilpotent variety by Sanov's theorem [124]. However, in spite of the fact that $G \leftthreetimes \langle\varphi\rangle$, being a finite 2-group, is nilpotent, the nilpotency class of the 2-generator group G increases unboundedly with n.

Chapter 7

Splitting automorphisms of prime order and finite p-groups admitting a partition

We have already come across the variety $\mathfrak{M}_p$ of operator groups consisting of all groups with a splitting automorphism φ of prime order p, that is, which satisfy the operator identities

$$x^{\varphi^p} = x \quad \text{and} \quad x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1.$$

Corollary 6.4.2 from the preceding chapter states that all soluble groups in $\mathfrak{M}_p$ are nilpotent. Corollary 6.4.5 gives a positive solution to the Restricted Burnside Problem for $\mathfrak{M}_p$: the locally nilpotent groups in $\mathfrak{M}_p$ form a subvariety or, equivalently, the nilpotency class of a d-generator nilpotent group in $\mathfrak{M}_p$ is bounded by a function depending only on d and p (there are simple examples showing that no such bound for the order exists). We give here another proof of this result which provides additional information on the associated Lie ring of the semidirect product $G \leftthreetimes \langle\varphi\rangle$, $G \in \mathfrak{M}_p$, and which illustrates some Lie ring technique. This proof uses Kostrikin's Theorem 1.3.1 on $(p-1)$-Engel Lie algebras of characteristic p together with generalizations of Higman's Theorem on regular automorphisms of prime order applied to the case of a finite p-group with an automorphism of order p – the Alperin-Khukhro Theorem 5.2.1. The main technical lemma is an analogue of the Magnus-Sanov Theorem 3.3.2 on the $(p-1)$-Engel condition for the associated Lie ring of a group of prime exponent p.

For finite p-groups, the study of splitting automorphisms of order p is equivalent to the study of groups admitting a proper partition (or groups which are distinct from their Hughes subgroup). Thus, theorems on splitting automorphisms give rise to a structural theory of finite p-groups admitting a partition. This theory includes a positive solution of the Hughes problem for almost all (in some precise sense) finite p-groups – in spite of the fact that there exist counterexamples to the Hughes conjecture. In the Comments in § 7.5 we indicate how the existence of these counterexamples is connected with new identities in Lie rings of free groups of prime exponent. In the opposite direction, we prove in § 7.4 another positive result bounding the index of the Hughes subgroup under a certain hypothesis on these

Lie rings. The proof here yields also unconditional results on any finite p-group whose Hughes subgroup has index p^k.

An important methodological principle in this chapter lies in the fact that we consider free groups of varieties or varieties of operator groups; Higman's Lemma is applied several times in different forms. We also introduce certain other universal groups analogous to free groups of varieties; for example, we prove the existence of universal counterexamples to the Hughes conjecture.

Different approaches to the study of finite p-groups with a partition are discussed in § 7.1 where we prove their equivalence. § 7.2 contains the positive solution to the Restricted Burnside Problem for $\mathfrak{M}_p$ and § 7.3 is devoted to its corollaries for finite p-groups with a partition and, in particular, for the Hughes problem.

§ 7.1 The connection between splitting automorphisms of prime order and finite p-groups admitting a partition

Definition. A group G is said to have a non-trivial (or proper) *partition* if it is the set-theoretic union of some of its proper subgroups with pairwise trivial intersections:

$$G = \bigcup_{\alpha} G_\alpha, \quad G_\alpha < G \quad \text{for all } \alpha, \quad G_\alpha \cap G_\beta = 1 \quad \text{for } \alpha \neq \beta.$$

This definition does not presuppose that G is finite; groups admitting a partition have been extensively studied both from a general point of view and for particular classes of groups – see the survey in § 7.5. We are primarily interested here in finite p-groups admitting a partition. Hitherto the study of such groups has been carried out exclusively in the context of the Hughes problem. This problem was posed in 1957 in [45]; to state it, we need the following definition.

Definition. Let p be a prime. The *Hughes subgroup* $H_p(G)$ of a group G (with respect to a given prime p) is the smallest subgroup of G outside of which all elements have order p, that is

$$H_p(G) = \langle x \in G \mid x^p \neq 1 \rangle.$$

(Here, as usual, a subgroup generated by the empty set is trivial by definition; in this case $H_p(G) = 1$ if G is a group of exponent p.)

Note that $H_p(G)$ is a characteristic subgroup of G by definition.

Of course, the Hughes subgroup often coincides with the whole group; on the other hand, we have observed that the Hughes subgroup is trivial in a group of

prime exponent p. D.R. Hughes asked: is it true that, if in a finite group $G \neq H_p(G) \neq 1$, then $|G : H_p(G)| = p$? We shall say that a group G *satisfies the Hughes conjecture* (for a given prime p) if either $H_p(G) = G$ or $H_p(G) = 1$, or $|G : H_p(G)| = p$.

The Hughes conjecture was proved for groups that are not p-groups in 1959 by Hughes and Thompson [46] on the basis of the fundamental work of Thompson on normal p-complements [140]; in particular, it was proved in [46] that a proper Hughes subgroup of a finite group is necessarily soluble. Kegel in [54] supplemented this result by showing that a proper Hughes subgroup of a finite group is nilpotent.

However, positive results on the Hughes problem for finite p-groups were of a partial nature. Most interesting were counterexamples to the Hughes conjecture first constructed by Wall [147] in 1965 for $p = 5$ with the value p^2 for the index of a non-trivial Hughes subgroup. (See the survey in § 7.5.)

In spite of the existence of counterexamples, one of the aims of this chapter is a positive solution to the Hughes problem for almost all (in some precise sense) finite p-groups. Another positive result bounds the index of the Hughes subgroup under a hypothesis that a certain conjecture on the Lie ring of a free group of prime exponent is true.

An approach connected with splitting automorphisms of prime order p, which originates from the works of Hughes and Thompson [46] and Kegel [54], turned out to be most productive both in formulating problems and in results. We recall an important definition.

Definition. An automorphism φ of a group G is called a *splitting automorphism of prime order p* if

$$\varphi^p = 1 \quad \text{and} \quad x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1$$

for all $x \in G$. Note that we do not exclude the case $\varphi = 1$ where, of course, G has exponent p.

The following proposition unites three approaches to the study of finite p-groups admitting a partition.

7.1.1 Proposition. *If P is a finite p-group then the following three conditions are equivalent:*

a) *P admits a non-trivial partition;*

b) *P is distinct from its Hughes subgroup: $H_p(P) \neq P$;*

c) *P may be expressed as a semidirect product $P_1 \leftthreetimes \langle \varphi \rangle$ where φ is a splitting automorphism of prime order p of P_1; here one can take any subgroup of index p containing $H_p(P)$ for P_1 and any element of P outside P_1 for φ.*

Proof. a) $\Rightarrow$ b). Let $P = \bigcup_\alpha P_\alpha$ where the P_α are proper subgroups of P with pairwise trivial intersections (called components). We prove that all elements of order greater than p belong to a single component. Then, obviously, the subgroup $H_p(P)$ generated by these elements is also contained in this component and hence is a proper subgroup. We fix an arbitrary element z of order p in the centre of P. Now, if x is any element of order greater than p, then $(xz)^p = x^p z^p = x^p \neq 1$. Since the components have pairwise trivial intersections, the elements x and xz, whose p-th powers are equal and non-trivial, must lie in the same component containing this p-th power. This component also clearly contains z. Hence all elements of order greater than p belong to the single component containing z.

b) $\Rightarrow$ a). If $H_p(P) \neq P$, then $H_p(P)$ together with cyclic subgroups of order p outside $H_p(P)$ obviously constitute a non-trivial partition of P.

b) $\Rightarrow$ c). Let $H_p(P) \neq P$ and let P_1 be an arbitrary subgroup of index p containing $H_p(P)$. The following identity holds for any elements φ and x in any group:

$$x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = (x\varphi^{-1})^p \varphi^p. \tag{7.1.2}$$

We choose for φ an arbitrary element from $P \setminus P_1$. Now, if x is any element of P_1, then neither $x\varphi^{-1}$ nor φ belong to P_1 and hence neither of them belongs to $H_p(P) \leq P_1$. By the definition of $H_p(P)$ these elements have order p and by (7.1.2) φ may be regarded as a splitting automorphism of order p of P_1. It is also clear that $P = P_1 \leftthreetimes \langle \varphi \rangle$.

c) $\Rightarrow$ b). Let $P = P_1 \leftthreetimes \langle \varphi \rangle$ where φ is a splitting automorphism of prime order p of P_1. By (7.1.2) we get $(x\varphi^{-1})^p = 1$ for any $x \in P_1$. Since any non-trivial element of P/P_1 which has order p is a power of the image of φ^{-1}, any element of $P \setminus P_1$ is of the form $y\varphi^{-k}$ for some $k \not\equiv 0 \pmod p$ and $y \in P_1$. Since P_1 is a normal subgroup, this implies that any element of $P \setminus P_1$ is a power of an element of the form $x\varphi^{-1}$, which has order p, and hence has order p itself. Hence by definition we have $H_p(P) \leq P_1 < P$.

The proposition is proved.

Finite p-groups admitting partitions may be regarded as a generalization of groups of prime exponent p. This makes sense from each of the three points of view represented in the statement of Proposition 7.1.1. Firstly, any group of prime exponent p admits a partition consisting of subgroups of order p. Secondly, in a finite p-group P admitting a partition, all elements outside $H_p(P)$ – and their number is at least $|P|(p-1)/p$ – have prime order p. Finally, the operator identity $x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}} = 1$ becomes the identity $x^p = 1$ on putting $\varphi = 1$. It is therefore not surprising that the techniques and results accumulated in the study of groups of prime exponent may be successfully applied in the study of finite p-groups admitting a partition. Thus the existence of counterexamples to the Hughes

conjecture is connected with new identities in the associated Lie rings of free groups of prime exponent and the proofs of the positive results in this chapter use Kostrikin's Theorem 1.3.1 on $(p-1)$-Engel Lie algebras and an analogue of the Magnus-Sanov Theorem 3.3.2 on the $(p-1)$-Engel condition for the Lie ring of a group of prime exponent. The approach connected with splitting automorphisms of prime order p is more "categorical" as it introduces the variety of operator groups $\mathfrak{M}_p$ defined by the $\langle\varphi\rangle$-identities $x^{\varphi^p} = x$ and $xx^{\varphi}x^{\varphi^2}\dots x^{\varphi^{p-1}} = 1$. We can illustrate the methodological advantages arising by comparing the result of Cody [16] with [56]. Cody proved that if a finite p-group P is metabelian and distinct from its Hughes subgroup (that is, admits a partition) then $H_p(P)$ is nilpotent of class $\leq p$. In [56] we proved, in particular, that, if in a finite p-group $P \neq H_p(P)$ and $H_p(P)$ is metabelian, then $H_p(P)$ is nilpotent of class $\leq p$. Comparing with Cody's theorem, we see that the hypothesis here is weaker while the conclusion is the same. However, more importantly, the result of [56] may be stated as follows: metabelian groups in $\mathfrak{M}_p$ are nilpotent of class $\leq p$. This more "categorical" formulation in [56] permits the use of induction on derived length to prove the nilpotency of soluble groups in $\mathfrak{M}_p$ – here we can refer to Corollary 2.3.4 or Theorem 2.3.5. (In Chapter 6 this result on nilpotency of soluble groups in $\mathfrak{M}_p$ was obtained as Corollary 6.4.2 to a more general theorem on varieties of operator groups.)

Although the nilpotency class of a d-generator nilpotent group in $\mathfrak{M}_p$ is bounded by a function depending on d and p only, there is an example showing that there is no analogous bound for the order (even if we restrict ourselves to finite p-groups in $\mathfrak{M}_p$, that is, even if the semidirect product is nilpotent). Namely, let

$$G = \langle a_0\rangle \times \langle a_1\rangle \times \dots \times \langle a_{p-1}\rangle$$

be a direct product of p copies of the infinite cyclic group and let φ be the automorphism of G cyclically permuting the factors: $a_i^{\varphi} = a_{i+1}$, where $i+1$ is a residue modulo p. It is easy to see that φ induces a splitting automorphism of order p of $G/\langle a_0a_1\dots a_{p-1}\rangle$ and hence induces a splitting automorphism of $P = G/(G^{p^s}\cdot\langle a_0a_1\dots a_{p-1}\rangle)$ for each $s \in \mathbb{N}$. It is easy to calculate that $|P| = p^{s(p-1)}$.

To conclude this section, we prove that the notions of splitting and regular automorphisms coincide for finite p'-groups.

7.1.3 Lemma. a) *If G is a finite p'-group admitting a splitting automorphism φ of prime order p then φ is a regular automorphism of G.*

b) *If φ is a regular automorphism of prime order p of a finite group G then φ is a splitting automorphism of G.*

Proof. a) If $x^{\varphi} = x$ then

$$1 = x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \dots \cdot x^{\varphi^{p-1}} = x^p,$$

whence $x = 1$ since x is a p'-element.

b) By a well-known formula (see, for example, Lemma 2.4.3) we have

$$|\{x^{-1}x^{\varphi} \mid x \in G\}| = |G : C_G(\varphi)| = |G|,$$

since $C_G(\varphi) = 1$ by hypothesis. Hence, since G is finite, these sets coincide: $\{x^{-1}x^{\varphi} \mid x \in G\} = G$. Therefore, for every $a \in G$, there is an $x \in G$ such that $a = x^{-1}x^{\varphi}$ whence

$$a \cdot a^{\varphi} \cdot a^{\varphi^2} \cdot \ldots \cdot a^{\varphi^{p-1}} = (x^{-1}x^{\varphi}) \cdot (x^{-1}x^{\varphi})^{\varphi} \cdot (x^{-1}x^{\varphi})^{\varphi^2} \cdot \ldots \cdot (x^{-1}x^{\varphi})^{\varphi^{p-1}} = 1.$$

The lemma is proved.

§ 7.2 The Restricted Burnside Problem for groups with a splitting automorphism of prime order

The purpose of this section is another proof of the following theorem.

7.2.1 Theorem. *For every prime p and every natural d there exists a number $f(d, p)$ such that the nilpotency class of any d-generator nilpotent group admitting a splitting automorphism of order p is not greater than $f(d, p)$.*

The scheme of this new proof is as follows. First of all there is an easy reduction to the case of a d-generator finite p-group P with a splitting automorphism φ of order p. As we have already said, Kostrikin's Theorem 1.3.1 on $(p-1)$-Engel Lie algebras is used. Note, however, that the associated Lie ring of a finite p-group from $\mathfrak{M}_p$ can be non-$(p-1)$-Engel when $p \geq 5$, as shown by the examples constructed in [57] using the existence of non-$(p-1)$-Engel identities in the associated Lie ring of a free group of prime exponent p.

Nevertheless, we show that the associated Lie ring of the semidirect product $P\langle\varphi\rangle$ satisfies "almost all" consequences of the $(p-1)$-Engel identity. More exactly, if this Lie ring is represented as the factor-ring L/A of a free $(d+1)$-generator Lie ring (over $\mathbb{Z}$) by an ideal A, then A contains pL and "almost all" of the $(p-1)$-Engel ideal $E = {}_{id}([x, \underbrace{y, \ldots, y}_{p-1}] \mid x, y \in L\rangle$. (Note that it is impossible to prove the inclusion $A \supseteq E + pL$ since the order of P cannot be bounded.) Next, we consider the "trace" of the subgroup $C_P(\varphi)$ in L/A – a subring L_1/A corresponding to $C_P(\varphi)$ which, although not the associated Lie ring of $C_P(\varphi)$, has the same order as $C_P(\varphi)$. Applying Kostrikin's Theorem 1.3.1 to the $(p-1)$-Engel Lie algebra $L/(E + pL)$ and using the information about $A \cap E$, we can

bound the order of L_1/A by a function of d and p. Next, by the Alperin-Khukhro Theorem 5.2.1, the bound on $|C_P(\varphi)|$ implies a bound on the derived length of P in terms of d and p. We finally need only apply Corollary 6.4.2 on the nilpotency of soluble groups from $\mathfrak{M}_p$. In addition, Theorem 5.2.1 allows us to assert that a finite d-generator p-group $P \in \mathfrak{M}_p$ also contains a subgroup of (d, p)-bounded index which is nilpotent of class $h(p)$.

We now proceed with a more detailed exposition.

Proof of Theorem 7.2.1. Suppose that G is a d-generator nilpotent group admitting a splitting automorphism φ of prime order p. If the k elements $a_1, a_2, \ldots a_k$ generate G as a $\langle\varphi\rangle$-group then the pk elements $a_i, a_i^{\varphi}, a_i^{\varphi^2}, \ldots, a_i^{\varphi^{p-1}}$, $i = 1, 2, \ldots, k$, generate G as an abstract group. Therefore, in order to prove the theorem we may assume that G is d-generator as a $\langle\varphi\rangle$-group.

It is not difficult to reduce to the case of finite p-groups, as follows. It is well-known that finitely generated nilpotent groups are residually finite (see, for example, [123]). If $\{N_\alpha\}$ is a family of normal subgroups of finite index in G with trivial intersection then $\left\{\bigcap_{i=1}^{p} N_\alpha^{\varphi^i}\right\}$ is a family of φ-invariant normal subgroups of finite index in G with trivial intersection. It is clearly sufficient to obtain the appropriate bound on the nilpotency class for each of the factor-groups $G/\bigcap_{i=1}^{p} N_\alpha^{\varphi^i}$. We may therefore assume that G is finite.

The Hall p'-subgroup $G_{p'}$ of the nilpotent group G is normal and φ-invariant. By Lemma 7.1.3 φ is regular on $G_{p'}$ so that $G_{p'}$ is nilpotent of class $\leq h(p)$, where $h(p)$ is Higman's function, by Higman's Theorem 5.1.1. If G_p is the Sylow p-subgroup of G then $G = G_p \times G_{p'}$. Hence G_p is a homomorphic image of the d-generator group G (by the natural homomorphism with kernel $G_{p'}$) and hence it is also d-generator. It is therefore sufficient to prove the theorem for a d-generator finite p-group and we shall consider this case in what follows.

In this situation the semidirect product $G\langle\varphi\rangle$ is also a finite p-group which is nilpotent of some nilpotency class n. Hence, $G\langle\varphi\rangle$ may be represented as the image of the free $(d+1)$-generator nilpotent group F of class n with free generators $y, x_1, x_2, \ldots, x_d$ under the homomorphism ϑ which extends the mapping of $x_1, x_2, \ldots, x_d$ onto the generators of G and of y onto φ. Clearly, the full inverse image of G under this homomorphism is the normal closure $\langle\{x_1, x_2, \ldots, x_d\}^F\rangle$ of $x_1, x_2, \ldots, x_d$. We put

$$N = \langle (x \cdot y^{-1})^p \mid x \in \langle\{x_1, x_2, \ldots, x_d\}^F\rangle\rangle.$$

It is easy to see that N is a normal subgroup of F: for any $g \in F$ we have

$$((x \cdot y^{-1})^p)^g = ((x \cdot y^{-1})^g)^p = (x^g \cdot (y^{-1})^g)^p = (x^g \cdot [g, y] \cdot y^{-1})^p$$

and $x^g[g, y] \in \langle\{x_1, x_2, \ldots, x_d\}^F\rangle$ since $F' \leq \langle\{x_1, x_2, \ldots, x_d\}^F\rangle$.

7.2.2 Lemma. *The homomorphism ϑ induces a homomorphism of F/N onto $G\langle\varphi\rangle$, that is, $N \leq \operatorname{Ker}\vartheta$.*

Proof. For elements generating N we have by (7.1.2)

$$\vartheta((x \cdot y^{-1})^p) = (\vartheta(x) \cdot \vartheta(y^{-1}))^p = (\vartheta(x) \cdot \varphi^{-1})^p =$$
$$\vartheta(x) \cdot \vartheta(x)^{\varphi} \cdot \vartheta(x)^{\varphi^2} \cdot \ldots \cdot \vartheta(x)^{\varphi^{p-1}} \cdot \varphi^{-p} = 1,$$

since $\vartheta(x) \in G$, and φ is a splitting automorphism of order p of G.

The lemma is proved.

7.2.3 Lemma. *The group F/N is a finite p-group.*

Proof. This group is nilpotent and is generated by $d + 1$ elements of order p, the images of $y^{-1}, x_1 y^{-1}, x_2 y^{-1}, \ldots, x_d y^{-1}$. Hence all factors of its lower central series are finite groups of exponent p (see Corollaries 2.5.4 and 2.5.6).

The lemma is proved.

Actually, F/N is a universal object analogous to a free group of a variety. In order to prove the theorem, it is sufficient by Lemma 7.2.2 to obtain a bound for the nilpotency class of the image in F/N of the normal closure $\langle\{x_1, x_2, \ldots, x_d\}^F\rangle$ which is the inverse image of G; of course, this bound must depend on d and p only.

We consider the associated Lie ring $L(F/N)$ of F/N. It is also nilpotent of class n and is generated by $d + 1$ elements, the images in the factor-group by the commutator subgroup of the generators of F/N which, in turn, are the images of $y, x_1, x_2, \ldots, x_d$ (see 3.2.2 and 3.2.4). Let L be a free n-generator nilpotent Lie ring of nilpotency class n (over $\mathbb{Z}$) with free generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$. There is a Lie ring homomorphism of L onto $L(F/N)$ which extends the mapping of the $\eta, \xi_1, \ldots, \xi_d$ onto the generators of $L(F/N)$, the aforementioned images of images of the $y, x_1, \ldots, x_d$. The kernel A of this homomorphism is called the ideal of relations of the Lie ring $L/A \cong L(F/N)$. The ideal A is homogeneous; for every $k \in \mathbb{N}$, its homogeneous component of weight k is generated by homogeneous elements of L of weight k which have the form $\sum_i \alpha_i c_i$ where $\alpha_i \in \mathbb{Z}$ and c_i are commutators of weight k in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ such that the product $\prod_i \tilde{c}_i^{\alpha_i}$ of the group commutators $\tilde{c}_i$ with the same bracket structure in the generators $y, x_1, x_2, \ldots, x_d$ belongs to $N \cdot \gamma_{k+1}(F)$ (here the order of the factors in the product is irrelevant since commutators of weight k commute modulo $\gamma_{k+1}(F)$).

Kostrikin's Theorem 1.3.1 states that the nilpotency class and, hence, also the order of the $(p-1)$-Engel $(d+1)$-generator Lie algebra $L/(pL+E)$ over $GF(p)$, is bounded in terms of d and p. In order to be able to use this, we shall show that the ideal A contains pL and the larger part of the $(p-1)$-Engel ideal

$$E = {}_{id}\langle [u, \underbrace{v, v, \ldots, v}_{p-1}] \mid u, v \in L \rangle.$$

Firstly we find generators of the additive group of E modulo pL.

7.2.4 Lemma. *The additive group of E is generated modulo pL by elements of the form*

$$\ll u, v_1, v_2, \ldots, v_{p-1} \gg = \sum_{\pi \in \mathbb{S}_{p-1}} [u, v_{\pi(1)}, v_{\pi(2)}, \ldots, v_{\pi(p-1)}],$$

where $u, v_1, v_2, \ldots, v_{p-1} \in L$.

