VOLUME 2

HANDBOOK OF ALGEBRA

M. HAZEWINKEL
EDITOR

NORTH-HOLLAND

HANDBOOK OF ALGEBRA
VOLUME 2

ELSEVIER
AMSTERDAM • LAUSANNE • NEW YORK • OXFORD • SHANNON • SINGAPORE • TOKYO

HANDBOOK OF ALGEBRA
Volume 2

edited by

M. HAZEWINKEL
CWI, Amsterdam

2000

ELSEVIER
AMSTERDAM • LAUSANNE • NEW YORK • OXFORD • SHANNON • SINGAPORE • TOKYO

ELSEVIER SCIENCE B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

First edition 2000

Library of Congress Cataloging-in-Publication Data

A catalog record from the Library of Congress has been applied for.

ISBN: 0 444 50396 X

⊗ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).

Printed in The Netherlands.

Preface

Basic philosophy

Algebra, as we know it today, consists of many different ideas, concepts and results. A reasonable estimate of the number of these different "items" would be somewhere between 50 000 and 200 000. Many of these have been named and many more could (and perhaps should) have a "name" or a convenient designation. Even the nonspecialist is likely to encounter most of these, either somewhere in the literature, disguised as a definition or a theorem, or to hear about them and