(The elements $\ll u, v_1, v_2, \ldots, v_{p-1} \gg$ are called *Kostrikin elements.*)

Proof. In a Lie algebra of characteristic $p > 0$, the identity $[u, \underbrace{v, v, \ldots, v}_{p-1}] = 0$ is equivalent by Lemma 3.3.1 to the identity

$$\sum_{\pi \in \mathbb{S}_{p-1}} [u_0, v_{\pi(1)}, v_{\pi(2)}, \ldots, v_{\pi(p-1)}] = 0.$$

Hence E is generated modulo pL (as a Lie algebra ideal) by elements of the form

$$\ll u_0, u_1, u_2, \ldots, u_{p-1} \gg = \sum_{\pi \in S_{p-1}} [u_0, u_{\pi(1)}, u_{\pi(2)}, \ldots, u_{\pi(p-1)}],$$
$$u_0, u_1, u_2, \ldots, u_{p-1} \in L.$$

For any $l \in L$ and $\pi \in \mathbb{S}_{p-1}$ by the Jacobi identity $[[a, b], c] = [[a, c], b] + [a, [b, c]]$ we have

$$\begin{aligned}
&[[u_0, u_{\pi(1)}, u_{\pi(2)}, \ldots, u_{\pi(p-1)}], l] = \\
&= [[[u_0, u_{\pi(1)}, u_{\pi(2)}, \ldots, u_{\pi(p-2)}], l], u_{\pi(p-1)}] + \\
&+ [[u_0, u_{\pi(1)}, u_{\pi(2)}, \ldots, u_{\pi(p-2)}], [u_{\pi(p-1)}, l]] = \ldots \\
&\ldots = \sum_{i=0}^{p-1} [u_0, u_{\pi(1)}, u_{\pi(2)}, \ldots, [u_{\pi(i)}, l], \ldots, u_{\pi(p-1)}].
\end{aligned}$$

Therefore summation gives

$$[\ll u_0, u_1, u_2, \ldots, u_{p-1}\gg, l] =$$

$$= \sum_{\pi\in\mathbb{S}_{p-1}} \sum_{i=0}^{p-1} [u_0, u_{\pi(1)}, u_{\pi(2)}, \ldots, [u_{\pi(i)}, l], \ldots, u_{\pi(p-1)}] =$$

$$= \sum_{i=0}^{p-1} \sum_{\pi\in\mathbb{S}_{p-1}} [v_{i.0}, v_{i.\pi(1)}, v_{i.\pi(2)}, \ldots, v_{i.\pi(p-1)}] =$$

$$= \sum_{i=0}^{p-1} \ll v_{i.0}, v_{i.1}, v_{i.2}, \ldots, v_{i.p-1}\gg,$$

where for each $i = 0, 1, 2, \ldots, p-1$

$$(v_{i.0}, v_{i.1}, v_{i.2}, \ldots, v_{i.p-1}) = (u_0, u_1, u_2, \ldots, [u_i, l], \ldots, u_{p-1}).$$

So, for every Kostrikin element $\ll u_0, u_1, \ldots, u_{p-1}\gg$ and for any $l \in L$, the commutator $[\ll u_0, u_1, \ldots, u_{p-1}\gg, l]$ is equal to a linear combination of Kostrikin elements. Hence the additive group of E is generated by Kostrikin elements.

The lemma is proved.

The following proposition (Theorem 1 of [75]) will substantially simplify the notation in what follows.

7.2.5 Lemma. *Kostrikin elements are symmetric modulo pL, that is,*

$$\ll u_1, u_2, \ldots, u_p\gg \equiv \ll u_{\pi(1)}, u_{\pi(2)}, \ldots, u_{\pi(p)}\gg \pmod{pL}$$

for any $u_1, u_2, \ldots, u_p \in L$ and for any permutation $\pi \in \mathbb{S}_p$.

Proof. Since the Kostrikin elements $\ll u_1, u_2, \ldots, u_p\gg$ are invariant under permutations of the elements $u_2, \ldots, u_p$ by definition, it is sufficient to prove that

$$\ll u_1, u_2, u_3, \ldots, u_p\gg \equiv \ll u_2, u_1, u_3, \ldots, u_p\gg \pmod{pL}.$$

We first establish this congruence in the case where $u_3 = \ldots = u_p = u$, and then apply "linearization" afterwards.

Recall the notation $[a, {}_k b] = [a, \underbrace{b, b, \ldots, b}_{k}]$. Note that

$$[a, [b, {}_n c]] = \sum_{i=0}^{n} (-1)^i C_n^i [[[a, {}_i c], b], {}_{n-i} c].$$

This formula is easily deduced from the Jacobi identity by induction. We apply it to obtain

$$\ll u_1, u_2, \underbrace{u, \ldots, u}_{p-2} \gg = (p-2)! \sum_{n=0}^{p-2} [[[u_1, {}_n u], u_2], {}_{p-2-n} u] =$$

$$= -(p-2)! \sum_{n=0}^{p-2} [[u_2, [u_1, {}_n u]], {}_{p-2-n} u] =$$

$$= (p-2)! \sum_{n=0}^{p-2} \sum_{i=0}^{n} (-1)^{i+1} C_n^i [[[u_2, {}_i u], u_1], {}_{p-2-i} u] =$$

$$= (p-2)! \sum_{i=0}^{p-2} (-1)^{i+1} \left(\sum_{n=i}^{p-2} C_n^i \right) [[[u_2, {}_i u], u_1], {}_{p-2-i} u] \equiv$$

$$\equiv \ll u_2, u_1, \underbrace{u, \ldots, u}_{p-2} \gg \pmod{pL},$$

since

$$(-1)^{i+1} \sum_{n=i}^{p-2} C_n^i \equiv (-1)^{i+1} C_{p-1}^{i+1} \equiv 1 \pmod{p}.$$

We now put $u = u_3 + \ldots + u_p$ in

$$\ll u_1, u_2, \underbrace{u, \ldots, u}_{p-2} \gg \equiv \ll u_2, u_1, \underbrace{u, \ldots, u}_{p-2} \gg \pmod{pL}.$$

Now let $u_1, u_2, \ldots, u_p$ be the free generators of a free Lie ring. Then this congruence implies the congruence of its multilinear components (which is the desired congruence). Since the u_i are free generators, the same congruence holds for arbitrary elements of any Lie ring.

The lemma is proved.

7.2.6 Lemma. *The ideal A contains pL.*

Proof. It is sufficient to prove that $pc \in A$ for any commutator c in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L. Let $\tilde{c}$ denote the commutator with the same bracket structure in the generators $y, x_1, x_2, \ldots, x_d$ of F and let k be its weight. If $\tilde{c} \in \langle \{x_1, x_2, \ldots, x_d\}^F \rangle$ then $(\tilde{c} \cdot y^{-1})^p$ and y^p belong to N by definition. Hence N contains their product $(\tilde{c} \cdot y^{-1})^p \cdot y^p$ which is congruent to $\tilde{c}^p$ modulo $\langle [\tilde{c}, y] \rangle$. But $\langle [\tilde{c}, y] \rangle$ is clearly contained in $\gamma_{k+1}(F)$, so that $\tilde{c}^p \in N \cdot \gamma_{k+1}(F)$. This means

that $pc \in A$. If, however, $\tilde{c} \notin \langle\{x_1, x_2, \ldots, x_d\}^F\rangle$, then $\tilde{c} = y$ (recall that $\tilde{c}$ is a commutator in $y, x_1, x_2, \ldots, x_d$), but $y^p \in N$ whence again $pc \in A$.

The lemma is proved.

The following main technical proposition is an analogue of the Magnus-Sanov Theorem 3.3.2. Its proof is similar, although it requires somewhat more complicated arguments. Those parts of the proof of Theorem 3.3.2 which fit here without much alteration will simply be quoted. It is worth noting that the original proof [61] was short since it used properties of the Baker-Hausdorff formula; here, since we are aiming for a more self-contained exposition, we cannot use this short cut.

Further we shall consider only Kostrikin elements $\ll u_0, u_1, \ldots, u_{p-1}\gg$ involving commutators u_i in generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L. We shall denote by $\ll \tilde{u}_0, \tilde{u}_1, \ldots, \tilde{u}_{p-1}\gg$ a product of commutators

$$\prod_{\pi \in \mathbb{S}_{p-1}} [\tilde{u}_0, \tilde{u}_{\pi(1)}, \tilde{u}_{\pi(2)}, \ldots, \tilde{u}_{\pi(p-1)}]$$

involving subcommutators $\tilde{u}_i$ corresponding to Lie ring commutators u_i, that is, $\tilde{u}_i$ is a group commutator with the same bracket structure as u_i in the generators $y, x_1, x_2, \ldots, x_d$ of F. The order of the factors here is irrelevant because these products will appear only in congruences where the factors commute modulo the corresponding normal subgroups.

7.2.7 Proposition. *Every Kostrikin element involving commutators in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L which is not of the form $\ll a, \underbrace{\eta, \ldots, \eta}_{p-1}\gg$ belongs to A, that is, $\ll u_1, u_2, \ldots, u_p\gg \in A$ provided $\{u_1, u_2, \ldots, u_p\} \neq \{a, \underbrace{\eta, \ldots, \eta}_{p-1}\}$.*

Proof. The difficulties which make the proof of this proposition different from the proof of Theorem 3.3.2 are connected with the special role of the generator y and with an asymmetry in the definition of N, which may be said to be "verbal" only with respect to the generators $x_1, x_2, \ldots, x_d$. In particular, applying homomorphisms to congruences modulo N, one has to take care that N is invariant under them.

For convenience we shall write

$$\ll u_1, u_2, \ldots, u_s, (p-s)\eta\gg = \frac{1}{(p-s)!}\ll u_1, u_2, \ldots, u_s, \underbrace{\eta, \eta, \ldots, \eta}_{p-s}\gg, \quad s \geq 1,$$

where u_i are commutators (in the generators $\eta, \xi_1, \ldots, \xi_d$ of L) ***different from*** η (the case of $s = p$ is not excluded). Here the coefficient $\frac{1}{(p-s)!}$ denotes an integer

whose image in the multiplicative group of $\mathbb{Z}/p\mathbb{Z}$ is inverse to the image of $(p-s)!$ (this is justified by the fact that L/A is a Lie algebra over $GF(p)$ by Lemma 7.2.6). In an analogous way we denote the corresponding product of group commutators by $\ll\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_s, (p-s)y\gg$.

For technical reasons, we shall need the set R_x of all commutators in the generators $y, x_1, x_2, \ldots, x_d$ of F which involve at least two equal elements of the set $\{x_1, x_2, \ldots, x_d\}$.

7.2.8 Lemma. a) *The subgroup $\langle R_x\rangle$ is a normal subgroup of F and each of its elements may be written in the form*

$$c_2 \cdot c_3 \cdot \ldots \cdot c_n, \tag{7.2.9}$$

where, for each i, the element c_i is a product of powers of commutators of weight i from R_x.

b) *If c is a commutator from R_x in $y, x_1, x_2, \ldots, x_d$ which has weight i_0 in y and weight i_s in x_s $(s = 1, 2, \ldots, d)$ then the commutator obtained from c by replacing y by y^μ and x_s by $x_s^{\nu_s}$ for $\mu, \nu_s \in \mathbb{N}$ is equal to $c^\kappa \cdot r$ where $\kappa = \mu^{i_0} \cdot \nu_1^{i_1} \cdot \nu_2^{i_2} \cdot \ldots \cdot \nu_d{}^{i_d}$ and r is a product of powers of commutators from R_x each of which has weight greater than that of c.*

Proof. a) If $c \in R_x$ then $[c, b] \in R_x$ for any $b \in \{y, x_1, x_2, \ldots, x_d\}$ so that $\langle R_x\rangle \trianglelefteq F$. Since commutators in commutators from R_x also lie in R_x, any product of powers of commutators from R_x may be transformed by formulae of the collecting process 2.7.5 to the required form (7.2.9). It is also clear that commutators from R_x have weight at least 2 (or even 3 since $[x_i, x_i] = 0$).

b) This follows from formulae 2.1.1 c), d).

The lemma is proved.

We return to the proof of the proposition. First, we prove that for every $s = 2, 3, \ldots, p$ the Kostrikin element $\ll\xi_1, \xi_2, \ldots, \xi_s, (p-s)\eta\gg$ belongs to A. We start with the element $(x_1x_2\ldots x_py)^p \in N$ and apply to the product

$$\underbrace{(x_1x_2\ldots x_py)(x_1x_2\ldots x_py)\ldots(x_1x_2\ldots x_py)}_{p}$$

a collecting process of the following form: we move all elements different from x_1 to the left preserving the order of their occurrence by the formula

$$[x_1, a_{i_1}, \ldots, a_{i_r}] \cdot b = b \cdot [x_1, a_{i_1}, \ldots, a_{i_r}] \cdot [x_1, a_{i_1}, \ldots, a_{i_r}, b],$$

where $b, a_{i_k} \in \{x_2, x_3, \ldots x_p, y\}$. So, $x_2, x_3, \ldots, x_p, y$ are transferred only by occurrences of x_1 and by commutators of the form $[x_1, a_{i_1}, \ldots, a_{i_r}]$ appearing at preceding steps. As a result, we obtain the product

$$(x_2 x_3 \ldots x_p y)^p \cdot \prod [x_1, a_{i_1}, a_{i_2}, \ldots, a_{i_r}],$$

where the product is taken over all commutators appearing (in the order in which they appear); here the case of $r = 0$ is not excluded, where $[x_1, a_{i_1}, \ldots, a_{i_r}] = x_1$ by definition.

Since $(x_2 x_3 \ldots x_p y)^p \in N$, the second factor also lies in N. We rewrite this fact as

$$\prod [x_1, a_{i_1}, a_{i_2}, \ldots, a_{i_r}] \equiv 1 \pmod{N}. \tag{7.2.10}$$

The number of occurrences of a given commutator $[x_1, a_{i_1}, \ldots, a_{i_r}]$ on the left of (7.2.10) may be calculated. Such a calculation is especially easy if it is conducted modulo p.

7.2.11 Lemma. a) *If the weight of* $[x_1, a_{i_1}, \ldots, a_{i_t}]$, *where* $a_{i_j} \in \{x_2, x_3, \ldots, x_p, y\}$ *is less than* p, *then the number of its occurrences in (7.2.10) is a multiple of* p.

b) *If the weight of* $[x_1, a_{i_1}, \ldots, a_{i_t}]$, *where* $a_{i_j} \in \{x_2, x_3, \ldots, x_p, y\}$ *is* p, *then the number of its occurrences in (7.2.10) is congruent to* 1 *modulo* p.

Proof. It follows from the description of the collecting process which produced the product (7.2.10) that the commutator $[x_1, a_{i_1}, \ldots, a_{i_t}]$, $a_{i_j} \in \{x_2, x_3, \ldots, x_p, y\}$, occurs in (7.2.10) exactly the same number of times as the subsequence $x_1, a_{i_1}, \ldots, a_{i_t}$ occurs in the sequence

$$\underbrace{(x_1, x_2, \ldots, x_p, y), (x_1, x_2, \ldots, x_p, y), \ldots, (x_1, x_2, \ldots, x_p, y)}_{p}. \tag{7.2.12}$$

We partition the subsequence by brackets according to the brackets in (7.2.12) (putting into a pair of brackets the segment of the subsequence which is contained in one bracket in (7.2.12)). For example, for $p = 5$, for the underlined subsequence in

$$\begin{gathered}(x_1, x_2, x_3, x_4, x_5, y), (\underline{x_1}, x_2, \underline{x_3}, x_4, x_5, y), (x_1, x_2, x_3, x_4, \underline{x_5}, \underline{y}), \\ (x_1, x_2, x_3, x_4, x_5, y), (x_1, \underline{x_2}, x_3, x_4, x_5, y)\end{gathered} \tag{7.2.13}$$

we obtain the following bracket structure: $(x_1, x_3), (x_5, y), (x_2)$. (Here, of course, not every bracket structure can appear: it is easy to see that a necessary and sufficient condition is that the order of elements within a pair of brackets in the subsequence matches the order $x_1 < x_2 < \ldots < x_p < y$.)

Thus the set of all subsequences in (7.2.12) which are equal to a given subsequence is partitioned into subsets which differ according to different bracket structures. The cardinalities of these subsets may be easily calculated modulo p. For a given bracket structure in the subsequence $x_1, a_{i_1}, \ldots, a_{i_t}$ (assuming k pairs of brackets) the occurrence of this subsequence in (7.2.12) is uniquely determined by the choice of those k brackets in (7.2.12) from which the elements of the corresponding brackets in the subsequence $x_1, a_{i_1}, \ldots, a_{i_t}$ are chosen. The existence and uniqueness of such an occurrence follows from the fact that each bracket in (7.2.12) contains exactly one of each of the elements $x_1, x_2, \ldots, x_p, y$. For example, if we choose the 1-st, 2-nd and 3-rd brackets in (7.2.13) (instead of the 2-nd, 3-rd and 5-th ones) then we obtain for $(x_1, x_3), (x_5, y), (x_2)$ the uniquely determined occurrence

$$(\underline{x_1}, x_2, \underline{x_3}, x_4, x_5, y), (x_1, x_2, x_3, x_4, \underline{x_5}, \underline{y}), (x_1, \underline{x_2}, x_3, x_4, x_5, y),$$
$$(x_1, x_2, x_3, x_4, x_5, y), (x_1, x_2, x_3, x_4, x_5, y).$$

Therefore, the cardinality of the subset of subsequences which are equal to a given one and have k pairs of brackets is equal to C_p^k. If $k < p$ then this number is divisible by p. If the weight of $[x_1, a_{i_1}, \ldots, a_{i_t}]$ is less than p then the number of brackets is certainly less than p for any admissible bracket structure. Hence the number of occurrences of this commutator in (7.2.10) is equal to a sum of certain C_p^k with $k < p$ and hence is a multiple of p.

If, however, the weight of $[x_1, a_{i_1}, \ldots, a_{i_{p-1}}]$ is equal to p then there is only one admissible bracket structure which gives a contribution not divisible by p to the multiplicity of the occurrence in (7.2.10) – namely, when the number of brackets is p: $(x_1), (a_{i_1}), (a_{i_2}), \ldots, (a_{i_{p-1}})$. Hence the number of occurrences of this commutator in (7.2.10) is congruent to 1 modulo p.

The lemma is proved.

Although at the moment our aim is to prove that

$$\ll x_1, x_2, \ldots, x_s, (p-s)y \gg \equiv 1 \pmod{N \cdot \gamma_{p+1}(F)} \tag{7.2.14}$$

for every $s \geq 2$, for technical reasons we switch to calculations modulo $\langle R_x \rangle \cdot N \cdot \gamma_{p+1}(F)$. So we first prove that

$$\ll x_1, x_2, \ldots, x_s, (p-s)y \gg \equiv 1 \pmod{\langle R_x \rangle \cdot N \cdot \gamma_{p+1}(F)}. \tag{7.2.15}$$

Since all commutators in (7.2.10) involve x_1, they commute modulo $\langle R_x \rangle$.

Note that by Lemma 7.2.11 the left-hand side of (7.2.15) is the multihomogeneous component of (7.2.10) of weight 1 in $x_i, i = 1, 2, \ldots, s$, and of weight $p - s$ in y (this means that the product of all commutators in (7.2.10) of weight

1 in $x_i, i = 1, 2, \ldots, s$, and of weight $p - s$ in y is equal to the left-hand side of (7.2.15), the order of commutators here being irrelevant as they commute modulo $\langle R_x \rangle$).

We rewrite (7.2.10) as a congruence modulo $\langle R_x \rangle \cdot N \cdot \gamma_{p+1}(F)$:

$$\prod [x_1, a_{i_1}, a_{i_2}, \ldots, a_{i_r}] \equiv 1 \pmod{\langle R_x \rangle \cdot N \cdot \gamma_{p+1}(F)} \tag{7.2.16}$$

and apply arguments similar to those used in the process of excluding commutators of smaller weight which appeared in the proof of Theorem 2.8.11. Namely, we apply to (7.2.16) the homomorphism σ of F which extends the mapping

$$x_i \to x_i^v \text{ for all } i, \quad y \to y^v$$

where $v \in \mathbb{N}$ is such that its image in $\mathbb{Z}/p\mathbb{Z}$ has order $p - 1$ with respect to multiplication (that is, it generates the multiplicative group of the field $GF(p) \cong \mathbb{Z}/p\mathbb{Z}$). Note that all subgroups $\langle R_x \rangle$, N, $\gamma_{p+1}(F)$ are σ-invariant. Then, taking an appropriate power v^{-k} of (7.2.16) and multiplying the congruences obtained, we shall get rid of commutators of weight $k \leq p - 1$ consecutively. Since $\gamma_p(F)^p \leq N \cdot \gamma_{p+1}(F)$ by Lemma 7.2.6, the exponents of powers of commutators of weight p in congruences modulo $\langle R_x \rangle \cdot N \cdot \gamma_{p+1}(F)$ may be regarded as residues modulo p.

Here, however, there are two obstacles which lead to differences between this and the proof of Theorem 2.8.11 which will be described now. First, it is important to preserve (modulo p) the multihomogeneous component of weight 1 in $x_i, i = 1, 2, \ldots, s$, and of weight $p - s$ in y. But applying formulae 2.1.1 c), d) to commutators of weight $< p$ may lead not only to the emergence of commutators from R_x but also to commutators of weight 1 in $x_i, i = 1, 2, \ldots, s$, and of weight $p - s$ in y. However, Lemma 7.2.11 a) and the fact that all commutators containing x_1 commute modulo $\langle R_x \rangle$ guarantee that additional factors from this multihomogeneous component lie in $\gamma_p(F)^p \leq N \cdot \gamma_{p+1}(F)$.

Secondly, here we cannot get rid of commutators of weight 1, that is, of x_1^p, simply by multiplying (7.2.16) by x_1^{-p} since this element does not lie in N. (Also, application of the homomorphism σ does not give the proper result for weight 1 since $v^p - v^1 \equiv 0 \pmod{p}$.) Instead, we multiply (7.2.16) from the outset by the element $(x_1^{-1}y)^p \cdot y^{-p}$ which lies in N. Since $(x_1^{-1}y)^p \cdot y^{-p} \equiv x_1^{-p} \pmod{\gamma_2(\langle x_1, y \rangle)}$, we get rid of commutators of weight 1. All additional factors will be powers of commutators in x_1 and y only and hence they do not change the multihomogeneous component of (7.2.10) of weight 1 in $x_i, i = 1, 2, \ldots, s$, and of weight $p - s$ in y (we remind the reader that $s \geq 2$).

As a result we obtain

$$\varkappa \cdot \ll x_1, x_2, \ldots, x_s, (p - s)y \gg \equiv 1 \pmod{\langle R_x \rangle \cdot N \cdot \gamma_{p+1}(F)}, \tag{7.2.17}$$

where $\varkappa$ denotes a product of powers of commutators of weight p which do not belong to the multihomogeneous component of weight 1 in $x_i, i = 1, 2, \ldots, s$, and of weight $p - s$ in y.

Transition from (7.2.17) to (7.2.15) is carried out by means of linearization in the variables $x_1, x_2, \ldots, x_s$, in practically the same way as in the proof of Theorem 2.8.11.

We apply to (7.2.17) the homomorphisms ϑ_j which extend, respectively, the mappings

$$y \to y, \quad x_j \to 1 \quad \text{and} \quad x_i \to x_i \text{ for } i \neq j$$

and leave the subgroups $\langle R_x \rangle$, N, $\gamma_{p+1}(F)$ invariant. The image of the left-hand side of (7.2.17) under ϑ_j is the product of all powers of commutators from (7.2.17) which do not depend on the x_j. Hence, firstly, all powers of commutators which do not involve any of the $x_1, x_2, \ldots, x_s$ may be dropped in (7.2.17), and, secondly, all powers of commutators involving any of the $x_{s+1}, x_{s+2}, \ldots, x_p$ may also be dropped in (7.2.17). As a result, the product of the powers of commutators from (7.2.17) which have weight 1 in each of the $x_1, x_2, \ldots, x_s$ and weight 0 in each of the $x_{s+1}, x_{s+2}, \ldots, x_p$ is congruent to 1. Since the total weight of each commutator is p, all of them also have weight $p - s$ in y. Hence (7.2.15) holds.

Using Lemma 7.2.8 a), we rewrite (7.2.15) as a congruence modulo $N \cdot \gamma_{p+1}(F)$:

$$c_2 \cdot c_3 \cdot \ldots \cdot c_p \cdot \ll x_1, x_2, \ldots, x_s, (p-s)y \gg \equiv 1 \pmod{N \cdot \gamma_{p+1}(F)}, \quad (7.2.18)$$

where, for every i, c_i is a product of powers of commutators from R_x of weight i. We apply to (7.2.18) the process of excluding commutators of smaller weight which was described in the proof of Theorem 2.8.11 and which was used just above. Now we have no difficulties either with weight 1 or with preserving the multihomogeneous component of weight 1 in $x_i, i = 1, 2, \ldots, s$, and of weight $p - s$ in y since here both all commutators of weight $\leq p - 1$ and all new additional factors appearing belong to $\langle R_x \rangle$ by Lemma 7.2.8.

As a result, we obtain

$$\varkappa \cdot \ll x_1, x_2, \ldots, x_s, (p-s)y \gg \equiv 1 \pmod{N \cdot \gamma_{p+1}(F)},$$

where $\varkappa$ denotes a product of powers of commutators of weight p which do not belong to the multihomogeneous component of weight 1 in $x_i, i = 1, 2, \ldots, s$, and of weight $p - s$ in y. Linearization in the variables $x_1, x_2, \ldots, x_s$, exactly as applied above to (7.2.17), transforms this congruence into (7.2.14), as required.

We have therefore proved that $\ll \xi_1, \xi_2, \ldots, \xi_s, (p-s)\eta \gg \in A$. Now our task is to replace the $\xi_1, \xi_2, \ldots, \xi_s$ here by arbitrary commutators $u_1, u_2, \ldots, u_s$ (distinct from η) in the variables $\eta, \xi_1, \xi_2, \ldots, \xi_d$.

We rewrite (7.2.14) as a congruence modulo N

$$\ll x_1, x_2, \ldots, x_s, (p-s)y \gg \equiv c_1^{r_1} \cdot \ldots \cdot c_t^{r_t} \pmod{N}, \tag{7.2.19}$$

where the c_i are commutators in generators $y, x_1, x_2, \ldots, x_d$ from $\gamma_{p+1}(F)$, that is, of weight $\geq p+1$. Before applying to (7.2.19) the homomorphism of F which extends the mapping

$$y \to y, \quad x_i \to \tilde{u}_i \text{ for } i = 1, 2, \ldots, s, \quad x_j \to x_j \text{ for } j > s,$$

we must first use an analogue of Higman's Lemma, as in the proof of Theorem 3.3.2.

7.2.20 Lemma. *A congruence of type (7.2.19) holds where all commutators c_i in the generators $y, x_1, x_2, \ldots, x_d$ have weight $\geq p+1$ and depend on each of the $x_1, x_2, \ldots, x_s$.*

Proof. We cannot simply refer to Corollary 1.10.6 and Lemma 1.10.1, since N is not a verbal subgroup. Instead, we apply arguments similar to the proofs of these results since $\gamma_{p+1}(F)$ and N are invariant under the homomorphisms ϑ_j defined above. If, for example, in (7.2.19) there are commutators among the c_i which do not depend on x_1, then all their powers may be collected at the start of the right-hand side. Application of ϑ_1 then shows that their product is congruent to 1 modulo N since the image of the left-hand side is 1 as also are the images of all commutators depending on x_1. We now repeat the process with respect to x_2, etc.

The lemma is proved.

We assume from now on that (7.2.19) satisfies Lemma 7.2.20.

We now complete the proof of the proposition. Let $u_1, u_2, \ldots, u_s$ be arbitrary commutators in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L which are different from η. If k_i is the weight of $u_i, i = 1, 2, \ldots, s$, then proving that

$$\ll u_1, u_2, \ldots, u_s, (p-s)\eta \gg \in A,$$

is equivalent to establishing

$$\ll \tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_s, (p-s)y \gg \equiv 1 \pmod{N \cdot \gamma_{k_1+k_2+\ldots+k_s+p-s+1}(F)} \tag{7.2.21}$$

where the $\tilde{u}_i$ are the corresponding group commutators in the generators $y, x_1, x_2, \ldots, x_d$.

We apply to (7.2.19) the homomorphism τ of F which extends the mapping

$$y \to y, \quad x_i \to \tilde{u}_i \text{ for } i = 1, 2, \ldots, s, \quad x_j \to x_j \text{ for } j > s.$$

Since the images of the $x_1, x_2, \ldots, x_d$ lie in the normal closure $\langle \{x_1, x_2, \ldots, x_d\}^F \rangle$ and the image of y is y, the subgroup N is τ-invariant. Hence we obtain

$$\ll \tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_s, (p-s)y \gg \equiv \tau(c_1)^{r_1} \cdot \tau(c_2)^{r_2} \cdot \ldots \cdot \tau(c_t)^{r_t} \pmod{N}.$$

We now need only show that $\tau(c_i)$ belongs to $\gamma_{k_1+k_2+\ldots+k_s+p-s+1}(F)$ for every i.

Since $[\gamma_i, \gamma_j] \leq \gamma_{i+j}$, a single occurrence in c_i of each of the x_i, $i = 1, 2, \ldots, s$, contributes k_i to the index of that term of the lower central series of F which contains $\tau(c_i)$. But the weight of c_i is at least $p+1$. Hence other occurrences of the $y, x_1, x_2, \ldots, x_d$, whose total number is at least $p-s+1$, increase that index by at least $p-s+1$. As a result we have

$$\tau(c_i) \in \gamma_{k_1+k_2+\ldots+k_s+p-s+1}(F)$$

for all i. The congruence (7.2.21) follows and hence $\ll u_1, u_2, \ldots, u_s, (p-s)\eta \gg \in A$, as required.

The proposition is proved.

It is clear that, by Lemma 7.2.4, Proposition 7.2.7 gives the following description of generators of the additive group $(E+A)/A$.

7.2.22 Corollary. *The additive group of the $(p-1)$-Engel ideal E is generated modulo the ideal A by the elements of the form $\ll a, (p-1)\eta \gg$, where a is a commutator in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L.*

Recall the notation introduced above:

$$\ll a, (p-1)\eta \gg = [a, \underbrace{\eta, \eta, \ldots, \eta}_{p-1}] = [a, {}_{p-1}\eta].$$

We now apply Kostrikin's Theorem 1.3.1.

7.2.23 Corollary. *The additive group of the $(p-1)$-Engel ideal E is generated modulo the ideal A by the elements of the form*

$$[b, {}_m\eta], \tag{7.2.24}$$

where $m \geq p-1$ and b is a commutator in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L whose weight is bounded by Kostrikin's function $k(d, p)$, depending on d and p only, which bounds the nilpotency class of the $(d+1)$-generator $(p-1)$-Engel Lie algebra $L/(E+pL)$ of characteristic p.

We fix notation $k(d, p)$ for this Kostrikin's function.

Proof. By the preceding corollary the additive group E is generated modulo A by elements of the form $[a, {}_{p-1}\eta]$, where a is a commutator in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L. By Theorem 1.3.1, the factor-ring $L/(E + pL)$ is nilpotent of (d, p)-bounded class $k(d, p)$. If the weight of a is greater than $k(d, p)$, then $a \in E + pL$ and the element a is also equal modulo A (which contains pL) to a linear combination of elements of the form $[a_1, {}_{p-1}\eta]$ where a_1 is a commutator in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L, and hence $[a, {}_{p-1}\eta]$ is equal to a linear combination of elements of the form $[a_1, {}_{2p-2}\eta]$. The same arguments may be applied to these elements as long as the weight of a_1 is greater than $k(d, p)$, etc. As a result, every element of the form $[a, {}_{p-1}\eta]$ may be represented modulo A as a linear combination of elements of the form (7.2.24) which therefore generate the additive group E modulo A.

The corollary is proved.

Corollary 7.2.23 yields a description of all homogeneous elements of E.

7.2.25 Corollary. *Every homogeneous element of E is equal modulo A to an element of the form $[b, {}_m\eta]$ where $m \geq p - 1$ and b is a homogeneous element of L whose weight is at most $k(d, p)$.*

Proof. By Corollary 7.2.23 an arbitrary homogeneous element $l \in E$ of weight s is equal modulo A to a linear combination of elements of the form (7.2.24):

$$l \equiv \sum_j \alpha_j [b_j, {}_{m_j}\eta] \pmod{A},$$

where $m_j \geq p - 1$ and the weight of b_j is at most $k(d, p)$. Since the element l and the ideals E and A are homogeneous, we may assume that the weight of b_j is $s - m_j$ for each j. If r is the maximal weight of the elements b_j, then clearly $s - m_j \leq r$ for all j, that is, $m_j - s + r \geq 0$. Hence we may write

$$l \equiv \left[\sum_j \alpha_j [b_j, {}_{m_j - s + r}\eta], {}_{s-r}\eta \right] \pmod{A},$$

where $b = \sum_j \alpha_j [b_j, {}_{m_j - s + r}\eta]$ is a homogeneous element of weight $r \leq k(d, p)$.

The corollary is proved.

Now our main goal is a (d, p)-bounded estimate for the order of the centralizer $C_{F/N}(\bar{y})$ where $\bar{y}$ is the image of y in F/N. We shall write $\gamma_k = \gamma_k(F/N), k \in \mathbb{N}$, for brevity.

We define the "trace" of the subgroup $C_{F/N}(\bar{y})$ in the additive group of the Lie ring L to be the full inverse image L_1 of the additive subgroup

$$\bigoplus_{i=1}^{n}((C_{F/N}(\bar{y})\cap\gamma_i)\cdot\gamma_{i+1})/\gamma_{i+1}$$

of the Lie ring $L(F/N)$, where the i-th summand is regarded as a subgroup of the homogeneous component of $L(F/N)$ of weight i. (In fact, as it is easy to see, L_1 is even a subring, but we do not need this fact.)

Although L_1/A is not the associated Lie ring of the group $C_{F/N}(\bar{y})$, their orders coincide.

7.2.26 Lemma. *We have* $|L_1/A| = |C_{F/N}(\bar{y})|$.

Proof. Consider the normal series

$$C_{F/N}(\bar{y}) \supseteq C_{F/N}(\bar{y})\cap\gamma_2 \supseteq C_{F/N}(\bar{y})\cap\gamma_3 \supseteq \ldots \supseteq C_{F/N}(\bar{y})\cap\gamma_n \supseteq 1,$$

of $C_{F/N}(\bar{y})$, whose factors are isomorphic to direct summands of L_1/A:

$$(C_{F/N}(\bar{y})\cap\gamma_i)/(C_{F/N}(\bar{y})\cap\gamma_{i+1}) =$$
$$= (C_{F/N}(\bar{y})\cap\gamma_i)/(C_{F/N}(\bar{y}))\cap\gamma_i\cap\gamma_{i+1}) \cong ((C_{F/N}(\bar{y})\cap\gamma_i)\cdot\gamma_{i+1})/\gamma_{i+1}.$$

Hence the order of $C_{F/N}(\bar{y})$ is equal to

$$\prod_{i=1}^{n}|((C_{F/N}(\bar{y})\cap\gamma_i)\cdot\gamma_{i+1})/\gamma_{i+1}|$$

– as is the order of L_1/A.

The lemma is proved.

The following is the key lemma for bounding the order of L_1/A.

7.2.27 Lemma. *If a is a homogeneous element of L_1 then* $[a,\eta]\in A$.

Proof. Since a is homogeneous of weight i, say, the hypothesis implies that the element $\tilde{a}\in F$ belongs to the inverse image of $(C_{F/N}(\bar{y})\cap\gamma_i)\cdot\gamma_{i+1}$; here $\tilde{a}$ denotes the product of powers of commutators in the generators $y, x_1, x_2, \ldots, x_d$ which corresponds to a as a linear combination of Lie ring commutators of weight i in the generators $\eta, \xi_1, \xi_2, \ldots, \xi_d$ of L. Denoting images in F/N by bars, we may write this as $\bar{\tilde{a}} = c\cdot g$ where $c\in C_{F/N}(\bar{y})\cap\gamma_i$ and $g\in\gamma_{i+1}$. Then we have

$$[\bar{\tilde{a}},\bar{y}] = [c\cdot g,\bar{y}] = [c,\bar{y}]^g\cdot[g,\bar{y}] = [g,\bar{y}]\in\gamma_{i+2}.$$

By the definition of multiplication in the Lie ring $L(F/N)$ the image of $[a, \eta]$ in L/A, which lies in the homogeneous component of $L(F/N)$ of weight $i+1$, is equal to the image of $[\bar{\bar{a}}, \bar{y}]$ in $\gamma_{i+1}/\gamma_{i+2}$ and hence is trivial. Hence $[a, \eta] \in A$.

The lemma is proved.

This lemma immediately implies

7.2.28 Corollary. *For every homogeneous element b, there is at most one value of $m \in \mathbb{N}$ such that the image of $[b, {}_m\eta]$ in L/A is a nontrivial element of L_1/A.*

Now we are ready to bound the order $|L_1/A|$. It is clear that

$$|L_1/A| = |L_1/L_1 \cap (E+A)| \cdot |(L_1 \cap (E+A))/A)|.$$

Since $L_1/L_1 \cap (E+A) \cong (L_1 + E + A)/(E+A)$, the order of the first factor is (d, p)-bounded, since $A \supseteq pL$ by Lemma 7.2.6 and the order of $L/(E+pL)$ is bounded in terms of d and p by Kostrikin's Theorem 1.3.1.

It therefore remains to bound the order of $(L_1 \cap (E+A))/A$. In view of the fact that the additive subgroups L_1, E and A are homogeneous, it is sufficient to bound the number of images of homogeneous elements in this additive factor-group.

By Corollary 7.2.25 all homogeneous elements of $L_1 \cap (E+A)$ have the form $[b, {}_m\eta]$ (mod A), where b is a homogeneous element of L whose weight is not greater than $k(d, p)$. For each such element b, there is, by Corollary 7.2.28, ***at most one*** value of m such that the image of $[b, {}_m\eta]$ is, modulo A, a non-trivial element of L_1. Since $A \supseteq pL$, this implies that the number of non-trivial images of homogeneous elements in $(L_1 \cap (E+A))/A$ does not exceed the number of homogeneous elements of L/pL of weight $\leq k(d, p)$ and hence it is (d, p)-bounded.

Thus we have proved that the order of L_1/A and hence also, by Lemma 7.2.26, the order of $C_{F/N}(\bar{y})$, are (d, p)-bounded.

By Theorem 5.2.1 the finite p-group F/N, which admits the automorphism of order p induced by conjugation by $\bar{y}$ which has a (d, p)-bounded number of fixed points $C_{F/N}(\bar{y})$, possesses a subgroup of (d, p)-bounded index which is nilpotent of class $\leq h(p)$, where $h(p)$ is Higman's function. In particular, this group is soluble of (d, p)-bounded derived length. By the definition of N and by Proposition 7.1.1 the element $\bar{y}$ induces a splitting automorphism of order p of the image of the normal closure $\langle \{x_1, x_2, \ldots, x_d\}^F \rangle$ in F/N. Hence, by Corollary 6.4.2, the nilpotency class of the image of $\langle \{x_1, x_2, \ldots, x_d\}^F \rangle$ is (d, p)-bounded.

The theorem is proved.

§ 7.3 The structure of finite p-groups admitting a partition and a positive solution of the Hughes problem

All results in this section describing the structure of finite p-groups admitting a partition are consequences of the main Theorem 7.2.1. For finite p-groups, it may be reformulated using Proposition 7.1.1 as

7.3.1 Theorem. *Every d-generator finite p-group admitting a partition contains a subgroup of index p which is nilpotent of (d, p)-bounded class not greater than $f(d-1, p)$ where f is the function from the statement of Theorem 7.2.1.*

Proof. If a d-generator finite p-group P admits a partition then by Proposition 7.1.1 it may be represented as a semidirect product $P = P_1 \leftthreetimes \langle\varphi\rangle$ where φ is a splitting automorphism of order p of P_1. The element φ may be included in a minimal system of generators of P in such a way that the remaining $d-1$ elements belong to P_1. (This follows from the Burnside Basis Theorem 2.8.5: it is clear that the image of φ in $P/\Phi(P)$ is non-trivial and together with the image of P_1 generates the whole factor-group.) These $d-1$ elements generate P_1 as a $\langle\varphi\rangle$-group. Since P_1 is a finite p-group, it is nilpotent and we may apply Theorem 7.2.1.

The theorem is proved.

It is easy to see that the proof of Theorem 7.2.1 yields also another result for finite p-groups based on Theorem 5.2.1.

7.3.2 Theorem. *If a d-generator finite p-group admits a splitting automorphism of prime order p then it has a subgroup of (d, p)-bounded index which is nilpotent of class $\leq h(p)$, where h is Higman's function.*

Proof. Let P be a d-generator finite p-group admitting a splitting automorphism φ of order p. Then, as it was shown in the proof of Theorem 7.2.1, the order of $C_P(\varphi)$ is bounded in terms of d and p. (Although, in the proof of Theorem 7.2.1, such a bound was obtained for the universal group F/N, that is, for the order of $C_{F/N}(\bar{y})$, the inequality $|C_P(\varphi)| \leq |C_{F/N}(\bar{y})|$ holds by Theorem 1.6.1.) An application of Theorem 5.2.1 completes the proof.

The theorem is proved.

A corresponding theorem for finite p-groups with a partition is derived from Theorem 7.3.2, in an analogous way using Proposition 7.1.1.

7.3.3 Theorem. *Every d-generator finite p-group admitting a partition contains a subgroup of (d, p)-bounded index which is nilpotent of class $\leq h(p)$, where h is Higman's function.*

We shall now prove that the Hughes conjecture holds for almost all finite p-groups. This is the content of the next two theorems.

7.3.4 Theorem. *For a given prime p, the Hughes conjecture holds for all finite p-groups which contain elements of sufficiently large order $p^{\alpha(p)}$, where α is a function depending on p only.*

7.3.5 Theorem. *For a given prime p and a given number d, the Hughes conjecture is valid for all finite d-generator p-groups of sufficiently large order $\geq p^{\beta(d,p)}$, where β is a function depending on d and p only.*

It makes sense to fix the number of generators in the statement of Theorem 7.3.5, since if there is a finite p-group which is a counterexample to the Hughes conjecture, then there exist d-generator counterexamples for all $d \geq 3$. This follows from the construction of universal counterexamples to the Hughes conjecture which will be given later.

Proof. It is clear that we may prove equivalently that any finite d-generator p-group P which is a counterexample to the Hughes conjecture has p-bounded exponent and (d, p)-bounded order. We shall in fact prove this under the formally weaker hypothesis $|P : H_p(P)| \geq p^2$. (Although this result includes a positive solution to the Restricted Burnside Problem for groups of prime exponent as a special case, we remind the reader that the proof of the main Theorem 7.2.1 is based on Kostrikin's Theorem 1.3.1.)

We choose a subgroup P_1 such that $P > P_1 > H_p(P)$ and $|P : P_1| = p$. By Proposition 7.1.1, the group P_1 admits a splitting automorphism φ of prime order p induced by conjugation by an element from $P \setminus P_1$. We prove that the exponent of $H_p(P)$ is bounded in terms of p. We fix an arbitrary element $a \in P_1 \setminus H_p(P)$. Then for any $x \in H_p(P)$, both a and ax do not lie in $H_p(P)$ and hence have order p. The subgroup generated by the $2p$ elements

$$a, a^{\varphi}, a^{\varphi^2}, \ldots, a^{\varphi^{p-1}}, ax, (ax)^{\varphi}, (ax)^{\varphi^2}, \ldots, (ax)^{\varphi^{p-1}}$$

(or, which is 2-generated as a $\langle\varphi\rangle$-subgroup by the a and ax) is φ-invariant, that is, it also admits a splitting automorphism of prime order p. Therefore, by Theorem 7.2.1 its nilpotency class is p-bounded by $f(2, p)$ where f is the function appearing in the statement of Theorem 7.2.1. Hence the exponent of this subgroup, which is generated by elements of order p, divides the p-bounded number $p^{f(2,p)}$ (see Corollary 2.5.4). But this subgroup contains $x = a^{-1} \cdot ax$. So, the order of any element $x \in H_p(P)$ divides $p^{f(2,p)}$. Since the exponent of $P/H_p(P)$ is p, the exponent of P itself divides the p-bounded number $p^{f(2,p)+1}$.

Theorem 7.3.4 is proved.

We now prove that the order of P is bounded in terms of d and p. As in the proof of Theorem 7.3.1, it is easy to see that P_1 may be generated by $d-1$ elements. We can choose these elements to lie outside $H_p(P)$ and hence having order p (if, in some set of generators of P_1, the elements $b_1, b_2, \ldots, b_k$ belong to $H_p(P)$ then they can be replaced by $ab_1, ab_2, \ldots, ab_k$ where a is an element of the set which does not belong to $H_p(P)$).

By Theorem 7.2.1 the nilpotency class of P_1 is at most $f(d-1, p)$, where f is as above. By Corollary 2.5.3 each factor of the lower central series of P_1 has exponent p and by Corollary 2.5.6 it is generated by a (d, p)-bounded number of elements. Hence the order of P_1, and therefore the order of P, is bounded in terms of d and p.

Theorem 7.3.5 is proved.

Theorem 7.3.5 yields the existence of universal counterexamples to Hughes conjecture which also has the form of a positive solution to the Restricted Burnside Problem for such groups. By the term, a *counterexample to the Hughes conjecture*, we shall always mean a finite p-group which is a counterexample to the Hughes conjecture.

7.3.6 Corollary. *Suppose that p is a prime for which there exists a counterexample to the Hughes conjecture. Then, for every $d \geq 3$, there exists a universal d-generator counterexample to the Hughes conjecture such that all d-generator counterexamples are its homomorphic images.*

Suppose that p is a prime for which there exists a 2-generator counterexample to the Hughes conjecture. Then there exists a universal 2-generator counterexample such that all 2-generator counterexamples are its homomorphic images.

Proof. At first, we describe a general construction for a universal group analogous to the group F/N in the proof of Theorem 7.2.1. Let p be a prime, n a natural number and let F be a free d-generator nilpotent group of class n with free generators $y_1, y_2, x_1, x_2, \ldots, x_{d-2}$. We set

$$N = \langle (y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^p \mid \varkappa \in F' \cdot \langle x_1, x_2, \ldots, x_{d-2} \rangle, \\ \text{where either } k_1 \not\equiv 0 \pmod{p} \text{ or } k_2 \not\equiv 0 \pmod{p} \rangle.$$

It is easy to see that N is a normal subgroup of F: for any $g \in F$

$$((y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^p)^g = ((y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^g)^p = (y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa \cdot [y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa, g])^p$$

and $\varkappa \cdot [y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa, g] \in F' \cdot \langle x_1, x_2, \ldots, x_{d-2} \rangle$.

Since F/N is nilpotent and is generated by elements of order p, for example, the images of $y_1, y_2, y_1x_1, y_1x_2, \ldots, y_1x_{d-2}$, it is a finite p-group (see Corollaries 2.5.4 and 2.5.6).

Put $F_1 = F' \cdot \langle y_1^p, y_2^p, x_1, x_2, \ldots, x_{d-2}\rangle$; then $|F : F_1| = p^2$, the factor-group F/F_1 has exponent p and hence $F_1 \geq F^p \geq N$. Every $g \in F\backslash F_1$ has the form $y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa$ where $\varkappa \in F' \cdot \langle x_1, x_2, \ldots, x_{d-2}\rangle$ and either $k_1 \not\equiv 0$ or $k_2 \not\equiv 0 \pmod p$, therefore $g^p \in N$ by construction. Hence all elements of $(F/N)\backslash(F_1/N)$ have order p, that is, $F_1/N \geq H_p(F/N)$.

Although $|F/N : H_p(F/N)| \geq p^2$, it may very well happen that $H_p(F/N) = 1$ and F/N is not a counterexample to the Hughes conjecture.

Suppose, however, that P is a counterexample to the Hughes conjecture. Since $|P : H_p(P)| \geq p^2$, we can choose a subgroup P_1 of index p^2 containing $H_p(P)$ and two elements b_1, b_2 of order p such that they, together with P_1, generate P. Since $H_p(P) \neq 1$, there is an element $a \in H_p(P)$ of order p^2.

Now if, in the definition of F/N given above, we put n equal to the nilpotency class of P and d arbitrary ≥ 3, then the mapping

$$y_1 \to b_1, \quad y_2 \to b_2, \quad x_1 \to a, \quad x_i \to 1 \text{ for } i \geq 2$$

extends to a homomorphism ϑ of F into P. Its kernel contains N. Indeed, we have

$$\vartheta((y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^p) = (b_1^{k_1} \cdot b_2^{k_2} \cdot \vartheta(\varkappa))^p,$$

and if either $k_1 \not\equiv 0$ or $k_2 \not\equiv 0 \pmod p$ then $b_1^{k_1} \cdot b_2^{k_2}$ does not belong to P_1 while $\vartheta(\varkappa) \in P_1$ since $\vartheta(\langle x_1, x_2, \ldots, x_{d-2}\rangle) \subseteq H_p(P) \subseteq P_1$ and $\vartheta(F') \subseteq P_1$. Hence $b_1^{k_1} \cdot b_2^{k_2} \cdot \vartheta(\varkappa)$ has order p whence

$$\vartheta((b_1^{k_1} \cdot b_2^{k_2} \cdot \varkappa)^p) = 1,$$

that is, $\operatorname{Ker}\vartheta \supseteq N$. So ϑ induces a homomorphism of F/N into P which is denoted by the same letter.

It is clear now that the image of x_1 in F/N, being an inverse image of a which has order p^2, has order $\geq p^2$. So, $H_p(F/N) \neq 1$.

By increasing the value of n in the definition of F/N, we shall also obtain d-generator counterexamples to the Hughes conjecture, the order of F/N being a non-decreasing function of n. But, by Theorem 7.3.5, the order of a d-generator counterexample to the Hughes conjecture is bounded in terms of d and p. Hence, the order of F/N stops increasing at some value of n and we shall assume thereafter that F/N is just the limit group in this sense.

Suppose now that P is any d-generator counterexample to the Hughes conjecture. We prove that P is a homomorphic image of F/N. Since F/N does not change with increasing n, we may assume that the nilpotency class of P is not greater

than n. We choose elements $b_1, b_2 \in P$ and a subgroup $P_1 \leq P$ as above. Since the images of $\langle b_1 \rangle$ and $\langle b_2 \rangle$ are distinct in $P/\Phi(P)$, as $\Phi(P) \subseteq P_1$, the elements b_1, b_2 may be supplemented by $d-2$ elements $a_1, a_2, \ldots, a_{d-2}$ of P_1 to give a minimal set of generators of P (by the Burnside Basis Theorem 2.8.5, see also the proof of Theorem 7.3.1).

The mapping

$$y_1 \to b_1, \quad y_2 \to b_2, \quad x_i \to a_i, \quad i = 1, 2, \ldots, d-2$$

obviously extends to a homomorphism ϑ of F onto P. We must now prove that its kernel contains N. We have $\vartheta((y_1^{k_1} y_2^{k_2} \varkappa)^p) = (b_1^{k_1} b_2^{k_2} \vartheta(\varkappa)))^p$ and if either $k_1 \not\equiv 0$ or $k_2 \not\equiv 0 \pmod{p}$ then $b_1^{k_1} b_2^{k_2}$ does not belong to P_1 while $\vartheta(\varkappa) \in P_1$ as $\vartheta(\langle x_1, x_2, \ldots, x_{d-2} \rangle) = \langle a_1, a_2, \ldots, a_{d-2} \rangle \subseteq P_1$ and $\vartheta(F') \subseteq P_1$. Hence, the order of $b_1^{k_1} b_2^{k_2} \vartheta(\varkappa)$ is p, that is, $\vartheta((b_1^{k_1} b_2^{k_2} \varkappa)^p) = 1$. So, $\operatorname{Ker} \vartheta \supseteq N$ and ϑ induces a homomorphism of F/N onto P, as required.

Suppose now that there exists a 2-generator counterexample T to the Hughes conjecture. Clearly, we have $H_p(T) \leq \Phi(T)$.

Put $d = 2$ in the definition of F/N, that is, $F = \langle y_1, y_2 \rangle$ and

$$N = \langle (y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^p | \ \varkappa \in F', \text{ and either } k_1 \not\equiv 0, \text{ or } k_2 \not\equiv 0 \pmod{p} \rangle.$$

Let $F_1 = F' \cdot \langle y_1^p, y_2^p \rangle$; it is easy to see that $F_1 \geq N$ and $|F : F_1| = p^2$. Also, we have $F_1/N \geq H_p(F/N)$ and $|F/N : H_p(F/N)| \geq p^2$.

We take n now to be the nilpotency class of T. If b_1, b_2 are generators of T then the mapping $y_1 \to b_1$, $y_2 \to b_2$ extends to a homomorphism ϑ of F onto T whose kernel contains N. Indeed, we have

$$\vartheta((y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^p) = (b_1^{k_1} \cdot b_2^{k_2} \cdot \vartheta(\varkappa))^p$$

and, if either $k_1 \not\equiv 0$ or $k_2 \not\equiv 0 \pmod{p}$, then $b_1^{k_1} b_2^{k_2}$ does not belong to $\Phi(T)$ while $\vartheta(\varkappa) \in T' \leq \Phi(T)$. Hence the order of $b_1^{k_1} b_2^{k_2} \vartheta(\varkappa)$ is p, so that $\vartheta((b_1^{k_1} \cdot b_2^{k_2} \cdot \varkappa)^p) = 1$. So $\operatorname{Ker} \vartheta \supseteq N$ and ϑ induces a homomorphism of F/N onto T which is denoted by the same letter. Since the exponent of T is greater than p, the exponent of F/N is also greater than p, so that F/N is a 2-generator counterexample to the Hughes conjecture.

Increasing the value of n in the definition of F/N, we shall also obtain 2-generator counterexamples to Hughes conjecture, the order of F/N being a non-decreasing function of n. But, by Theorem 7.3.5, the order of a 2-generator counterexample to the Hughes conjecture is bounded in terms of 2 and p. Hence, the order of F/N stabilizes at some value of n, the corresponding limit group F/N in this sense is the desired universal 2-generator counterexample. The fact that any

2-generator counterexample T to the Hughes conjecture is a homomorphic image of F/N has, in fact, already been established above.

The corollary is proved.

Finally, we give an alternative proof of the Hughes conjecture for almost all finite p-groups. Theorems 7.3.4 and 7.3.5 may be derived from the following theorem which is an exact analogue of the Magnus-Sanov Theorem 3.3.2 for counterexamples to the Hughes conjecture. Actually, the positive solution to the Hughes conjecture follows from this theorem directly, via Kostrikin's Theorem 1.3.1, but its proof uses Lemma 7.2.6 and Proposition 7.2.7 from the proof of Theorem 7.2.1.

7.3.7 Theorem. *The associated Lie ring of any counterexample to the Hughes conjecture has characteristic p and satisfies the $(p-1)$-Engel condition.*

Proof. Standard arguments show that it is sufficient to prove the theorem for a universal group F/N, where F is a free nilpotent group of sufficiently large class n with free generators $y_1, y_2, x_1, x_2, \ldots$ and

$$N = \langle (y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^p \mid \varkappa \in F' \cdot \langle x_1, x_2, \ldots \rangle, \\ \text{where either } k_1 \not\equiv 0 \pmod{p} \text{ or } k_2 \not\equiv 0 \pmod{p} \rangle.$$

Let L/A be the associated Lie ring of F/N, where L is a free nilpotent $\mathbb{Z}$-Lie ring of class n with free generators $\eta_1, \eta_2, \xi_1, \xi_2, \ldots$ and A its ideal spanned by the $(N \cap \gamma_i(F))\gamma_{i+1}(F)/\gamma_{i+1}(F)$ regarded as additive subgroups of homogeneous components of $L \cong L(F)$.

The subgroup N obviously contains both

$$N_1 = \langle (y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^p \mid \varkappa \in F' \cdot \langle x_1, x_2, \ldots \rangle, \quad \text{where } k_1 \not\equiv 0 \pmod{p} \rangle$$

and

$$N_2 = \langle (y_1^{k_1} \cdot y_2^{k_2} \cdot \varkappa)^p \mid \varkappa \in F' \cdot \langle x_1, x_2, \ldots \rangle, \quad \text{where } k_2 \not\equiv 0 \pmod{p} \rangle.$$

It is easy to see that both factor-groups F/N_1 and F/N_2 are universal groups in the sense of the proof of Theorem 7.2.1 (where the images of y_1 and y_2 take the role of the splitting automorphism of prime order p, respectively). Hence, by Lemma 7.2.6, A contains pL. Furthermore, by Proposition 7.2.7 applied to F/N_1, A contains all Kostrikin elements in commutators in the generators $\eta_1, \eta_2, \xi_1, \xi_2, \ldots$ which are not of the form $\ll a, \underbrace{\eta_1, \ldots, \eta_1}_{p-1} \gg$. But the latter ones are not of the form $\ll a, \underbrace{\eta_2, \ldots \eta_2}_{p-1} \gg$ and therefore belong to A, by Proposition 7.2.7 applied to F/N_2.

As a result, we get that all Kostrikin elements in commutators in the generators $\eta_1, \eta_2, \xi_1, \xi_2, \ldots$ are contained in A, which means that the Lie ring L/A satisfies the $(p-1)$-Engel identity (see Lemma 7.2.4).

The theorem is proved.

The positive solution to the Hughes problem now easily follows.

7.3.8 Corollary. a) *The nilpotency class of a d-generator counterexample to the Hughes conjecture does not exceed $k(d, p)$, where $k(d, p)$ is Kostrikin's function depending on d and p only, which bounds the nilpotency class of a d-generator $(p-1)$-Engel Lie algebra of characteristic p.*

b) *The order of a d-generator counterexample to the Hughes conjecture is bounded on terms of d and p.*

c) *The exponent of a counterexample to the Hughes conjecture is bounded in terms of p only.*

Proof. Let P be a d-generator counterexample to the Hughes conjecture. The required bounds for the nilpotency class and the order of P follow from Kostrikin's Theorem 1.3.1 applied to the associated Lie ring of P whose nilpotency class and order coincide with those of P. It follows that the exponent of any 3-generator counterexample is bounded in terms of p only. Now, if b_1 and b_2 are chosen as in the proof of Corollary 7.3.6 (so that they are "independent" elements outside $H_p(P)$), then for any $a \in H_p(P)$ the subgroup $\langle a, b_1, b_2 \rangle$ is also a counterexample to the Hughes conjecture whenever $a^p \neq 1$. Hence the order of any $a \in H_p(P)$ is bounded in terms of p only.

§ 7.4 Bounding the index of the Hughes subgroup

We have already mentioned that the existence of finite p-groups which are counterexamples to the Hughes conjecture is connected with new identities in the Lie rings of free groups of prime exponent. All multilinear identities of these Lie rings were described by Vaughan-Lee [144] and it is conjectured that all of their relations are consequences of these multilinear identities. Wall has proved that the index of the Hughes subgroup in the counterexamples to the Hughes conjecture may be the larger, the more of the Vaughan-Lee's multilinear identities are not consequences of those of smaller degree (see comments in § 7.5). However, little is known about which of these identities are really new in this sense. We prove in this section that if $|P : H_p(P)| = p^k$ in a finite p-group P then the associated Lie ring of P satisfies all multilinear identities of degrees $\leq (k-1)(p-1)+1$ of the associated Lie ring of a free group of exponent p. This result is contained in the proof of

the following theorem; another consequence of this proof is a lower bound for the nilpotency class of P in terms of the structure of $P/H_p(P)$, provided $H_p(P) \neq 1$.

7.4.1 Theorem. *If the associated Lie ring of a free countably-generated group of prime exponent p is a relatively free Lie ring all of whose identities follow from its multilinear identities of degree at most $1+k_0(p-1)$ then the index of the non-trivial Hughes subgroup $H_p(G) \neq 1$ in an arbitrary finite group G is at most p^{k_0}.*

Before proceeding with the proof of the theorem, we quote some properties of the multilinear identities of the associated Lie rings of groups of prime exponent p from the works of Vaughan-Lee [144] and Wall [152] and fix some notation. It is convenient to consider the factor-group F/F^p of a free countably-generated (nilpotent of class c, say) group F with free generators $x_1, x_2, \ldots, y_1, y_2, \ldots$. Then its associated Lie ring $L(F/F^p)$ may be represented as a factor-ring of a free Lie ring L (over $\mathbb{Z}$) with free generators $\xi_1, \xi_2, \ldots, \eta_1, \eta_2, \ldots$ by an ideal I, the images of the generators $\xi_1, \xi_2, \ldots, \eta_1, \eta_2, \ldots$ being identified with the images of $x_1, x_2, \ldots, y_1, y_2, \ldots$, respectively, in $(F/F^p)/(F/F^p)' \cong F/(F^p \cdot F')$ as the homogeneous component of $L(F/F^p)$ of weight 1. According to [144], for each $k \in \mathbb{N}$, there is a multilinear identity V_k of degree k which holds in $L(F/F^p)$ and all multilinear identities of $L(F/F^p)$ are exhausted by the consequences of these V_k. The hypothesis of the theorem means that the ideal I is a verbal ideal generated by the Lie ring words $V_k(\xi_1, \xi_2, \ldots, \xi_k)$ for $k \leq 1 + k_0(p-1)$. Since these words are multilinear, the additive group I is generated by the values

$$V_k(\mu_1, \mu_2, \ldots, \mu_k), \quad k \leq 1 + k_0(p-1)$$

where $\mu_i \in L$ and, moreover, we can take the μ_i to be commutators in the generators $\xi_1, \xi_2, \ldots, \eta_1, \eta_2, \ldots$ of L.

(Note that the same description is valid also for the associated Lie ring of a free group of exponent p with a finite number of generators. An appropriate finitely generated free group $\bar{F}$ embeds in F by including its free generators in a set of free generators of F and the corresponding free Lie ring $\bar{L}$ embeds in L in an analogous way. If a homogeneous element μ belongs to $\bar{I}$ where $L(\bar{F}/\bar{F}^p) = \bar{L}/\bar{I}$, then it also belongs to I since both $\bar{F}^p$ and $\gamma_i(\bar{F})$ are contained in F^p and $\gamma_i(F)$, respectively. Hence μ is a linear combination of the values of the V_k at elements of L. It remains to show that it is sufficient to take only the values of the V_k at elements of $\bar{L}$. This is obvious after applying to the resultant equality the homomorphism of L which is identical on $\bar{L}$ and which takes the generators of L which are outside $\bar{L}$ to 0.)

We do not need the explicit form of the identities V_k, but only some of their properties. We denote the ideal whose additive group is generated by the values $V_k(\mu_1, \mu_2, \ldots, \mu_k)$ of the V_k for $k \leq s$, by I_s.

7.4.2 Lemma [144, Lemma 5]. *For every $k \in \mathbb{N}$ the Lie ring polynomial V_k is a symmetric function modulo I_{k-1}, that is,*

$$V_k(\mu_{\pi(1)}, \mu_{\pi(2)}, \ldots, \mu_{\pi(k)}) \equiv V_k(\mu_1, \mu_2, \ldots, \mu_k) \pmod{I_{k-1}}$$

for any $\mu_i \in L$ and any $\pi \in \mathbb{S}_k$.

7.4.3 Lemma [144, Lemma 8]. *For every $k \in \mathbb{N}$ the value of the polynomial $V_k(\mu_1, \mu_2, \ldots, \mu_k)$ belongs to I_{k-1} if there are at least p equal elements among the arguments $\mu_1, \mu_2, \ldots, \mu_k$.*

According to this lemma we can reformulate the hypothesis of the theorem as follows: the additive group I is generated by the values

$$V_k(\mu_1, \mu_2, \ldots, \mu_k), \quad k \leq 1 + k_0(p-1) \tag{7.4.4}$$

where the μ_i are commutators in the generators $\xi_1, \xi_2, \ldots, \eta_1, \eta_2, \ldots$ of L and neither of the μ_i occurs more than $p-1$ times.

We continue to use the following convention. Let $u(\xi_1, \xi_2, \ldots, \eta_1, \eta_2, \ldots)$ be a homogeneous Lie ring polynomial of degree s; it is a linear combination of commutators of weight s in generators $\xi_1, \xi_2, \ldots, \eta_1, \eta_2, \ldots$. We shall denote by $\tilde{u}(x_1, x_2, \ldots, y_1, y_2, \ldots)$ a product of powers of group commutators with the same bracket structure in generators $x_1, x_2, \ldots, y_1, y_2, \ldots$ and with exponents equal to the coefficients in the corresponding Lie ring commutators (the order in which this product is taken is arbitrary since it will occur only in congruences modulo $\gamma_{s+1}(F)$). So $u(\xi_1, \xi_2, \ldots, \eta_1, \eta_2, \ldots)$ is the image of $\tilde{u}(x_1, x_2, \ldots, y_1, y_2, \ldots)\gamma_{s+1}(F)$ under the canonical isomorphism of $\gamma_s(F)/\gamma_{s+1}(F)$ onto the additive group of the homogeneous component of L of weight s induced by the mapping $x_i \to \xi_i$, $y_i \to \eta_i$, $i = 1, 2, \ldots$.

Proof of Theorem 7.4.1. To prove the theorem it is sufficient to show that for any finite group P the hypothesis $|P : H_p(P)| \geq p^{k_0+1}$ implies that $H_p(P) = 1$, that is, that P is a group of exponent p. Since the Hughes conjecture is valid for finite groups which are not p-groups, we may assume that P is a finite p-group.

Here it is also convenient to deal with some universal group instead of P. However, unlike the construction of the preceding section, we are not able here to settle for one universal group, but rather we have to construct a universal group related to P.

We choose a minimal system of generators of P in such a way that its elements $a_1, a_2, \ldots, a_m$ do not lie in $H_p(P)$ while the remainder (possibly, empty) is contained in $H_p(P)$. Let c be the nilpotency class of P. Let $\bar{F}$ be a free nilpotent group of class c with free generators $x_1, x_2, \ldots, x_m, y_1$. For any element $b \in P$

the mapping

$$x_i \to a_i, \quad i = 1, 2, \ldots, m, \quad y_1 \to b$$

extends to a homomorphism ϑ_b of F into P. We denote by ϑ the restriction of ϑ_b to the subgroup $\langle x_1, x_2, \ldots, x_m \rangle$ which clearly does not depend on the choice of b. We define the subgroup $\bar{N}$ of $\bar{F}$ as the normal closure in $\bar{F}$ of all elements of the form

$$(xy)^p \tag{7.4.5}$$

where $x \in \langle x_1, x_2, \ldots, x_m \rangle$ is such that $\vartheta(x) \notin H_p(P)$ and y is an arbitrary element of the normal closure $\langle y_1^{\bar{F}} \rangle$ of y_1 in $\bar{F}$. The factor-group $\bar{F}/\bar{N}$ is the required universal group. This means that the following lemma holds.

7.4.6 Lemma. *For any $b \in P$, the homomorphism ϑ_b induces a homomorphism of $\bar{F}/\bar{N}$ into P. If P is not a group of exponent p, then $\bar{F}/\bar{N}$ is also not a group of exponent p.*

Proof. The second statement follows from the first one: it is sufficient to take b to be an element of order p^2 from $H_p(P)$. Then the image of y_1 in $\bar{F}/\bar{N}$ has order divisible by p^2 since it is an inverse image of b.

To prove the first statement we consider the image of any element (7.4.5) under ϑ_b. We have

$$\vartheta_b(xy) = \vartheta_b(x)\vartheta_b(y) = \vartheta(x)\vartheta_b(y). \tag{7.4.7}$$

By the definition of $\bar{N}$ we also have $\vartheta(x) \notin H_p(P)$ and $\vartheta_b(y) \in \langle b^p \rangle \leq H_p(P)$ as $y \in \langle y_1^{\bar{F}} \rangle$ and $b \in H_p(P)$. Hence the element (7.4.7) lies outside $H_p(P)$ and therefore its p-th power, which is equal to the image of the element (7.4.5) under ϑ_b, is equal to 1. This means that $\bar{N} \leq \operatorname{Ker} \vartheta_b$ and the lemma is proved.

So, in order to prove the theorem, it is sufficient to show that $\bar{F}/\bar{N}$ is a group of exponent p, that is, $\bar{N} \geq \bar{F}^p$. Let $\bar{L}$ be a free nilpotent Lie ring of class c over $\mathbb{Z}$ with free generators $\xi_1, \xi_2, \ldots, \xi_m, \eta_1$. Let $\bar{L}/\bar{I}$ be the associated Lie ring of $\bar{F}/\bar{F}^p$ and let $\bar{L}/\bar{J}$ be the associated Lie ring of $\bar{F}/\bar{N}$, in both cases we identify the images of the generators $\xi_1, \xi_2, \ldots, \xi_m, \eta_1$ with the images of the $x_1, x_2, \ldots, x_m, y_1$, respectively, as described at the beginning of the section. By construction, we have $\bar{N} \leq \bar{F}^p$ so that $\bar{J} \leq \bar{I}$. The group $\bar{F}/\bar{F}^p$ is finite and its order is equal to that of $\bar{L}/\bar{I}$. Therefore to prove that $\bar{N} \geq \bar{F}^p$ (actually, that $\bar{N} = \bar{F}^p$) it is sufficient to show that $\bar{J} \geq \bar{I}$. By the description of $\bar{I}$ it is sufficient to prove that $\bar{J}$ contains all elements of the form (7.4.4).

For technical reasons we have to increase the number of generators. Changing notation slightly, we consider a free nilpotent group F of class c with free generators $x_1, x_2, \ldots, x_m, y_1, y_2, \ldots$, the corresponding free generators of a free Lie ring

L being $\xi_1, \xi_2, \ldots, \xi_m, \eta_1, \eta_2, \ldots$. We consider $\bar{F}$ as a subgroup of F and $\bar{L}$ as a Lie subring of L. We define the subgroup N of F as the normal closure in F of all elements of the form $(xy)^p$ where $x \in \langle x_1, x_2, \ldots, x_m \rangle$ is such that $\vartheta(x) \notin H_p(P)$ and y is an arbitrary element of the normal closure $\langle \{y_1, y_2, \ldots\}^F \rangle$.

We choose commutators $a_{m+1}, a_{m+2}, \ldots, a_d$ ($d \geq m$, that is, possibly an empty set) of increasing weights in the generators $a_1, a_2, \ldots, a_m$ in such a way that, taken together, the images of the $a_1, a_2, \ldots, a_m, a_{m+1}, a_{m+2}, \ldots, a_d$ form a Mal'cev basis for $P/H_p(P)$ (that is, their images generate different factors of some chief series of $P/H_p(P)$). This is possible because, by the choice of the $a_1, a_2, \ldots, a_m$, their images generate $P/H_p(P)$ – see § 2.7. Note that $d \geq 1 + k_0$ by hypothesis. We denote by $x_{m+1}, x_{m+2}, \ldots, x_d$, respectively, the same commutators as $a_{m+1}, a_{m+2}, \ldots, a_d$, but in $x_1, x_2, \ldots, x_m$. Since $P/H_p(P)$ has exponent p,

$$a_1^{i_1} \cdot a_2^{i_2} \cdot \ldots \cdot a_d^{i_d} \in H_p(P)$$

only if each of the $i_1, i_2, \ldots, i_d$ is a multiple of p. In other words,

$$\vartheta(x_1^{i_1} \cdot x_2^{i_2} \cdot \ldots \cdot x_d^{i_d}) \notin H_p(P) \tag{7.4.8}$$

if at least one of the $i_1, i_2, \ldots, i_d$ is not divisible by p.

We denote by $\deg x_i$ the weight of the commutator x_i, $i = 1, 2, \ldots, d$, and we put $w = (p-1) \sum_{i=1}^{d} \deg x_i$.

7.4.9 Lemma. *The factor-group $F/(N \cdot \gamma_{w+1}(F))$ has exponent p.*

Proof. Every $f \in F$ may be written in the form $f = f_x \cdot f_y$ where $f_x \in \langle x_1, x_2, \ldots, x_m \rangle$ and $f_y \in \langle \{y_1, y_2, \ldots\}^F \rangle$. If $\vartheta(f_x) \notin H_p(P)$ then $(f_x \cdot f_y)^p \in N$ by definition.

Now suppose that $\vartheta(f_x) \in H_p(P)$. We shall prove that in this case $(f_x \cdot f_y)^p \in N \cdot \gamma_{w+1}(F)$. By Corollary 1.10.8 of Higman's Lemma we have

$$(x_1^{p-1} \cdot x_2^{p-1} \cdot \ldots \cdot x_d^{p-1} \cdot f)^p = h \cdot \prod_{0 \leq a_i \leq p-1} (x_1^{a_1} \cdot x_2^{a_2} \cdot \ldots \cdot x_d^{a_d} \cdot f^{\varepsilon})^{\pm p}, \tag{7.4.10}$$

where $\varepsilon = 0$ or 1, the product involves the factor $f^{\pm p}$ and h is a product of powers of commutators each of which contains at least $p-1$ occurrences of every element $x_1, x_2, \ldots, x_d$ and at least one occurrence of f. In particular, $h \in \gamma_{w+1}(F)$.

Every element in the brackets in (7.4.10) other than $f^{\pm p}$ has the form $(x_1^{a_1} \cdot x_2^{a_2} \cdot \ldots \cdot x_d^{a_d} \cdot f_x^{\varepsilon} \cdot f_y^{\varepsilon})$ where $0 < a_i \leq p-1$ for at least one i. In the situation under

consideration, $\vartheta(f_x^\ell) \in H_p(P)$ and therefore by (7.4.8) we have

$$\vartheta(x_1^{a_1} \cdot x_2^{a_2} \cdot \ldots \cdot x_d^{a_d} \cdot f_x^\ell) \notin H_p(P).$$

Hence the p-th power of any element in brackets other than $f^{\pm p}$ belongs to N by definition. It now follows from (7.4.10) that $f^p \in N \cdot \gamma_{w+1}(F)$.

The lemma is proved.

If L/J is the associated Lie ring of F/N then it follows from Lemma 7.4.9, according to the description of the associated Lie ring of F/F^p, that for all k

$$V_k(\mu_1, \mu_2, \ldots, \mu_k) \in J \quad \text{if} \quad \sum_{i=1}^{k} \deg \mu_i \leq w \tag{7.4.11}$$

where μ_i are homogeneous elements of L.

The desired result that $\bar{J} \geq \bar{I}$ will be obtained from (7.4.11) by applying homomorphisms of F onto $\bar{F}$. In order to construct these homomorphisms we have to "linearize" using Higman's Lemma. That is why we had to increase the number of generators.

Let $V_k(\mu_1, \mu_2, \ldots, \mu_k)$ be an arbitrary element of the form (7.4.4). We shall distinguish the variables μ_i as commutators in the ξ_i and the η_j. By Lemma 7.4.2 without loss of generality, the first r of them $\mu_1, \mu_2, \ldots, \mu_r$, $k \geq r \geq 0$, may be assumed to be such that $\tilde{\mu}_i = x_{j(i)}$ for all $i \leq r$ where $j(i) \in \{1, 2, \ldots, d\}$. (See the definition of the operator $\tilde{}$ at the beginning of the section.) We change notation for the remaining $k - r$ variables to $\nu_i = \mu_{r+i}$, $1 \leq i \leq k - r$. Each of the ν_i by definition either involves at least one η_j or is a commutator in the ξ_s such that $\vartheta(\tilde{\nu}_i) \in H_p(P)$. Recall that there may be at most $p - 1$ equal elements among the variables μ_i, ν_j. We regard this notation as now fixed.

7.4.12 Lemma. *The following inclusion holds:*

$$V_k(\mu_1, \ldots, \mu_r, \eta_1, \ldots, \eta_{k-r}) \in J. \tag{7.4.13}$$

Proof. According to (7.4.11), it is sufficient to verify that

$$\sum_{i=1}^{r} \deg \mu_i + k - r \leq w = (p-1) \sum_{i=1}^{d} \deg x_i.$$

This inequality follows from the hypothesis that $k \leq 1 + k_0(p-1)$, $d \geq 1 + k_0$ and from the fact that among the μ_i (whose degrees are equal to $\deg \tilde{\mu}_i = \deg x_{j(i)}$) there may be at most $p - 1$ repetitions of any element. Indeed, the maximum of

the sum $\sum_{i=1}^{r} \deg \mu_i + k - r$ is achieved when among the μ_i there are as many elements of greater weights as possible. We now find this maximum. On dividing we get: $k = q(p-1) + t$ where $0 \leq t \leq p-1$. Suppose that $q \leq d-1$. Then, since the order on the commutators x_i agrees with the increase of their weight, the maximum of $\sum_{i=1}^{r} \deg \mu_i + k - r$ is achieved when $k = r$ and is

$$(p-1) \sum_{i=d-q+1}^{d} \deg x_i + t \cdot \deg x_{d-q}.$$

This number clearly does not exceed $(p-1) \sum_{i=1}^{d} \deg x_i = w$.

So, it remains to show that, indeed, $q \leq d-1$. We have

$$q = \left[\frac{k}{p-1}\right] \leq \left[\frac{k_0(p-1)+1}{p-1}\right] = k_0 \leq d-1$$

since $k \leq 1 + k_0(p-1)$ and $d \geq 1 + k_0$.

The lemma is proved.

Set $n = \sum_{i=1}^{r} \deg \mu_i + k - r$. Then, by definition of the associated Lie ring, (7.4.13) is equivalent to

$$\tilde{V}_k(\tilde{\mu}_1, \ldots, \tilde{\mu}_r, y_1, \ldots, y_{k-r}) \equiv 1 \pmod{N \cdot \gamma_{n+1}(F)}.$$

This may be rewritten as a congruence modulo N

$$\tilde{V}_k(\tilde{\mu}_1, \ldots, \tilde{\mu}_r, y_1, \ldots, y_{k-r}) \equiv c_1^{k_1} \cdot c_2^{k_2} \cdot \ldots \cdot c_s^{k_s} \pmod{N} \tag{7.4.14}$$

where the c_i are commutators in the generators x_i and y_j which have weight greater than $\sum_{i=1}^{r} \deg \mu_i + k - r$. The following lemma is a variation of Higman's Lemma, it is analogous to Lemma 7.2.20.

7.4.15 Lemma. *In (7.4.14) all commutators c_i may be assumed to depend on all of $y_1, \ldots, y_{k-r}$.*

Proof. Again, we cannot simply refer to Corollary 1.10.6 and Lemma 1.10.1, since N is not a verbal subgroup. Instead, we can apply arguments similar to the proofs of these results, since $\gamma_{n+1}(F)$ and N are invariant under the homomorphisms ϑ_t

which extend mappings

$$y_t \to 1, \quad y_j \to y_j \text{ for } j \neq t, \quad x_i \to x_i \text{ for all } i.$$

If, for example, in (7.4.14) there are commutators among c_i not depending on y_1 then all their powers may be collected at the beginning of the right-hand side. Then applying ϑ_1 we see that their product is congruent to 1 modulo N since the image of the left-hand side is 1, as well as the images of all commutators depending on y_1. The same thing may now be done relative to y_2, etc.

The lemma is proved.

We are now ready to prove that $\bar{J}$ contains every element of the form (7.4.4) and to do this we use the detailed notation $V_k(\mu_1, \ldots, \mu_r, \nu_1, \ldots, \nu_{k-r})$ where μ_i and ν_j are commutators in the generators $\xi_1, \xi_2, \ldots, \xi_m, \eta_1$ (see above). The mapping

$$y_j \to \tilde{\nu}_j \text{ for } j = 1, 2, \ldots, k-r, \quad y_j \to 1 \text{ for } j > k-r.$$
$$x_i \to x_i \text{ for all } i$$

extends to a homomorphism ρ of F into $\bar{F}$. It is easy to see that the image of N under ρ is contained in $\bar{N}$. So ρ induces a homomorphism of F/N into $\bar{F}/\bar{N}$ which is denoted by the same letter.

We apply ρ to the congruence (7.4.14) which satisfies Lemma 7.4.15, to obtain

$$\tilde{V}_k(\tilde{\mu}_1, \ldots, \tilde{\mu}_r, \tilde{\nu}_1, \ldots, \tilde{\nu}_{k-r}) \equiv \rho(c_1)^{k_1} \cdot \rho(c_2)^{k_2} \cdot \ldots \cdot \rho(c_s)^{k_s} \pmod{\bar{N}}. \quad (7.4.16)$$

We show that the weight of every $\rho(c_t)$, as a commutator in the x_i and y_1, is greater than $\sum_{i=1}^{r} \deg \mu_i + \sum_{i=1}^{s-r} \deg \nu_i$. Indeed, the weight of c_t is greater than $\sum_{i=1}^{r} \deg \mu_i + k - r$ and replacing an occurrence of each of the y_j $(j = 1, 2, \ldots, k-r)$ by the $\tilde{\nu}_j$ contributes to the weight of $\rho(c_t)$ as much as does replacing 1 by $\deg \nu_j$. So, all of the $\rho(c_t)$ belong to $\gamma_{u+1}(\bar{F})$, where $u = \sum_{i=1}^{r} \deg \mu_i + \sum_{i=1}^{s-r} \deg \nu_i$ is the weight of those commutators, the product of whose powers is the left-hand side of (7.4.16). That is

$$\tilde{V}_k(\tilde{\mu}_1, \ldots, \tilde{\mu}_r, \tilde{\nu}_1, \ldots, \tilde{\nu}_{k-r}) \equiv 1 \pmod{\bar{N} \cdot \gamma_{u+1}(\bar{F})}.$$

By the definition of the associated Lie ring $\bar{L}/\bar{J}$ of $\bar{F}/\bar{N}$ this means that

$$V_k(\mu_1, \ldots, \mu_r, \nu_1, \ldots, \nu_{k-r}) \in \bar{J}.$$

We have therefore proved that $\bar{J} \supseteq \bar{I}$, as required.

The theorem is proved.

The hypothesis of Theorem 7.4.1 may not be satisfied. Its proof, nevertheless, yields information on the structure of finite p-groups with large indices of the Hughes subgroup. The following result generalizes Theorem 7.3.7.

7.4.17 Theorem. *If the index of the Hughes subgroup $H_p(G)$ in a finite p-group G is p^k then the associated Lie ring of G satisfies all multilinear identities of degrees at most $(k-1)(p-1)+1$ which hold in the associated Lie ring of the free countably generated group of exponent p.*

Proof. This has actually already been proved in the course of proving Theorem 7.4.1 for the universal groups F/N. But it may be easily shown that the associated Lie ring of a factor-group of a group is a homomorphic image of the associated Lie ring of the group.

The following theorem may find an application in bounding the nilpotency class of the factor-group by the *non-trivial* Hughes subgroup in terms of p only (see Comments in § 7.5).

7.4.18 Theorem. *Suppose that the Hughes subgroup $H_p(G)$ in a finite p-group G is non-trivial and that the minimal number of generators of $G/H_p(G)$ is m and the nilpotency class of $G/H_p(G)$ is k. Then the nilpotency class of G is at least $(p-1)(m+k(k+1)/2-1)+1$.*

Proof. We recall that the construction of the universal group F/N for the given group G in the proof of Theorem 7.4.1 starts with a free nilpotent group F whose nilpotency class c is exactly the nilpotency class of G. By Lemma 7.4.9, in the notation of the proof of Theorem 7.4.1, the factor-group $F/N\gamma_{w+1}(F)$ has exponent p, where $w = (p-1)\sum_{i=1}^{d} \deg x_i$ and the x_i are commutators in the generators $x_1, x_2, \ldots, x_m$. Since the nilpotency class of $G/H_p(G)$ is k, there is at least one such commutator of each of the weights $1, 2, \ldots, k$. Hence $w \geq (p-1)(m-1+1+2+\ldots+k) = (p-1)(m+k(k+1)/2-1)$. Since F/N is not a group of exponent p by the hypothesis, we have $\gamma_{w+1}(F) \neq 1$ which means that the nilpotency class of G is at least $(p-1)(m+k(k+1)/2-1)+1$, as required.

§ 7.5 Comments

Infinite groups admitting a partition. The definition of a group admitting a partition does not presuppose its finiteness. However, although there are some papers (for example, of P.G. Kontorovich [74]) devoted to arbitrary groups with a partition, substantial results have only been obtained for particular classes.

Among infinite groups with a partition there are the groups of prime exponent p. For them, on the one hand, Kostrikin [76] gave an affirmative solution to the Restricted Burnside Problem: the nilpotency class (or, equivalently, the derived length or the order) of an m-generated finite (or, equivalently, soluble or finite) group of exponent p is bounded in terms of d and p. On the other hand, the Adian-Novikov Theorem [1, 116] states that for sufficiently large p, free m-generator Burnside groups $B(m, p)$ are infinite (and nonsoluble). Further important results on the properties of $B(m, p)$ were also obtained in subsequent works of Adian and others.

Among infinite groups of prime exponent there are counterexamples, constructed by Ol'shanskii [117], to the problem of O.Yu. Shmidt, that is, infinite groups all of whose proper subgroups have order p. There also exist variations of these groups in which all elements outside the commutator subgroup have order p while the commutator subgroup is neither periodic nor soluble. Such groups clearly also admit a partition.

Another well-known class of groups with a partition is the class of so-called Frobenius groups, that is, semidirect products of the form $G = N \leftthreetimes A$ where $A \cap A^n = 1$ for all $n \in N\backslash\{1\}$ and $G = N \cup \bigcup_{n \in N} A^n$. Up to now, all that is known about infinite Frobenius groups, is what can be obtained by relatively easy deduction from the theory of finite Frobenius groups: if G is locally finite then N is nilpotent of class $\leq h(p)$ where p is the least prime divisor of elements of A; if N is locally soluble and A has elements of finite order then N is nilpotent, etc. (see, for example, [24, 55]).

Abelian groups admitting a partition are interesting in connection with the geometrical structures which they define (some generalizations of these geometries are defined also by non-abelian finite p-groups admitting a partition).

In view of the above negative results of Adian-Novikov and Ol'shanskii, there is no real hope for any general theory of infinite groups admitting a partition.

Finite groups admitting a partition. Much greater progress has been achieved in the study of finite groups admitting a partition. The works of Hughes and Thompson [46], Baer [4, 5] and Kegel [52–54] give a classification of those finite groups admitting a partition which have a proper Fitting subgroup (that is, the groups are non-nilpotent and contain a non-trivial nilpotent normal subgroup). Suzuki, using his classification of finite simple groups with nilpotent centralizers [138] and his discovery of a new series of finite simple groups [137], completed the classification of non-soluble finite groups with a partition in [139].

Finite Frobenius groups may be characterized as semidirect products of the form $N \leftthreetimes A$, where A is a group of automorphisms of N such that each of its elements is a regular automorphism of N. The structure of such groups is very well understood. By Thompson's Theorem [140] a finite group with a regular automorphism of

prime order is nilpotent and the Higman-Kreknin-Kostrikin Theorem provides a bound for the nilpotency class of N in terms of the least prime divisor of the order of A (Corollary 4.3.8 and Theorem 5.1.1). Beside this, since all elements of A are regular, all abelian subgroups of A are cyclic and this implies strong restrictions on the structure of A.

It is remarkable that the theory of finite groups admitting a partition seems to be isomorphically embedded in the theory of finite groups itself. Non-soluble finite groups with a partition are classified and this classification lies at the foundation of the classification theory of all finite simple groups.

Soluble finite groups with a partition modulo nilpotent groups are exhausted mainly by Frobenius groups and by semidirect products of the form $N \leftthreetimes \langle \varphi \rangle$ where φ is a splitting automorphism of prime order p of N (see § 7.1). We recall that Hughes and Thompson [46], using Thompson's fundamental work on normal p-complements [140], proved that in this situation N is always soluble. Kegel [54] complemented this result by proving that N is even nilpotent.

At the same time, as noted by Busarkin and Gorchakov in their book "Finite groups admitting a partition" [13], for a long time almost nothing was known about finite nilpotent groups with a partition except for certain counterexamples to the Hughes conjecture. In 1959 Hughes and Thompson [46] proved the Hughes conjecture for finite groups which are not p-groups. It was also proved in [46] that the proper Hughes subgroup of a finite group is soluble and in Kegel's work [54] it was proved that the proper Hughes subgroup of a finite group is nilpotent.

The Hughes problem for finite p-groups. As far as finite p-groups are concerned, positive results on the Hughes problem have been partial. The Hughes conjecture was proved for $p = 2$ (Hughes [44]), for $p = 3$ (Straus and Szekeres [136]), for metabelian groups (Hogan and Kappe [43]) and for p-groups of nilpotency class $2p - 2$ (Macdonald [99]).

The construction of counterexamples to the Hughes conjecture has been of great interest. The first one was constructed by Wall [147] in 1965 for $p = 5$ with the index of a non-trivial Hughes subgroup p^2. Later in 1973, Wall [148] showed that the existence of such counterexamples is connected with new non-$(p-1)$-Engel identities for the associated Lie algebra $L(B(n, p))$ (of characteristic p) of the free group $B(n, p)$ of exponent p. In fact, Wall discovered a new identity of degree $2p - 1$ which holds in this Lie algebra. However, the fact that it is really new in the sense that it is not a consequence of the $(p-1)$-Engel identity, has to date been established only for $p = 5, 7, 11$ with the aid of computer calculations [14]. (It is easy to show, that for $p = 2, 3$ all identities of $L(B(n, p))$ follow from the $(p-1)$-Engel identity which here implies nilpotency of class 1 and 3, respectively.)

The method of constructing universal counterexamples to the Hughes conjecture allowed us to achieve in [57] the best possible value $2p - 1$ for the nilpotency class of a counterexample (under the same hypothesis on $L(B(n, p))$. In [58] we

also showed that the consequence of weight $2p$ of the Wall's identity may be not a consequence of the $(p-1)$-Engel identity also in $L(B(2, p))$ and confirmed this for $p = 7$ with the aid of computer. (Note that in [147] the number of generators was ≥ 3, that for a 2-generator group there are no new relations in $L(B(2, p))$ of degree $2p - 1$, as shown by Kostrikin [75], and that for $p = 5$ there are no new relations in $L(B(2, 5))$ whatever, as shown by the computer aided calculations of Havas, Wall and Wamsley [38].) This new relation of degree $2p$ was a starting point for the construction in [58] of even a 2-generator counterexample to the Hughes conjecture which is a kind of a "monster" in the theory of finite p-groups, a finite p-group all of whose elements outside the Frattini subgroup have order p while it itself has elements of order p^2 (also for $p = 7$). Earlier in 1969 Macdonald attracted attention in a special article [98] to the question of the existence of such a group by showing that its existence would give negative answers to some other problems on finite p-groups (arising in [42, 50, 98]).

For 2-generator counterexamples the class $2p$ is best possible, according to Macdonald [99], and the prime $p = 7$ is least possible according to Vaughan-Lee's computer-aided calculations (private communication).

Recently, Vaughan-Lee [144] found all multilinear identities which hold in the Lie algebra $L(B(n, p))$. Wall [152] showed that only those of degrees $k(p-1)+1$ may be really new, that is, not consequences of those of smaller degrees. In the same paper Wall proved that if, for some r and all $k = 1, 2, \ldots, r$, Vaughan-Lee's identity of degree $k(p-1)+1$ is really new in this sense, then there exists a finite p-group with non-trivial Hughes subgroup of index p^r.

It is conjectured that all relations in $L(B(n, p))$ follow from its multilinear identities. To date little is known about which of Vaughan-Lee's identities are really new: only the above cases of Wall's identity of degree $2(p-1)+1$ for $p = 5, 7, 11$ and, of course, the $(p-1)$-Engel identity for all p (which is in fact equivalent to a multilinear one, as we have seen).

Of course, our Theorem 7.4.1 makes sense only if its hypothesis is satisfied for the prime p, that is, if all relations of $L(B(n, p))$ follow from a finite number of multilinear ones. It is quite unclear whether this is true or not. Note that recently G. Havas, M.F. Newman and M.R. Vaughan-Lee used a computer to show that for $p = 5$ the identity of degree $3(p-1)+1$ is not new in $L(B(3, 5))$, but this may be only because the number of generators is small.

There is also a possibility that the proof of Theorem 7.4.1 may be modified to yield bounds for the index of the Hughes subgroup in some special classes of p-groups.

We remark in addition that the interesting problem of describing the relations in $L(B(n, p))$ is connected also with other problems in the theory of finite p-groups. Under the same hypothesis as in Wall's theorem (that for some r for all $k = 1, 2, \ldots, r$ Vaughan-Lee's identity of degree $k(p-1)+1$ is really new) we have constructed in [62] examples of the so-called "secretive" finite p-groups P

of rank $r(p-1)+1$, that is, such that $|P^p| = p$ and $|P : \Omega_1(P)| = p^{r(p-1)+1}$ with $\Omega_1(P) = \Phi(P)$. So, for $p = 5, 7, 11$ and $r = 2$, this refutes the conjecture of Blackburn and Espuelas in [12] that if $|P^p| = p$ in a finite p-group P then $|P : \Omega_1(P)| \leq p^p$ (they had proved this to be so in the case of P metabelian). Earlier Wall [149] had constructed examples of secretive p-groups of rank p. It is worth noting that, although we were able to bound the index of the Hughes subgroup in Theorem 7.4.1, we could not bound the index $|P : \Omega_1(P)|$ in a p-group P with $|P^p| = p$ under the same hypothesis: that all relations of $L(B(n, p))$ follow from a finite number of multilinear ones.

In [62], as a corollary to the main theorem, we constructed examples of finite soluble, non-nilpotent groups in which the generalized Hughes subgroup

$$H_{p^2}(G) = \langle x \in G \mid x^{p^2} \neq 1 \rangle$$

is non-trivial and has index $p^{r(p-1)+1}$. This contrasts with the positive solution to the Hughes problem for finite groups which are not p-groups. Under the same assumptions it is also possible to construct examples of finite p-groups in which the generalized Hughes subgroup $H_{p^2}(P)$ is non-trivial and has index $p^{r(p-1)+1}$.

We note that Scoppola [128] showed that any finite p-group H may occur as a section of $G/H_{p^k}(G)$ for some finite p-group G and for some k *depending on* H.

We exhibit now the evidence for the existence of a bound in terms of p only for the nilpotency class of the factor-group by the *non-trivial* Hughes subgroup of a finite p-group. The hypothesis of the following theorem is a well-known conjecture which is known to be true for $p = 5$ (while the cases $p = 2$ and $p = 3$ are trivial, as we have seen above). This theorem appeared in the e-mail exchanges of the author with E.I. Zel'manov, its proof is based on Theorem 7.4.18.

7.5.1 Theorem. *If, for a given prime number p, there is a linear function of d bounding the nilpotency class of a d-generator $(p-1)$-Engel Lie algebra of characteristic p then the nilpotency class of a finite p-group by the non-trivial Hughes subgroup is bounded in terms of p only.*

Proof. Let $\alpha(p)d$ be the linear function from the hypothesis of the theorem. Suppose that P is a finite p-group such that the Hughes subgroup $H_p(P)$ is non-trivial and the nilpotency class of $P/H_p(P)$ is k. We fix some minimal system of generators of $P/H_p(P)$ and consider a non-trivial commutator of weight k in these generators. The number m of the generators involved in this commutator is obviously at most k. These m generators together with an element of order p^2 from $H_p(P)$ generate a group G which satisfies the hypothesis of Theorem 7.4.18 with parameters m and k for the number of generators of $G/H_p(G)$ and the class of $G/H_p(G)$, respectively.

By Theorem 7.4.18, the nilpotency class of G is at least $(p-1)(m+k(k+1)/2-1)+1$. On the other hand, the nilpotency class of G is at most $\alpha(p)(m+1)$, since

we may assume that $m \geq 2$ which implies, by Theorem 7.3.7, that the associated Lie ring of G is a $(p-1)$-Engel Lie algebra over $GF(p)$. Taking into account that $m \leq k$ we obtain that

$$(p-1)(m + k(k+1)/2 - 1) + 1 \leq \alpha(p)(m+1) \leq \alpha(p)(k+1).$$

Since the left-hand side is a quadratic function of k and the right-hand side is a linear one, this inequality implies that k is bounded in terms of p only.

7.5.2 Corollary. *There is a number c such that the nilpotency class of the factor-group of any 5-group by the non-trivial Hughes subgroup does not exceed c.*

Proof. It is known from the work of G. Highman on the Restricted Burnside Problem for exponent 5 that the hypothesis of Theorem 7.5.1 is satisfied for $p = 5$ with $\alpha(5) = 25$. (We note that recently G. Havas, M.F. Newman and M.R. Vaughan-Lee used a computer to show that $\alpha(5) = 6$ which, together with some additional information, yields that the factor-group of any 5-group by the non-trivial Hughes subgroup is, in fact, always abelian.)

Periodic compact groups. We shall consider topological compact groups all of whose elements have finite orders. Every such a group is profinite, that is, it is an inverse limit of a spectrum of finite groups. It is easy to show that every periodic compact group contains an open subset consisting of elements of the same order. This means that there is a subgroup of finite index such that some of its cosets consists of elements of equal orders. The following are well-known conjectures:

a) every periodic compact group is locally finite;

b) every periodic compact group is a group of finite exponent.

Recently, Zel'manov [161] proved conjecture a) using his solution of the Restricted Burnside Problem for groups of prime-power exponent p^k in [158-160]. Actually, he considered periodic pro-p-groups, that is, inverse limits of spectra of finite p-groups, while the reduction to this case is due to Wilson [154].

Conjecture b) remains unproved.

For periodic compact groups containing an open subset consisting of elements of prime order, Theorem 7.2.1 may be applied to prove both conjectures for this special case in a strengthened "uniform" setting.

7.5.3 Corollary. *Suppose that a periodic compact group contains an open subset consisting of elements of the same prime order. Then it is locally finite, is a group of bounded exponent and contains a subgroup of finite index which belongs to a locally nilpotent variety.*

Proof. If G is a group satisfying the hypothesis then it contains a subgroup K of finite index such that all elements in some coset Kg of K have prime order p. By

Poincaré's Theorem, K contains a normal subgroup N of finite index and, clearly, for the same g, all elements of the coset Ng also have order p. By (7.1.2) we obtain that for any $n \in N$

$$n \cdot n^{g^{-1}} \cdot n^{g^{-2}} \cdot \ldots \cdot n^{g^{-p+1}} = (ng)^p g^{-p} = 1.$$

Hence N admits a splitting automorphism φ of prime order p induced by conjugation under g^{-1}.

By definition of profinite groups, N is residually finite. Replacing the subgroups N_α of finite index by the φ-invariant subgroups $\bigcap_{i=1}^{p} N_\alpha^{\varphi^i}$, we may assume that there is a family of φ-invariant normal subgroups $\{K_\alpha\}$ of finite index with trivial intersection. Every finite factor-group N/K_α admits a splitting automorphism φ of prime order p and therefore is nilpotent by the Thompson-Hughes-Kegel Theorem [46, 54]. Hence, by Theorem 7.2.1, the nilpotency class of any d-generator subgroup of such a factor-group is (d, p)-bounded by $f(d, p)$. (More precisely, if a subgroup S is generated by d elements then $\langle S^{\langle\varphi\rangle}\rangle$ is φ-invariant and is generated by dp elements – or by the same d elements as a $\langle\varphi\rangle$-group. Theorem 7.2.1 may be applied to $\langle S^{\langle\varphi\rangle}\rangle$ and the nilpotency class of S does not exceed that of $\langle S^{\langle\varphi\rangle}\rangle$.) Since $\bigcap_\alpha K_\alpha = 1$, the same property is enjoyed by the whole group N. Thus N is the desired subgroup of finite index which belongs to a locally nilpotent variety.

Since N is also a periodic compact group, it contains an open subset consisting of elements of the same order n. This means that there is a subgroup $L \leq N$ of finite index such that all elements of some coset Lh, $h \in N$, have orders equal to n. Now, let l be an arbitrary element of L. As we have shown, the subgroup $\langle\{h, lh\}^{\langle\varphi\rangle}\rangle$ is nilpotent of p-bounded class $f(2, p)$ (or $f(2p, p)$ at the other interpretation of the number of generators). This subgroup is generated by elements of order n and hence it is a group of exponent dividing $n^{f(2,p)}$ by Corollary 2.5.4. But $\langle\{h, lh\}^{\langle\varphi\rangle}\rangle$ obviously contains l. Thus the order of an arbitrary element $l \in L$ divides the (n, p)-bounded number $n^{f(2,p)}$. Since the index of L in G is finite, G is a group of bounded exponent.

The local finiteness of G follows from the fact that it is periodic and almost locally nilpotent.

The corollary is proved.

We conjecture that, in general, every periodic compact group contains a subgroup of finite index which belongs to a locally soluble variety of bounded exponent.

Splitting automorphisms of composite order. This conjecture on periodic compact group would, perhaps, follow from the following generalization of Theorem 7.2.1 if it held: if, for a prime number p and for natural k and d, a nilpotent d-generator group G admits a splitting automorphism φ of order p^k, that is, such

that

$$x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p^k-1}} = 1$$

for all $x \in G$, then the derived length of G is bounded in terms of p, k and d. As in the case of prime order, the notion of a splitting automorphism of order p^k combines the notion of a group of exponent p^k with that of a group with a regular automorphism of order p^k. It seems that any proof of the generalization just stated must also involve using the properties of nilpotent groups from these two classes. Now, due to the work of Zel'manov, there is a positive solution to the Restricted Burnside Problem for groups of exponent p^k. However, as we noted in § 5.4, up to now, no upper bound for the derived length of nilpotent periodic groups with a regular automorphism of order p^k has yet been obtained. All the same, even in the case of order 4, where both properties of groups of exponent 4 and of groups with a regular automorphism of order 4 are well-known, we could not yet combine them to settle a conjecture on a splitting automorphism of order 4.

We note that E. Jabara has recently proved that a finite group with a splitting automorphism of order 4 is soluble.

Splitting automorphisms of prime order. If an arbitrary group admits a splitting automorphism φ of order 2 or 3 then it is nilpotent of class 1 or 3, respectively. For $|\varphi| = 2$ this is a simple exercise, while for $|\varphi| = 3$ it follows from calculations similar to those which prove that a group of exponent 3 is nilpotent of class ≤ 3.

The theories of nilpotent groups of prime exponent and of nilpotent groups with a regular automorphism are involved not only in the definition of a splitting automorphism of prime order, but also in the (second) proof of Theorem 7.2.1 in § 7.2 (we have noted before that the first proof in § 6.4 also gives an *a posteriori* bound for the order of the centralizer of the automorphism). This gives rise to the following interesting question. By Theorem 7.2.1 all locally nilpotent groups in the variety $\mathfrak{M}_p$ of $\langle\varphi\rangle$-groups (of groups with operators $\langle\varphi\rangle$) which consists of all groups with a splitting automorphism φ of prime order p, form a subvariety $LN\mathfrak{M}_p$. Is it true that $LN\mathfrak{M}_p$ is a join of a nilpotent subvariety $\mathfrak{N}_{c(p)} \cap \mathfrak{M}_p$ (where the nilpotency class $c(p)$ is, of course, p-bounded) and a subvariety $\mathfrak{B}_p \cap LN\mathfrak{M}_p$ of locally nilpotent groups of exponent p? The *join* of varieties is, by definition, the smallest variety which contains them.

This question may be formulated equivalently in terms of a free $\langle\varphi\rangle$-group F. Let N be the verbal $\langle\varphi\rangle$-subgroup of F generated as a verbal subgroup by the word $x \cdot x^{\varphi} \cdot x^{\varphi^2} \cdot \ldots \cdot x^{\varphi^{p-1}}$, that is, N is the normal closure

$$N = \langle \{ g \cdot g^{\varphi} \cdot g^{\varphi^2} \cdot \ldots \cdot g^{\varphi^{p-1}} \mid g \in F \}^F \rangle.$$

Is it true that there exists a p-bounded number $c(p)$ such that

$$N\gamma_n(F) = N\gamma_{c(p)+1}(F) \cap N\gamma_n(F)F^p$$

for all $n \geq c(p)+1$? If the answer to this question is affirmative then, in particular,

$$(\gamma_{c(p)+1}(G))^p = 1 \quad \text{and} \quad \gamma_{c(p)+1}(G^p) = 1$$

for any nilpotent group G in $\mathfrak{M}_p$. These statements are also interesting conjectures in their own right. At the moment we are only able to prove a somewhat weaker proposition in which, however, the bounds for the nilpotency class are best possible.

7.5.4 Corollary. *There exist p-bounded numbers $k(p)$ and $l(p)$ such that the following identities hold in any locally nilpotent group G in $\mathfrak{M}_p$:*

a) $[x_1^{p^{k(p)}}, x_2^{p^{k(p)}}, \ldots, x_{h+1}^{p^{k(p)}}] = 1$

(this means that $G^{p^{k(p)}}$ is nilpotent of class h, that is, $\gamma_{h+1}(G^{p^{k(p)}}) = 1$) and

b) $[x_1, x_2, \ldots, x_{h+1}]^{p^{l(p)}} = 1$,

where $h = h(p)$ is Higman's function.

The author does not know whether b) here may be replaced by an equation of the form $(\gamma_{h+1}(G))^{p^{m(p)}} = 1$.

Proof. By virtue of Mal'cev's Local Theorem, G may be assumed finitely generated and hence nilpotent and therefore residually finite. If $\{N_\alpha\}$ is a family of normal subgroups of finite indices with trivial intersection then $\left\{\bigcap_{i=1}^{p} N_\alpha^{\varphi^i}\right\}$ is a family of φ-invariant normal subgroups of finite indices with trivial intersection. It is clearly sufficient to prove the corollary for each of the finite factor-groups $G/\bigcap_{i=1}^{p} N_\alpha^{\varphi^i}$ which also belong to $\mathfrak{M}_p$. So, we need only prove the corollary assuming that G is finite and nilpotent.

The Hall p'-subgroup of G is nilpotent of class $\leq h(p)$, since φ is regular on it. We may therefore assume that G is a finite p-group.

We denote Higman's function by $h = h(p)$.

Let $x_1, x_2, \ldots, x_{h+1}$ be arbitrary elements of G. We apply Theorem 7.3.2 to the φ-invariant $p(h+1)$-generated nilpotent p-subgroup $H = \langle \{x_1, x_2, \ldots, x_{h+1}\}^{\langle\varphi\rangle} \rangle$ to obtain that H contains a subgroup of p-bounded index r which is nilpotent of class $\leq h$. By Poincaré's Theorem it contains a normal subgroup of index $\leq r!$ which, by Lagrange's Theorem, contains all elements $x_1^{p^{k(p)}}, x_2^{p^{k(p)}}, \ldots, x_{h+1}^{p^{k(p)}}$ for some p-bounded number $k(p) \leq r!$. Hence

$$[x_1^{p^{k(p)}}, x_2^{p^{k(p)}}, \ldots, x_{h+1}^{p^{k(p)}}] = 1.$$

But the elements of the form $x^{p^{k(p)}}$, $x \in G$, generate $G^{p^{k(p)}}$. Therefore this subgroup is nilpotent of class $\leq h$, since the identity of nilpotency may be verified on generators. We have proved a).

In view of the existence of a homomorphism of the tensor product of abelian groups

$$\underbrace{\gamma_1(H)/\gamma_2(H) \otimes \ldots \otimes \gamma_1(H)/\gamma_2(H)}_{h+1}$$

onto the factor-group $\gamma_{h+1}(H)/\gamma_{h+2}(H)$, the obtained identity a) implies that

$$[x_1, x_2, \ldots, x_{h+1}]^{p^{(h+1)k(p)}} \equiv 1 \pmod{\gamma_{h+2}(H)}.$$

Hence $\gamma_{h+1}(H)/\gamma_{h+2}(H)$ has exponent dividing $p^{(h+1)k(p)}$ and the same holds for all factors of the lower central series of H, starting from the $(h+1)$-st one. Since the nilpotency class of H is p-bounded by Theorem 7.2.1, $\gamma_{h(p)+1}(H)$ also has a p-bounded exponent. In particular, for some p-bounded number $l(p)$, the equality b) holds.

The corollary is proved.

This result also gives corresponding corollaries on the structure of finite p-group admitting a non-trivial partition by Proposition 7.1.1.

We finally point out one more application of Theorem 7.2.1 obtained by Kovács [80]: if a group G is locally a residually SI^*-group and admits an automorphism φ of prime order p such that $G = \{g^{-1} \cdot g^{\varphi} \mid g \in G\}$ then G is nilpotent of class $h(p)$. (The property SI^* is one of the generalizations of solubility.)

Chapter 8

Nilpotent p-groups admitting automorphisms of order p^k with few fixed points

The results of Chapters 5 and 7 give a complete, in a certain sense, picture of the structure of nilpotent groups with automorphisms of prime order close to regular (splitting or almost regular). As we indicated in the Comments in § 5.4 and 7.5, much less is known about nilpotent groups with such automorphisms of composite order. This chapter contains the first major breakthrough in this direction. In the "modular" case where a (locally) nilpotent p-group P admits an automorphism of order p^k with p^n fixed points, it is proved that P is almost soluble with a strong bound, in terms of p and k only, on the derived length of a subgroup of bounded index.

The proof is based on Kreknin's Theorem on Lie rings from Chapter 4. It uses a general group-theoretic corollary to Kreknin's Theorem, obtained with the help of the Mal'cev correspondence given by the Baker-Hausdorff formula. More precisely, it is proved in § 8.1 that if a nilpotent group G of class c admits an automorphism φ of finite order m then, for some (m, c)-bounded number $N = N(m, c)$ the subgroup $(G^N)^{(f(m))}$ is contained in $\langle C_G(\varphi)^G \rangle$, where $f(m)$ is the value of Kreknin's function from Theorem 4.3.2. Another important technique comes from the theory of powerful p-groups, especially, from Shalev's work [129], where a weak bound, in terms of p, k and n, for the derived length of P was obtained. Standard arguments show that, in the proof of the main theorem, P may be assumed to be a powerful finite p-group. We then prove, using Kreknin's Theorem again, that $P^{(f(p^k))}$ is nilpotent of (p, k, n)-bounded class. This allows us to apply the general corollary mentioned above. It may seem amazing that a combination of two "weak" results yields a "strong" one, as if one managed to lift oneself up by pulling the laces of one's boots.

The resulting bound for the derived length of a subgroup, which is actually obtained in the proof of the main theorem, is $2f(p^k)$. This is close to the apparently best possible value $f(p^k)$ (if we require that the function depends on the order of the automorphism only, compare with the problem at the end of § 8.4). On the other hand, we did not give ourself labour to record an explicit upper bound for the index of the subgroup, though such a bound may be easily obtained on the basis of the bound for Kreknin's function in Theorem 4.3.2, of the bound for the

index of a powerful subgroup in a finite p-group of a given characteristic rank in [96] and of the p-adic estimation of the denominators of the coefficients in the Baker-Hausdorff formula, which may be used for bounding the number $N(c, m)$ above.

The necessary preliminary material on the Mal'cev correspondence and on powerful p-groups is included without proofs in § 8.1 and 8.2, respectively.

§ 8.1 An application of the Mal'cev correspondence

In this section we derive the following corollary to Kreknin's theorem 4.3.2.

8.1.1 Theorem. *Suppose that G is a nilpotent group of class c and φ an automorphism of G of finite order m. Then, for some (c, m)-bounded number $N = N(c, m)$, the subgroup $(G^N)^{(f(m))}$ is contained in the normal closure $\langle C_G(\varphi)^G \rangle$ of the centralizer $C_G(\varphi)$.*

The proof of Theorem 8.1.1 uses the theory of Mal'cev completions of torsion-free nilpotent groups and the Mal'cev correspondence, which is the so-called categorical isomorphism between the category of radicable (complete) torsion-free nilpotent groups and the category of nilpotent $\mathbb{Q}$-Lie algebras and which is given by the Baker-Hausdorff formula, see [89, 106, 153]. It is actually more convenient to use a special form of this correspondence, where only π-roots are adjoined for an appropriate set of primes $\pi = \pi(c)$, which is sufficient for applying the Baker-Hausdorff formula to nilpotent groups of class c.

The Baker-Hausdorff formula appears when a free (nilpotent) group is presented in a free (nilpotent) associative non-commutative algebra $\mathcal{A}$ of formal power series over $\mathbb{Q}$. Let $x_1, x_2, \ldots$ be free generators of such an algebra $\mathcal{A}$. For any $x \in \mathcal{A}$, we set

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^i}{i!} + \ldots.$$

Then

$$e^x \cdot e^y = e^{H(x,y)},$$

where, as usual,

$$H(x, y) = \log(e^x \cdot e^y),$$

and where, by definition,

$$\log(1 + z) = z - \frac{z^2}{2} + \frac{z^3}{3} + \ldots + (-1)^{i+1}\frac{z^i}{i} + \ldots.$$

It is well-known that if $\mathcal{A}$ is a free nilpotent algebra of class c (that is, all monomials of degrees greater than c are 0), then the Lie algebra $\mathcal{L}$ generated by the x_i with respect to the Lie multiplication $[a, b] = ab - ba$ (where on the right-hand side the operations are in $\mathcal{A}$), is a free nilpotent Lie algebra of class c and the e^{x_i} freely generate a free nilpotent group F of class c (with respect to the multiplication in $\mathcal{A}$). The elements of $\mathcal{L}$ are called *Lie elements.*

The "Baker-Hausdorff formula" usually means a theorem stating that $H(x, y)$ is a Lie element in x and y, that is, that $H(x, y)$ belongs to the $\mathbb{Q}$-Lie algebra generated by x and y. Therefore, all elements of F have the form $e^l = e^{l(x_1, x_2, \ldots)}$, where $l = l(x_1, x_2, \ldots)$ is a Lie element.

We remark that taking an r-th power of e^l is equivalent to multiplying l by r: $(e^l)^r = e^{rl}$. One can adjoin to F all elements of the form e^{rl}, $e^l \in F$, $r \in \mathbb{Q}$. The resulting set $\tilde{F}$ may be shown to be a group which is nilpotent of the same class c. Since the same formula $(e^l)^r = e^{rl}$ holds for any $l \in \mathcal{L}$, $r \in \mathbb{Q}$, this group $\tilde{F}$ is *radicable* (or *complete),* which means that for every $a \in \tilde{F}$ and every $k \in \mathbb{Z}$ there is $b \in \tilde{F}$ such that $b^k = a$. Since $\tilde{F}$ is a torsion-free nilpotent group, the root b here is unique, by Theorem 2.6.1 a), so that $\tilde{F}$ may be also called a *group with unique roots*. The group $\tilde{F}$ may be regarded as an algebraic system, which has, besides group operations, unary operations of taking powers in $\mathbb{Q}$. Moreover, $\tilde{F}$ is, in fact, a free radicable nilpotent group of class c and it may be regarded as an abstract universal radicable closure of the free nilpotent group F which is unique up to isomorphism (a special case of the *Mal'cev completion*).

It can be shown that the mapping $l \to e^l$ (*the exponential map*) is a one-to-one correspondence between $\mathcal{L}$ and $\tilde{F}$. While the Baker-Hausdorff formula expresses the group operations in $\tilde{F}$ in terms of the Lie ring operation in $\mathcal{L}$, there are inversions of the Baker-Hausdorff formula which reconstruct the operations in the Lie algebra $\mathcal{L}$, addition and Lie bracket, in terms of the group operations in $\tilde{F}$. The exponential map is a so-called categorical isomorphism of $\mathcal{L}$ and $\tilde{F}$: every statement in terms of a $\mathbb{Q}$-Lie algebra $\mathcal{L}$ may be translated into the statement in terms of the radicable group $\tilde{F}$ and vice versa. In particular, the (radicable) subgroups in $\tilde{F}$ correspond precisely to the subalgebras in $\mathcal{L}$, a subgroup is normal in $\tilde{F}$ if and only if the corresponding subalgebra is an ideal in $\mathcal{L}$, the automorphisms of $\tilde{F}$ are exhausted by those which are induced by the automorphisms of $\mathcal{L}$, and so on. It may also be easily shown that the normal closure of a subset in $\tilde{F}$ corresponds to the ideal in $\mathcal{L}$ generated by this subset and that the terms of the lower central series of $\tilde{F}$ correspond to those of $\mathcal{L}$. The same is true for the terms of the derived series.

Since the operations are expressed by formulae in free nilpotent groups admitting exponents in $\mathbb{Q}$ and free nilpotent $\mathbb{Q}$-Lie algebras, the same correspondence may be established for arbitrary nilpotent radicable torsion-free groups and nilpotent $\mathbb{Q}$-Lie algebras. The sets may be simply identified: for any nilpotent $\mathbb{Q}$-Lie algebra L, one can define the structure of a radicable torsion-free nilpotent group G on

the same set $G = L$ with respect to the group operation $x \cdot y = H(x, y)$ and, vice versa, the structure of a $\mathbb{Q}$-Lie algebra may be defined on any radicable torsion-free nilpotent group using the inversions of the Baker-Hausdorff formula.

It is clear that if $\mathcal{A}$ is a nilpotent algebra of class c, then the denominators in the above formulae are divisible only by the primes not greater than c. It may be shown that the same is true for the inverse formulae. This allows us to establish an analogous category isomorphism between the π-radicable nilpotent torsion-free groups of class $\leq c$, where π is the set of all primes not greater than c, and the nilpotent $\mathbb{Q}_\pi$-Lie algebras of class $\leq c$, where $\mathbb{Q}_\pi$ is the subring of $\mathbb{Q}$ consisting of all rational numbers whose denominators are divisible only by the primes in π. (This generalization of the Mal'cev correspondence is due to Lazard [89].)

For a given set of prime numbers π, a *π-completion* of a free nilpotent group F may be constructed within the algebra $\mathcal{A}$ in essentially the same way: $\{e^{kl} \mid e^l \in F,\ k \in \mathbb{Q}_\pi\}$.

We note at last that the Mal'cev completions and the Mal'cev-Lazard correspondence may be defined for free nilpotent groups with operators, since for any group of operators Ω a free (nilpotent) Ω-group, as an abstract group, is also a free nilpotent group (see § 1.9).

Proof of Theorem 8.1.1. Let F be a free nilpotent $\langle\varphi\rangle$-group of class c on free generators $x_1, x_2, \ldots, x_{2^f}$, where $f = f(m)$ is the value of Kreknin's function from Theorem 4.3.2, that is, f is a number such that $H^{(f)} \subseteq {}_{id}\langle C_H(\psi)\rangle$ for any Lie ring H with an automorphism ψ of order m, provided $mH = H$ (one can take, in particular, $f(m) = 2^{m-1} - 1$). We set $\pi = \pi(c!) \cup \pi(m)$, where $\pi(c!)$ is the set of all primes not exceeding c and $\pi(m)$ is the set of prime divisors of m. Let $\hat{F}$ denote the π-completion of F and let L be the $\mathbb{Q}_\pi$-Lie algebra which corresponds to $\hat{F}$ under the category isomorphism of Mal'cev-Lazard given by the Baker-Hausdorff formula and its inverses. We may regard φ as an automorphism of L acting on the set $\hat{F} = L$ in the same way.

We have $L^{(f)} \leq {}_{id}\langle C_L(\varphi)\rangle$ by Kreknin's Theorem 4.3.2, since $mL = L$ by the choice of π. Since $C_{\hat{F}}(\varphi) = C_L(\varphi)$, we have also ${}_{id}\langle C_L(\varphi)\rangle = \langle C_{\hat{F}}(\varphi)^{\hat{F}}\rangle$, so that we get $\hat{F}^{(f)} \leq \langle C_{\hat{F}}(\varphi)^{\hat{F}}\rangle$ in terms of $\hat{F}$. In particular, the normal closure $\langle C_{\hat{F}}(\varphi)^{\hat{F}}\rangle$ contains the commutator $\delta_f(x_1, x_2, \ldots, x_{2^f})$ of weight 1 in each of the $x_1, x_2, \ldots, x_{2^f}$, which is the left-hand side of the identity of solubility of derived length f. We rewrite this fact as

$$\delta_f(x_1, x_2, \ldots, x_{2^f}) = c_1^{g_1} \cdot c_2^{g_2} \cdot \ldots \cdot c_s^{g_s}, \tag{8.1.2}$$

where $c_i \in C_{\hat{F}}(\varphi)$, $g_i \in \hat{F}$. This equation is, in fact, an identity in a free π-radicable nilpotent group $\hat{F}$ and the elements c_i and g_i depend only on c and m.

We claim that, for some (c, m)-bounded $n = n(c, m)$, the value $\delta_f(x_1^n, x_2^n, \ldots, x_{2f}^n)$ of the same commutator δ_f belongs to $\langle C_F(\varphi)^F \rangle$. We need the following

8.1.3 Lemma. *Let $g = g(x_1, x_2, \ldots, x_{2f})$ be an arbitrary element of $\hat{F}$ regarded as a group word in $x_1, x_2, \ldots, x_{2f}$ with exponents in $\mathbb{Q}_\pi$. Then there is a natural π-number s, depending on g (and on c and m), such that replacing all of the x_i by their s-th powers transforms g into an element $\tilde{g}$ of F. If $g \in C_{\hat{F}}(\varphi)$, then $\tilde{g} \in C_F(\varphi)$. All multiples of s inherit these properties of s.*

Proof. We use induction on the nilpotency class c. If $c = 1$, then $g(x_1^s, x_2^s, \ldots, x_{2f}^s) = g^s$, and $g^s \in F$ for some π-number s by the definition of the π-completion $\hat{F}$. It is clear that every multiple of s has the same property.

Now let $c > 1$. By induction hypothesis applied to the factor-group $\hat{F}/\gamma_c(\hat{F})$, which may be identified with the completion of $F/\gamma_c(F) = F/(F \cap \gamma_c(\hat{F})) \cong F\gamma_c(\hat{F})/\gamma_c(\hat{F})$, there is a π-number t such that the image of $g(x_1^t, x_2^t, \ldots, x_{2f}^t)$ belongs to the image of F. Hence

$$g(x_1^t, x_2^t, \ldots, x_{2f}^t) \cdot h(x_1, x_2, \ldots, x_{2f}) = z \in \gamma_c(\hat{F})$$

where $h(x_1, x_2, \ldots, x_{2f}) \in F$ (and $h(x_1, x_2, \ldots, x_{2f})$ may be taken to be a group word in the x_i with exponents in $\mathbb{Z}$) and where $z = z(x_1, x_2, \ldots, x_{2f})$ is a product of powers of commutators of weight c with exponents in $\mathbb{Q}_\pi$. It follows from the standard commutator identities that $z(x_1^s, x_2^s, \ldots, x_{2f}^s) = z^{s^c}$ in a nilpotent group of class c (see, for example, Lemma 6.1.11). Since $z^{s^c} \in F$ for some π-number s, we have

$$\tilde{g} = g(x_1^{st}, x_2^{st}, \ldots, x_{2f}^{st}) = h^{-1}(x_1^s, x_2^s, \ldots, x_{2f}^s) \cdot z^{s^c} \in F,$$

so that st is a π-number with the required property. Every multiple of st has the form stu, and we also have

$$g(x_1^{stu}, x_2^{stu}, \ldots, x_{2f}^{stu}) = h^{-1}(x_1^{su}, x_2^{su}, \ldots, x_{2f}^{su}) \cdot z^{s^c u^c} \in F.$$

In order to show that $\tilde{g} \in C_F(\varphi)$ if $g \in C_{\hat{F}}(\varphi)$, we consider the set $C_L(\varphi) = C_{\hat{F}}(\varphi)$ as a subalgebra of L. Recall that taking the s-th power in $\hat{F}$ is equivalent to multiplying by s in L: $F \ni x^s = sx \in L$. Since L is a free nilpotent $\langle\varphi\rangle$-Lie algebra over $\mathbb{Q}_\pi$, it is clear that $C_L(\varphi)$ is a homogeneous subalgebra of L. Hence any $g \in C_L(\varphi)$ may be expressed in the form

$$g = g_1 + g_2 + \ldots + g_c,$$

where g_i is a homogeneous element of weight i and $g_i \in C_L(\varphi)$ for all i. We now have

$$g(sx_1, sx_2, \ldots, sx_{2f}) = sg_1 + s^2 g_2 + \ldots + s^c g_c \in C_L(\varphi).$$

Hence $\tilde{g} \in C_L(\varphi)$ if $g \in C_L(\varphi)$. As a result $\tilde{g} \in C_{\tilde{F}}(\varphi) \cap F = C_F(\varphi)$, as required. The same argument is also valid for any multiple of s.

The lemma is proved.

We now apply Lemma 8.1.3 to each of the elements c_i, g_i occurring on the right-hand side of (8.1.2). Let N_1 be the least common multiple of all numbers which are given for the c_i, g_i by Lemma 8.1.3. Then

$$\delta_f(x_1^{N_1}, x_2^{N_1}, \ldots, x_{2f}^{N_1}) = \tilde{c}_1^{\tilde{g}_1} \cdot \tilde{c}_2^{\tilde{g}_2} \cdot \ldots \cdot \tilde{c}_s^{\tilde{g}_s}, \tag{8.1.4}$$

where $\tilde{c}_i \in C_F(\varphi)$, $\tilde{g}_i \in F$ for all i. This is also an identity in the free nilpotent group F.

It is natural to try to obtain consequences of the equation (8.1.4) by substituting arbitrary elements of the subgroup F^{N_1} instead of the $x_i^{N_1}$. However, although F^{N_1} is generated by the elements of the form g^{N_1}, the derived subgroup $(F^{N_1})^{(f)}$ may not be generated by the values of δ_f at these generating elements of F^{N_1} only. In order to overcome this difficulty, we use the following lemma of Blackburn [N. Blackburn, Conjugacy in nilpotent groups, *Proc. Amer. Math. Soc.* **16** (1965), 143-148] which was also anticipated in the work of Mal'cev [106].

8.1.5 Blackburn's Lemma. *For every prime number p and natural number c there exists a (c, p)-bounded number $b(c, p)$ such that, in any nilpotent group of class $\leq c$, any product of p^r-th powers of its elements is a $p^{r-b(c,p)}$-th power for any $r \geq b(c, p)$.*

We return to the proof of Theorem 8.1.1. If $N_1 = p_1^{k_1} p_2^{k_2} \ldots p_s^{k_s}$ is the decomposition of N_1 in the product of prime-powers, we put

$$N = p_1^{k_1 + b(c,p_1)} \cdot p_2^{k_2 + b(c,p_2)} \ldots p_s^{k_s + b(c,p_s)},$$

where the $b(c, p_i)$ are as in Lemma 8.1.5.

It is convenient to prove Theorem 8.1.1 at first for a free nilpotent $\langle\varphi\rangle$-group F_1 of class c. It is easy to see that any product of N-th powers of elements of F_1 is an N_1-th power of an element of F_1. Indeed, by Lemma 8.1.5, such a product is a $p_i^{k_i}$-th power for each i. We need only apply the following elementary lemma.

8.1.6 Lemma. *Suppose that for coprime numbers u and v an element g of F_1 is both a u-th and a v-th power. Then g is also a uv-th power.*

Proof. The completion $\tilde{F}_1$ of F_1 is a group with unique roots by Theorem 2.6.1. Let $h^{uv} = g$ for some $h \in \tilde{F}_1$. It suffices to show that $h \in F_1$. We have $(h^u)^v = g$ and hence $h^u \in F_1$ by the hypothesis, and also $(h^v)^u = g$ and hence $h^v \in F_1$ by the hypothesis. But there exist integers s and t such that $su + tv = 1$. Hence $h = h^{su+tv} \in F_1$, as required.

The lemma is proved.

The subgroup $(F_1^N)^{(f)}$ is generated by the values of δ_f at the products of N-th powers of the elements of F_1. Since, as shown above, such products are N_1-th powers, the subgroup $(F_1^N)^{(f)}$ is contained in the subgroup generated by the elements of the form $\delta_f(g_1^{N_1}, g_2^{N_1}, \ldots, g_{2^f}^{N_1})$, $g_i \in F_1$. An application of the homomorphism of F into F_1, which extends the mapping $x_i \to g_i$, $i = 1, 2, \ldots, 2^f$, to (8.1.4) shows that $\delta_f(g_1^{N_1}, g_2^{N_1}, \ldots, g_{2^f}^{N_1})$ belongs to $\langle C_{F_1}(\varphi)^{F_1}\rangle$ (which contains the image of $\langle C_F(\varphi)^F\rangle$) for any $g_1, g_2, \ldots, g_{2^f} \in F_1$. Hence $(F_1^N)^{(f)} \leq \langle C_{F_1}(\varphi)^{F_1}\rangle$.

Finally let G be an arbitrary nilpotent group of class $\leq c$ admitting an automorphism φ of order p^k. Since there is a homomorphism ϑ of F_1 onto G, we obtain that

$$(G^N)^{(f)} = \vartheta((F_1^N)^{(f)}) \leq \vartheta(\langle C_{F_1}(\varphi)^{F_1}\rangle) \leq \langle C_G(\varphi)^G\rangle,$$

as required.

The theorem is proved.

§ 8.2. Powerful p-groups

The theory of powerful p-groups, anticipated in Lazard's work [90] on analytic pro-p-groups, was recently created by Lubotzki and Mann [96]; see also the book [18]. It has already found a number of applications in the theories of pro-p-groups, residually finite groups, groups of finite rank and p-groups of given coclass. The notion of a powerful p-group seems to be a more successful attempt in defining a "more linear" class of finite p-groups than that of regular p-groups (and it is actually, in a sense, more general, since the p-th power of a regular p-group is a powerful p-group). One can say that the theory of powerful p-groups reflects the properties of the "linear part" of a finite p-group all of whose abelian sections have a given bound for their ranks. The relevance to finite p-groups admitting p-automorphisms with few fixed points is due to the fact that their ranks are bounded, by Corollary 1.7.4.

Definition. A normal subgroup N of a finite p-group G is said to be *powerfully embedded* in G if $N^p \geq [N, G]$ for p odd (if $N^4 \geq [N, G]$ for $p = 2$).

Definition. A finite p-group G is called *powerful* if it is powerfully embedded in itself.

The following basic lemma will be frequently used, often without reference.

8.2.1 Lemma [96, Theorem 1.2]. *If M and N are powerfully embedded subgroups of a finite p-group, then $[M, N]$ and M^p are also powerfully embedded subgroups.*

Shalev has strengthened Theorem 1.6 of [96] about powerfully embedded subgroups having proved the following remarkable formula (see also the proof of Lemma 2.6 in [A. Shalev, The structure of finite p-groups and a constructive proof of the coclass conjecture, *Preprint, Jerusalem Univ.*, 1992]).

8.2.2 Lemma [129, Lemma 3.1]. *If M and N are powerfully embedded subgroups of a finite p-group, then $[M, N]^p = [M^p, N]$.*

Definition. A group is said to be a group of (*special* or *sectional*, or *Mal'cev*) *rank r* if all of its finitely generated sections may be generated by r elements.

It is clear that the rank of a finite abelian group coincides with the minimal number of its generators (and so it was defined in § 1.1).

The following theorem is of fundamental importance for the study of groups of given rank.

8.2.3 Theorem [96]. a) *If G is a finite d-generator powerful p-group, then each subgroup of G may be generated by d elements.*

b) *If all characteristic subgroups of a finite p-group G may be generated by r elements then G contains a subgroup of (p, r)-bounded index, which is a powerful p-group of rank at most r.*

Although in [96] an analogue of b) is proved in which the bound on the ranks is imposed on all subgroups of G, it may be seen from the proof that it suffices to restrict the ranks of the characteristic subgroups only. We note also that in [96] there is an explicit upper bound for the index of a powerful subgroup which appears in Theorem 8.2.3 b).

8.2.4 Corollary. *If a finite p-group P admits an automorphism φ of order p^k having exactly p^n fixed points, then it contains a subgroup of (p, k, n)-bounded index, which is a powerful p-group of (p, k, n)-bounded rank.*

Proof. If H is a characteristic subgroup of P then, by Theorem 1.6.1, we have $|C_{H/\Phi(H)}(\varphi)| \leq p^n$ so that the rank of $H/\Phi(H)$ is at most np^k by Corollary 1.7.4.

Thus, all characteristic subgroups of P may be generated by np^k elements and, by Theorem 8.2.3 b), P contains a subgroup of (p, k, n)-bounded index, which is a powerful p-group of rank at most np^k.

We recall some further properties of powerful p-groups.

8.2.5 Lemma [96]. *If G is a powerful p-group then, for each i, the subgroup G^{p^i} generated by the p^i-th powers consists, in fact, of p^i-th powers, and $(G^{p^i})^{p^j} = G^{p^{i+j}}$ for all i, j.*

The sections $G^{p^i}/G^{p^{i+1}}$ are elementary abelian for all i. Taking the p-th powers of elements induces a homomorphism of $G^{p^i}/G^{p^{i-1}}$ onto $G^{p^{i+1}}/G^{p^{i+2}}$ and therefore $|G^{p^i}/G^{p^{i+1}}| \geq |G^{p^{i+1}}/G^{p^{i+2}}|$ for all i.

We now give the definition of a special class of powerful p-groups which possess even more linear properties.

Definition. Suppose that p^t is the exponent of a powerful p-group G, that is, t is the least number such that $G^{p^t} = 1$. If $|G^{p^i}/G^{p^{i+1}}| = |G^{p^{i+1}}/G^{p^{i+2}}|$ for all $i \leq t-2$ then G is said to be *uniformly powerful*.

We shall need the following property of such groups.

8.2.6 Lemma [18, § 4.1]. *Let G be a uniformly powerful p-group of exponent p^t. Then $x^{p^j} \in G^{p^j}$ implies $x \in G^{p^{j-i}}$ whenever $0 \leq i \leq j \leq t$.*

§ 8.3 A weak bound for the derived length

We consider here a special case, where P is a uniformly powerful finite p-group and φ an automorphism of P of order p^k having exactly p^n fixed points. For the rest of the chapter we fix notation $f = f(p^k)$ for the value of Kreknin's function from Theorem 4.3.2 (we may actually put $f = f(p^k) = 2^{p^k-1} - 1$).

8.3.1 Theorem. *Suppose that P is a uniformly powerful p-group admitting an automorphism φ of order p^k with exactly p^n fixed points. Then the f-th derived subgroup $P^{(f)}$ is nilpotent of (p, k, n)-bounded class.*

Proof. Taking s arbitrary we apply Kreknin's Theorem 4.3.2 to the associated Lie ring $L = L(P^{p^s})$ of the subgroup P^{p^s} which also admits the automorphism φ of order p^k . We obtain that

$$(p^k L)^{(f)} \subseteq {}_{id}\langle C_L(\varphi)\rangle. \tag{8.3.2}$$

Since $|C_{\gamma_i(P)/\gamma_{i+1}(P)}(\varphi)| \leq p^n$ for all i by Theorem 1.6.1, we have $p^n C_L(\varphi) = 0$ in additive notation and therefore $p^n id(C_L(\varphi)) = 0$. Applying this to (8.3.2) we get $p^{n+k2^f} L^{(f)} = 0$. In terms of the group P^{p^s} this means that

$$((P^{p^s})^{(f)})^{p^{n+k2^f}} \leq \gamma_{2^f+1}(P^{p^s}). \tag{8.3.3}$$

Repeated application of Lemmas 8.2.1 and 8.2.2 yields that

$$(P^{p^s})^{(b)} = (P^{(b)})^{p^{sb}} \qquad \text{and} \qquad \gamma_c(P^{p^d}) = (\gamma_c(P))^{p^{cd}}$$

for all $a, b, c, d \in \mathbb{N}$. Hence it follows from (8.3.3) that

$$(P^{(f)})^{p^{s2^f - n + k2^f}} \leq (\gamma_{2^f+1}(P))^{p^{s2^f - s}} \leq P^{p^{s2^f - s}}. \tag{8.3.4}$$

Let now p^t be the exponent of P. We take s to be the maximal number satisfying $s2^f + s \leq t$, that is, $s = \lfloor t/(2^f + 1) \rfloor$. It may be assumed that $s2^f + n + k2^f \leq s2^f + s$. Otherwise the exponent of P is bounded in terms of p, k and n. Since the rank is (p, k, n)-bounded, the order of P is then also (p, k, n)-bounded (see Lemma 8.2.5) and the result follows.

Thus, we may apply Lemma 8.2.6 to "cancel" the exponent p^{s2^f} in (8.3.4), that is, to obtain that

$$(P^{(f)})^{p^{n+k2^f}} \leq P^{p^s}.$$

By Lemma 8.2.2 we get also that $(P^{p^r})^{(f)} \leq P^{p^s}$ for some (p, k, n)-bounded number r.

We may assume that $s(2^f + 2) > (s + 1)(2^f + 1) \Leftrightarrow s > 2^f + 1$ (otherwise the order of P is (p, k, n)-bounded – see above). Applying Lemma 8.2.2 again we have

$$[P, \underbrace{(P^{p^r})^{(f)}, \ldots, (P^{p^r})^{(f)}}_{2^f+2}] \leq [P, \underbrace{P^{p^s}, \ldots, P^{p^s}}_{2^f+2}] \leq P^{p^{s2^f - 2s}} = 1.$$

The last equality holds since $s(2^f + 2) > (s + 1)(2^f + 1) > t$ by the choice of s. By Lemma 8.2.2, we now have

$$1 = [P, \underbrace{(P^{p^r})^{(f)}, \ldots, (P^{p^r})^{(f)}}_{2^f+2}] = [P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{2^f+2}]^{p^h},$$

where h is a (p, k, n)-bounded number. Since the rank is also (p, k, n)-bounded, this gives a (p, k, n)-bound for the order of $[P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{2^f+2}]$. Since P is a

nilpotent p-group, this normal subgroup is therefore contained in an appropriate term of the upper central series with (p, k, n)-bounded number. Hence

$$[P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{u}] = 1$$

for some (p, k, n)-bounded number u, which implies that $P^{(f)}$ is nilpotent of class u.

The theorem is proved.

This theorem already implies a weak bound, in terms of p, k and n, for the derived length of an arbitrary nilpotent p-group P admitting an automorphism of order p^k having exactly p^n fixed points. Indeed, P may be assumed to be a finite p-group (by Mal'cev's Local Theorem). By Corollary 8.2.4, P may be also assumed to be a powerful p-group. We now consider the inequalities

$$|P/P^p| \geq |P^p/P^{p^2}| \geq \ldots \geq |P^{p^i}/P^{p^{i+1}}| \geq \ldots \tag{8.3.5}$$

which hold for the orders of the elementary abelian sections $P^{p^i}/P^{p^{i+1}}$ by Lemma 8.2.5. Since, by Theorem 1.6.1, φ has at most p^n fixed points acting on each of the $P^{p^i}/P^{p^{i+1}}$, the ranks of the φ-invariant sections $P^{p^i}/P^{p^{i+1}}$ are bounded in terms of p, k and n. Therefore, strict inequalities separate the chain (8.3.5) into (p, k, n)-boundedly many segments with equalities. This means that P possesses a series of (p, k, n)-bounded length with uniformly powerful sections. Each of them satisfies Theorem 8.3.1 and this clearly gives the desired bound for the derived length of P.

Such a (p, k, n)-bounded solubility of P was earlier proved by Shalev [129] using a Lie ring of another kind defined by a uniformly powerful p-group (this construction comes from the theory of analytic pro-p-groups, see [18]). It is claimed in [129] that the usual associated Lie ring cannot be used in this situation. However, it is the usual associated Lie ring that we were using above, and it seems that both constructions carry essentially the same information. We note also that the idea of using the "cancellation property" of uniformly powerful p-groups given by Lemma 8.2.6 which is used in our proof goes back to Shalev [129].

§ 8.4 A strong bound for the derived length of a subgroup of bounded index

In order to obtain a strong bound for the derived length of a subgroup of bounded index, we first of all extend Theorem 8.3.1 to arbitrary powerful p-groups.

8.4.1 Theorem. *Suppose that P is a powerful p-group admitting an automorphism φ of order p^k with exactly p^n fixed points. Then the f-th derived subgroup $P^{(f)}$ is nilpotent of (p, k, n)-bounded class.*

Proof. Just like at the end of the preceding section, we consider inequalities

$$|P/P^p| \geq |P^p/P^{p^2}| \geq \ldots \geq |P^{p^i}/P^{p^{i+1}}| \geq \ldots \tag{8.4.2}$$

which hold for the orders of the elementary abelian sections $P^{p^i}/P^{p^{i+1}}$ by Lemma 8.2.5. Since, by Theorem 1.6.1, φ has at most p^n fixed points acting on each of the $P^{p^i}/P^{p^{i+1}}$, the ranks of the φ-invariant sections $P^{p^i}/P^{p^{i+1}}$ are bounded in terms of p, k and n. Therefore, strict inequalities separate the chain (8.4.2) into (p, k, n)-boundedly many segments with equalities. This means that P possesses a series of (p, k, n)-bounded length with uniformly powerful sections.

We use induction on the number of uniformly powerful sections of P, that is, on the number of strict inequalities in (8.4.2). The case where P is uniformly powerful itself is covered by Theorem 8.3.1.

Let now P^{p^m} be a uniformly powerful subgroup of P such that P/P^{p^m} has fewer "steps" – uniformly powerful sections – than P. Then by the induction hypothesis we have

$$[P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{c}] \leq P^{p^m} \tag{8.4.3}$$

for some (p, k, n)-bounded number c.

On the other hand, by Theorem 8.3.1, we have

$$[P^{p^m}, \underbrace{(P^{p^m})^{(f)}, \ldots, (P^{p^m})^{(f)}}_{u}] = 1$$

for a (p, k, n)-bounded number u. By repeated application of Lemma 8.2.2, we get

$$[P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{u}]^{p^{m(1+2^f u)}} = 1.$$

One further application of Lemma 8.2.2 yields

$$[P^{p^{m(1+2^f u)}}, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{u}] = 1. \tag{8.4.4}$$

But taking several, (p, k, n)-bounded, times the mutual commutator with $P^{(f)}$ of both parts of (8.4.3) we get

$$[P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{2c}] \leq [P^{p^m}, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{c}] = [P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{c}]^{p^m} \leq P^{p^{2m}}$$

and so on. By an obvious induction,

$$[P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{v}] \leq P^{p^{m(1+2^f u)}}$$

for the (p, k, n)-bounded number $v = c(1 + 2^f u)$. We need only use (8.4.4) to get

$$[P, \underbrace{P^{(f)}, \ldots, P^{(f)}}_{w}] = 1$$

for the (p, k, n)-bounded number $w = u + v$, as required.

The theorem is proved.

We now state the main result of this chapter.

8.4.5 Theorem. *If a locally nilpotent p-group P admits an automorphism of order p^k having exactly p^n fixed points, then it contains a subgroup of (p, k, n)-bounded index which is soluble of (p, k)-bounded derived length $2f(p^k)$.*

Proof. By Mal'cev's Local Theorem, we may assume P to be finitely generated and therefore finite. By Corollary 8.2.4, P may be assumed to be a powerful p-group. Then the derived subgroup $P^{(f)}$ is nilpotent of (p, k, n)-bounded class by Theorem 8.4.1.

Applying Theorem 8.1.1 to $P^{(f)}$ we get $((P^{(f)})^{p^l})^{(f)} \leq \langle C_{P^{(f)}}(\varphi)^{P^{(f)}} \rangle$ for some (p, k, n)-bounded number l. But the subgroup $\langle C_{P^{(f)}}(\varphi)^{P^{(f)}} \rangle$ is generated by the elements conjugate to the elements of $C_{P^{(f)}}(\varphi)$ all of which have order $\leq p^n$. Since it is also nilpotent of (p, k, n)-bounded class, the exponent of $\langle C_{P^{(f)}}(\varphi)^{P^{(f)}} \rangle$ is (p, k, n)-bounded by Corollary 2.5.4.

This implies that

$$(((P^{(f)})^{p^l})^{(f)})^{p^m} \leq \langle C_{P^{(f)}}(\varphi)^{P^{(f)}} \rangle^{p^m} = 1$$

for some (p, k, n)-bounded l and m. An application of Lemma 8.2.2 yields

$$(P^{p^q})^{(2f)} \leq (((P^{(f)})^{p^l})^{(f)})^{p^m} = 1$$

for some (p, k, n)-bounded number q. Since both the rank and the exponent of P/P^{p^q} are (p, k, n)-bounded, the order of $|P/P^{p^q}|$ is also (p, k, n)-bounded. Thus P^{p^q} is the desired subgroup of (p, k, n)-bounded index which is soluble of (p, k)-bounded derived length $2f$.

The theorem is proved.

Concluding this chapter, we remark that in the particular case where $n = 1$, that is, where the number of fixed points of an automorphism of order p^k acting on a finite p-group P is p, Kiming [72] and McKay [109] proved that P contains a subgroup of (p, k)-bounded index which is nilpotent of class 2. This gives rise to the following problem: Does there exist a function $d(m)$ depending on m only, such that every finite p-group admitting an automorphism of order p^k with exactly p^m fixed points contains a subgroup of (p, m, k)-bounded index which is soluble of derived length $d(m)$?

References

1. S.I. Adyan, *The Burnside problem and identities in groups*, "Nauka", Moscow, 1975 (Russian); English transl., Springer-Verlag, 1978.
2. J. Alperin, Automorphisms of solvable groups, *Proc. Amer. Math. Soc.* **13** (1962), 175-180.
3. R. Baer, Representation of groups as quotient groups. I, *Trans. Amer. Math. Soc.* **58** (1945), 295-419.
4. R. Baer, Partitionen endlicher Gruppen, *Math. Z.* **75**, (1960/61), 333-372.
5. R. Baer, Einfache Partitionen endlicher Gruppen mit nicht-trivialer Fittingscher Untergruppe, *Arch. Math. (Basel)* **12** (1961), 81-89.
6. G. Baumslag, *Lecture notes on nilpotent groups*, Providence, 1971.
7. S.D. Bell, B. Hartley, A note on fixed-point-free actions of finite groups, *Quart. J. Math. Oxford Ser. (2)* **41** (1990), 127-130.
8. T. Berger, Nilpotent fixed point free automorphism groups of solvable groups, *Math. Z.* **131** (1973), 305-312.
9. T. Berger, Hall-Higman type theorems. I, *Canad J. Math.*, **29** (1974), 513-531; II, *Trans. Amer. Math. Soc.* **205** (1975), 47-69; III, *Trans. Amer. Math. Soc.* **228** (1977), 47-83; IV, *Proc. Amer. Math. Soc.* **37** (1973), 317-325; V, *Pacific J. Math.* **73** (1977), 1-62; VI, J. Algebra, **51** (1978), 416-424; VII, *Proc. London Math. Soc. (3)* **31** (1975), 21-54.
10. N. Blackburn, On a special class of p-groups, *Acta Math.* **100** (1958), 45-92.
11. N. Blackburn, Some remarks on Černikov p-groups, *Illinois J. Math.* **6** (1962), 421-433.
12. N. Blackburn, A. Espuelas, The power structure of metabelian p-groups, *Proc. Amer. Math. Soc.* **92** (1984), 478-484.
13. V.M. Busarkin, Yu.M. Gorchakov, *Finite groups admitting a partition*, "Nauka", Moscow, 1968 (Russian).
14. J. Cannon, Some combinatorial and symbol manipulation programs in group theory, in: *Computational Problems in Abstract Algebra (Proc. Conf., Oxford, 1967)*, Pergamon Press, 1970, 199-203.
15. K.-H. Clemens, Fixpunktfreie Automorphismen endlicher Gruppen, *Diplomarbeit*, Mathematisches Institut der Albert-Ludwigs-Universität, Freiburg i. Br., 1978.
16. W. Cody, A Meier-Wunderli theorem for H_p-subgroups, *Arch. Math.* (Basel) **44** (1985), 493-502.
17. E.C. Dade, Carter subgroups and Fitting heights of finite soluble groups, *Illinois J. Math.* **13** (1972), 347-369.
18. J.D. Dixon, M.P.F. du Sautoy, A. Mann, D. Segal, *Analytic pro-p-groups*, Cambridge Univ. Press, 1991.

19. S. Donkin, Space groups and groups of prime power order. VIII. Pro-p-groups of finite coclass and p-adic Lie algebras, *J. Algebra* **111** (1987), 316-342.
20. W. Feit, J.G. Thompson, Solvability of groups of odd order, *Pacific J. Math.* **13** (1963), 773-1029.
21. H. Finken, J. Neubüser, W. Plesken, Space groups and groups of prime power order. II. Classification of space groups by finite factor-groups, *Arch. Math. (Basel)* **35** (1980), 203-209.
22. P. Fong, On orders of finite groups and centralizers of p-elements, *Osaka J. Math.* **13** (1976), 483-489.
23. Yu.M. Gorchakov, On the existence of abelian subgroups of infinite rank in locally soluble groups, *Dokl. Akad. Nauk SSSR* **156** (1964), 17-20 (Russian); English transl. *Soviet Math. Dokl.* **5** (1964), 591-594.
24. Yu.M. Gorchakov, On infinite Frobenius groups, *Algebra i Logika* **4** (1965), 15-29 (Russian).
25. D. Gorenstein, *Finite Groups*, Harper and Row, NY, 1968.
26. F. Gross, Solvable groups admitting a fixed-point-free automorphism of prime power order, *Proc. Amer. Math. Soc.* **17** (1966), 1440-1446.
27. M. Hall, *The theory of groups*, NY, MacMillan, 1959.
28. P. Hall, A contribution to the theory of groups of prime-power order, *Proc. London Math. Soc. Ser. (2)* **36** (1934), 29-95.
29. P. Hall, Finite-by-nilpotent groups, *Math. Proc. Cambridge Philos. Soc.* **52** (1956), 611-616.
30. P. Hall, Some sufficient conditions for a group to be nilpotent, *Illinois J. Math.* **2** (1958), 787-801.
31. P. Hall, On the finiteness of certain soluble groups, *Proc. London Math. Soc. (3)* **9** (1959), 595-622.
32. P. Hall, G. Higman, The p-length of a p-soluble group and reduction theorems for Burnside's problem, *Proc. London Math. Soc. (3)* **6** (1956), 1-42.
33. B. Hartley, Centralizers in locally finite groups, in: *Proc. Conf. Group Theory, Brixen/Bressanone, 1986, Lecture Notes in Math.* **1281**, Springer, 1987, 36-51.
34. B. Hartley, I.M. Isaacs, On characters and fixed points of coprime operator groups, *J. Algebra* **131** (1990), 342-458.
35. B. Hartley, Th. Meixner, Periodic groups in which the centralizer of an involution has bounded order, *J. Algebra* **64** (1980), 285-291.
36. B. Hartley, Th. Meixner, Finite soluble groups containing an element of prime order whose centralizer is small, *Arch. Math. (Basel)* **36** (1981), 211-213.
37. B. Hartley, V. Turau, Finite soluble groups admitting an automorphism of prime power order with few fixed points, *Math. Proc. Cambridge Philos. Soc.* **102** (1987), 431-441.
38. G. Havas, G.E. Wall, J.W. Wamsley, The two-generator restricted Burnside group of exponent five, *Bull. Austral. Math. Soc.* **10** (1974), 459-470.
39. P.J. Higgins, Lie rings satisfying the Engel condition, *Math. Proc. Cambridge Philos. Soc.* **50** (1954), 8-15.
40. G. Higman, Groups and rings which have automorphisms without non-trivial fixed elements, *J. London Math. Soc. (2)* **32** (1957), 321-334.

41. G. Higman, Some remarks on varieties of groups, *Quart. J. Math. Oxford Ser. (2)* **10** (1959), 165-178.
42. C.R. Hobby, Nearly regular p-groups, *Canad. J. Math.* **19** (1967), 520-522.
43. G.T. Hogan, W.P. Kappe, On the H_p-problem for finite p-groups, *Proc. Amer. Math. Soc.* **20** (1969), 450-454.
44. D.R. Hughes, Partial difference sets, *Amer. J. Math.* **78** (1956), 650-677.
45. D.R. Hughes, A research problem in group theory, *Bull. Amer. Math. Soc.* **63** (1957), 209.
46. D.R. Hughes, J.G. Thompson, The H_p-problem and the structure of H_p-groups, *Pacific J. Math.* **9** (1959), 1097-1101.
47. I. Hughes, Groups with fixed-point-free automorphisms, *C. R. Math. Rep. Acad. Sci. Canada* **7** (1985), 61-66.
48. B. Huppert, *Endliche Gruppen* I, Berlin et al., Springer, 1967.
49. B. Huppert, N. Blackburn, *Finite Groups* II, Berlin et al., Springer, 1982.
50. W. Kappe, Properties of groups related to the second center, *Math. Z.* **101** (1967), 356-368.
51. M.I. Kargapolov, Yu.I. Merzlyakov, *Fundamentals of the theory of groups*, 3-rd ed., "Nauka", Moscow, 1982 (Russian); English transl. of the 2-nd ed., Springer-Verlag, 1979.
52. O.H. Kegel, Nicht-einfache Partitionen endlicher Gruppen, *Arch. Math. (Basel)* **12** (1961), 170-175.
53. O.H. Kegel, Aufzählung der Partitionen endlicher Gruppen mit trivialer Fittingscher Untergruppe, *Arch. Math. (Basel)* **12** (1961), 409-412.
54. O.H. Kegel, Die Nilpotenz der H_p-Gruppen, *Math. Z.* **75** (1960/61), 373-376.
55. E.I. Khukhro, A soluble group admitting a regular splitting automorphism of prime order is nilpotent, *Algebra i Logika* **17** (1978), 611-618 (Russian); English transl., *Algebra and Logic* **17** (1979), 402-405.
56. E.I. Khukhro, Nilpotency of soluble groups admitting a splitting automorphism of prime order, *Algebra i Logika* **19** (1980), 118-129 (Russian); English transl., *Algebra and Logic* **19** (1981), 77-84.
57. E.I. Khukhro, On a connection between the Hughes conjecture and relations in finite groups of prime exponent, *Mat. Sb.* **116** (1981), 253-264 (Russian); English transl., *Math. USSR Sb.* **44** (1983), 227-237.
58. E.I. Khukhro, On the associated Lie ring of a free 2-generator group of prime exponent and on the Hughes conjecture for 2-generator p-groups, *Mat. Sb.* **118** (1982), 567-575 (Russian); English transl., *Math. USSR Sb.* **46** (1983), 571-579.
59. E.I. Khukhro, Finite p-groups admitting an automorphism of order p with a small number of fixed points, *Mat. Zametki* **38** (1985), 652-657 (Russian); English transl., *Math. Notes* **38** (1986), 867-870.
60. E.I. Khukhro, A new identity in the Lie ring of a free group of prime exponent, *Izv. Akad. Nauk SSSR Ser. Mat.* **50** (1986), 1308-1325 (Russian); English transl., *Math. USSR-Izv.* **29** (1987), 659-676.
61. E.I. Khukhro, Locally nilpotent groups admitting a splitting automorphism of prime order, *Mat. Sb.* **130** (1986), 120-127 (Russian); English transl., *Math. USSR-Sb.* **58** (1987), 119-126.

62. E.I. Khukhro, Finite p-groups close to groups of prime exponent, *Algebra i Logika* **25** (1986), 227-240 (Russian); English transl., *Algebra and Logic* **25** (1987), 143-153.
63. E.I. Khukhro, Nilpotent periodic groups with an almost regular automorphism of prime order, *Algebra i Logika* **26** (1987), 502-517 (Russian); English transl., *Algebra and Logic* **26** (1988), 299-310.
64. E.I. Khukhro, On the Hughes problem for finite p-groups, *Algebra i Logika* **26** (1987), 642-646 (Russian); English transl., *Algebra and Logic* **26** (1988), 398-401.
65. E.I. Khukhro, A remark on periodic compact groups, *Sibirsk. Mat. Zh.* **30** (1989), 187-190 (Russian); English transl., *Siberian Math. J.* **30** (1990), 493-496.
66. E.I. Khukhro, On the structure of finite p-groups admitting a partition, *Sibirsk. Mat. Zh.* **30** (1989), 208-218 (Russian); English transl., *Siberian Math. J.* **30** (1990), 1010-1019.
67. E.I. Khukhro, Groups and Lie rings admitting almost regular automorphisms of prime order, in: *Proc. Int. Conf. Theory of Groups, Bressanone/Brixen (Italy), June 11-17, 1989, Suppl. Rend. Circ. Mat. Palermo*, 1990, 183-192.
68. E.I. Khukhro, Groups and Lie rings admitting an almost regular automorphism of prime order, *Mat. Sb.* **181** (1990), 1207-1219 (Russian); English transl., *Math. USSR-Sb.* **71** (1992), 51-63.
69. E.I. Khukhro, Nilpotency in varieties of groups with operators, *Mat. Zametki* **50** (1991), 142-145 (Russian).
70. E.I. Khukhro, Local nilpotency in varieties of groups with operators, *Preprint, Univ. of Freiburg i. Br.*, Freiburg, 1992; *to appear in Mat. Sb.*, 1993.
71. E.I. Khukhro, Finite p-groups admitting p-automorphisms with few fixed points, *Preprint, Univ. of Freiburg i. Br.*, Freiburg, 1992; *to appear in Mat. Sb.*, 1993.
72. I. Kiming, Structure and derived length of finite p-groups possessing an automorphism of p-power order having exactly p fixed points, *Math. Scand.* **62** (1988), 153-172.
73. Yu.A. Kolmakov, Varieties of groups whose elements satisfy some conditions close to solubility, *Mat. Zametki* **35** (1984), 735-738 (Russian); English transl., *Math. Notes* **35** (1985), 389-391.
74. P.G. Kontorovich, On groups with bases of partition. I, *Mat. Sb.* **12** (1943), 56-70; II, *ibid.*, **19** (1946), 287-308; III, *ibid.*, 22 (1948), 79-100; IV, *ibid.*, **26** (1950), 311-320 (Russian).
75. A.I. Kostrikin, On the connection between periodic groups and Lie rings, *Izv. Akad. Nauk SSSR Ser. Mat.* **21** (1957), 289-310 (Russian); English transl., *Amer. Math. Soc. Transl. Ser. II* **45** (1965), 165-189.
76. A.I. Kostrikin, On the problem of Burnside, *Izv. Akad. Nauk SSSR Ser. Mat.* **23** (1959), 3-34 (Russian); *Amer. Math. Soc. Transl. Ser. II* **36** (1964), 63-99.
77. A.I. Kostrikin, *Introduction to algebra*, "Nauka", Moscow, 1977 (Russian); English transl., Springer-Verlag, 1982.
78. A.I. Kostrikin, *Around Burnside*, "Nauka", Moscow, 1986 (Russian); English transl., Springer-Verlag, 1990.
79. L.G. Kovács, Groups with regular automorphisms of order four, *Math. Z.* **75** (1961), 277-294.
80. L.G. Kovács, Groups with uniform automorphisms, in: *"Group theory and combinatorial geometry", Rend. Circ. Mat. Palermo (2)*, Suppl. no. 19 (1988), 125-133.

81. L.G. Kovács, J. Neubüser, B.H. Neumann, On finite groups with "hidden" primes, *J. Austral. Math. Soc. Ser. A* **12** (1971), 287-300.
82. V.A. Kreknin, A.I. Kostrikin, Lie algebras with regular automorphisms, *Dokl. Akad. Nauk SSSR* **149** (1963), 249-251 (Russian); English transl., *Soviet Math. Dokl.* **4** (1963), 355-358.
83. V.A. Kreknin, The solubility of Lie algebras with regular automorphisms of finite period, *Dokl. Akad. Nauk SSSR,* **150** (1963), 467-469 (Russian); English transl., *Soviet Math. Dokl.* **4** (1963), 683-685.
84. V.A. Kreknin, The solubility of Lie algebras with a regular automorphism, *Sibirsk. Mat. Zh.* **8** (1967), 715-716 (Russian); English transl., *Siberian Math. J.* **8** (1968), 536-537.
85. A.G. Kurosh, *The theory of groups, 3-rd ed.*, "Nauka", Moscow, 1967 (Russian); English transl. of the 1-st ed., Chelsey, New York, 1955.
86. A.G. Kurosh, *A course in higher algebra*, "Nauka", Moscow, 1973 (Russian).
87. H. Kurzweil, p-Automorphismen von auflösbaren p-Gruppen, *Math. Z.* **120** (1971), 326-354.
88. T.J. Laffey, The Hughes problem for exponent nine, *Math. Proc. Cambridge Philos. Soc.* **87** (1980), 393-399.
89. M. Lazard, Sur les groupes nilpotents et les anneaux de Lie, *Ann. Sci. École Norm. Supr.* **71** (1954), 101-190.
90. M. Lazard, Groupes analytiques p-adiques, *Inst. Hautes Études Sci. Publ. Math.* **26** (1965), 389-603.
91. C.R. Leedham-Green, Pro-p-groups of finite coclass, *Preprint.*
92. C.R. Leedham-Green, The structure of finite p-groups, *Preprint.*
93. C.R. Leedham-Green, S. McKay, On p-groups of maximal class. I, *Quart. J. Math. Oxford Ser. (2)* **27** (1976), 297-311.
94. C.R. Leedham-Green, S. McKay, W. Plesken, Space groups and groups of prime power order. V. A bound to the dimension of space groups with fixed coclass, *Proc. London Math. Soc.* **52** (1986), 73-94.
95. C.R. Leedham-Green, M.F. Newman, Space groups and groups of prime power order. I, *Arch. Math. (Basel)* **35** (1980), 193-202.
96. A. Lubotzky, A. Mann, Powerful p-groups. I: finite groups, *J. Algebra* **105** (1987), 484-505; II: p-adic analytic groups, *ibid.*, 506-515.
97. I.V. L'vov, E.I. Khukhro, Letter to the Editors on problem 5.56, *Kourovka Notebook*, 6-th ed., Novosibirsk, 1978 (Russian).
98. I.D. Macdonald, The Hughes problem and others, *J. Austral. Math. Soc. Ser. A* **10** (1969), 457-459.
99. I.D. Macdonald, Solution of the Hughes problem for finite p-groups of class $2p-2$, *Proc. Amer. Math. Soc.* **27** (1971), 39-42.
100. W. Magnus, A connection between the Baker-Hausdorff formula and a problem of Burnside, *Ann. of Math. (2)* **52** (1950), 111-126;
101. W. Magnus, *Errata, Ann. of Math. (2)* **57** (1953), 606.
102. W. Magnus, A. Karras, D. Solitar, *Combinatorial Group Theory*, Wiley, New York, 1966.
103. N.Yu. Makarenko, On almost regular automorphisms of prime order, *Sibirsk. Mat. Zh.* **33** (1992), 206-208 (Russian).

104. N.Yu. Makarenko, On nilpotent non-periodic groups with an almost regular automorphism of prime order, *to appear in Sibirsk. Mat. Zh.* (Russian).
105. A.I. Mal'cev, *Fundamentals of linear algebra*, "Nauka", Moscow, 1969 (Russian); English transl. of the 2-nd ed., Freeman, San Francisco – London, 1963.
106. A.I. Mal'cev, Nilpotent groups without torsion, *Izv. Akad. Nauk SSSR Ser. Mat.* **13** (1949), 201-212 (Russian).
107. A.I. Mal'cev, *Selected works. V. 1, Classical algebra; V. 2, Mathematical logic and the general theory of algebraic systems;* "Nauka", Moscow, 1976 (Russian).
108. A. Mann, Space groups and groups of prime power order. VII. Powerful p-groups and uncovered p-groups, *Bull. London Math. Soc.*, **24** (1992), 271-276.
109. S. McKay, On the structure of a special class of p-groups, *Quart. J. Math. Oxford Ser. (2)* **38** (1987), 489-502.
110. S. McKay, On the structure of a special class of p-groups. II, *Quart. J. Math. Oxford Ser. (2)* **41** (1990), 431-448.
111. Ju.A. Medvedev, Groups and Lie rings with an almost regular automorphism of prime order, *to appear in J. Algebra.*
112. H. Meier-Wunderli, Metabelsche Gruppen, *Comment. Math. Helv.* **25** (1951), 1-10.
113. Th. Meixner, *Über endliche Gruppen mit Automorphismen, deren Fixpunktgruppen beschränkt sind*, Dissertation, Univ. Erlangen-Nürnberg, 1979.
114. Th. Meixner, Solvable groups admitting an automorphism of prime power order whose centralizer is small, *J. Algebra* **99** (1986), 181-190.
115. Yu.I. Merzlyakov, On locally soluble groups of finite rank, *Algebra i Logika* **3** (1964), 5-16 (Russian).
116. P.S. Novikov, S.I. Adyan, Infinite periodic groups. I, II, III, *Izv. Akad. Nauk SSSR Ser. Mat.* **32** (1968), 212-244, 251-524, 709-731 (Russian); English transl., *Math. USSR Izv.* **2** (1969), 209-236, 241-480, 665-685.
117. A.Yu. Ol'shanskii, Groups of bounded exponent with subgroups of prime orders, *Algebra i Logika* **21** (1982), 553-618 (Russian); English transl., *Algebra and Logic* **21**, (1983), 369-418.
118. B.A. Panfërov, On nilpotent groups with lower central factors of minimal ranks, *Algebra i Logika* **19** (1980), 701-706 (Russian); English transl., *Algebra and Logic* **19** (1981), 455-458.
119. Yu.P. Razmyslov, On Engel Lie algebras, *Algebra i Logika* **10** (1971), 33-44 (Russian); English transl., *Algebra i Logika* **10** (1972), 21-29.
120. Yu.P. Razmyslov, On a problem of Hall and Higman, *Izv. Akad. Nauk SSSR Ser. Mat.* **42** (1978), 833-847 (Russian); English transl., *Math. USSR Izv.* **13** (1979), 133-146.
121. Yu.P. Razmyslov, *Identities of algebras and of their representations*, "Nauka", Moscow, 1989 (Russian).
122. D.J.S. Robinson, *Finiteness conditions and generalized soluble groups. Part 1*, Berlin et. al., Springer, 1972.
123. D.J.S. Robinson, *Finiteness conditions and generalized soluble groups. Part 2*, Berlin et. al., Springer, 1972.
124. I.N. Sanov, A solution of the Burnside problem for period 4, *Leningrad. Gos. Univ. Uchen. Zap. Ser. Mat. Nauk* **10** (1940), 166-170 (Russian).

125. I.N. Sanov, On a some system of relations in periodic groups with exponent a power of a prime number, *Izv. Akad. Nauk SSSR Ser. Mat.* **15** (1951), 477-502 (Russian).
126. I.N. Sanov, Establishment of a connection between periodic groups with period a prime number and Lie rings, *Izv. Akad. Nauk SSSR Ser. Mat.* **16** (1952), 23-58 (Russian).
127. I. Schur, Über die Darstellungen der endlichen Gruppen durch gebrochene lineare Substitutionen, *J. Reine Angew. Math.* **127** (1904), 20-50.
128. C.M. Scoppola, Groups of prime-power order as Frobenius-Wielandt complements, *Trans. Amer. Math. Soc.* **325** (1991), 855-874.
129. A. Shalev, On almost fixed point free automorphisms, *J. Algebra, to appear.*
130. A. Shalev, E.I. Zel'manov, Pro-p-groups of finite coclass, *Math. Proc. Cambridge Philos. Soc.* **111**, (1992), 417-421.
131. R. Shepherd, *Ph. D. Thesis*, Univ. of Chicago, 1971.
132. A.I. Shirshov, On free Lie rings, *Mat. Sb.* **45** (1958), 113-122 (Russian).
133. E. Shult, On groups admitting fixed point free abelian operator groups, *Illinois J. Math.* **9** (1965), 701-720.
134. P.V. Shumiatskii, Groups with regular elementary abelian 2-group of automorphisms, *Algebra i Logika* **27** (1988), 715-730 Russian); English transl., *Algebra and Logic* **27** (1989), 447-457.
135. V.P. Shunkov, On periodic groups with an almost regular involution, *Algebra i Logika* **11** (1972), 470-493 (Russian); English transl., *Algebra and Logic* **11** (1973), 260-272.
136. E.G. Straus, G. Szekeres, On a problem of D.R. Hughes, *Proc. Amer. Math. Soc.* **9** (1958), 157-158.
137. M. Suzuki, A new type of simple groups of finite order, *Proc. Nat. Acad. Sci. U.S.A.* **46** (1960), 868-870.
138. M. Suzuki, Finite groups with nilpotent centralizers, *Trans. Amer. Math. Soc.* **99** (1961), 425-470.
139. M. Suzuki, On a finite group with a partition, *Arch. Math. (Basel)* **12** (1961), 241-254.
140. J.G. Thompson, Finite groups with fixed-point-free automorphisms of prime order, *Proc. Nat. Acad. Sci. U.S.A.* **45** (1959), 578-581.
141. J.G. Thompson, Automorphisms of solvable groups, *J. Algebra* **1** (1964), 259-267.
142. A. Turull, Fitting heights of groups and of fixed points, *J. Algebra* **86** (1984), 555-566.
143. A. Turull, Groups of automorphisms and centralizers, *to appear.*
144. M.R. Vaughan-Lee, The restricted Burnside problem, *Bull. London Math. Soc.* **17** (1985), 113-133.
145. M.R. Vaughan-Lee, *The restricted Burnside problem*, Oxford, Clarendon Press, 1990.
146. M.R. Vaughan-Lee, J. Wiegold, Countable locally nilpotent groups of finite exponent without maximal subgroups, *Bull. London Math. Soc.* **14** (1981), 45-46.
147. G.E. Wall, On Hughes' H_p-problem, in: *Proc. Int. Conf. Theory of Groups, Canberra, 1965*, NY, Gordon and Breach, 1967, 357-362.
148. G.E. Wall, On the Lie ring of a group of prime exponent, in: *Proc. Second Int. Conf. Theory of Groups, Canberra, 1973, Lecture Notes in Math.* **372**, Springer, 1974, 667-690.
149. G.E. Wall, Secretive prime-power groups of large rank, *Bull. Austral. Math. Soc.* **12** (1975), 363-369.

150. G.E. Wall, On the Lie ring of a group of prime exponent. II, *Bull. Austral. Math. Soc.* **19** (1978), 11-28.
151. G.E. Wall, Lie methods in group theory, in: *Topics in Algebra (Proc. Eighteenth Summer Res. Inst. Austral. Math. Soc., Canberra, 1978), Lecture Notes in Math.* **697**, Springer, 1978, 137-173.
152. G.E. Wall, On the multilinear identities which hold in the Lie ring of a group of prime-power exponent, *J. Algebra* **104** (1986), 1-22.
153. R.B. Warfield, *Nilpotent groups*, Lecture Notes in Math. **513**, Springer, 1976.
154. J. Wilson, On the structure of compact torsion groups, *Monatsh. Math.* **96** (1983), 57-66.
155. H. Zassenhaus, Ein Verfahren, jeder endlichen p-Gruppe einen Lie-Ring mit der Charakteristik p zuzuordnen, *Abh. Math. Sem. Univ. Hamburg* **13** (1939), 200-207.
156. H. Zassenhaus, Liesche Ringe mit Primzahlcharacteristik, *Abh. Math. Sem. Univ. Hamburg* **13** (1939), 1-100.
157. E.I. Zel'manov, On Engel Lie algebras, *Sibirsk. Mat. Zh.* **29** (1988), 112-117 (Russian); English transl., *Siberian Math. J.* **29** (1989), 777-781.
158. E.I. Zel'manov, On some problems of the theory of groups and Lie algebras, *Mat. Sb.* **180** (1989), 159-167 (Russian); English transl., *Math. USSR Sb.* **66** (1990), 159-168.
159. E.I. Zel'manov, A solution of the Restricted Burnside Problem for groups of odd exponent, *Izv. Akad. Nauk SSSR Ser. Mat.* **54** (1990), 42-59 (Russian); English transl., *Math. USSR Izv.* **36** (1991), 41-60.
160. E.I. Zel'manov, A solution of the Restricted Burnside Problem for 2-groups, *Mat. Sb.* **182** (1991), 568-592 (Russian).
161. E.I. Zel'manov, On compact periodic groups, to appear in *Israel J. Math.*, 1992.

Index of Names

Subject Index

Walter de Gruyter
Berlin · New York

Klaus Doerk
Trevor Hawkes

Finite Soluble Groups

1992. XIV, 891 pages. 17 x 24 cm
ISBN 3-11-012892-6

de Gruyter Expositions in Mathematics, Vol. 4
Editors: *O. H. Kegel, V. P. Maslov, W. D. Neumann,* and *R. O. Wells, Jr.*

Contents:

Walter de Gruyter
Berlin · New York

Yu. A. Bahturin, A. A. Mikhalev,
V. M. Petrogradsky, M. V. Zaicev

Infinite Dimensional Lie Superalgebras

1992. 17 x 24 cm. X, 250 pages. Cloth DM 158,-
ISBN 3-11-012974-4

de Gruyter Expositions in Mathematics, Volume 7

Editors: *O. H. Kegel – V. P. Maslov – W. D. Neumann – R. O. Wells, Jr.*

Contents: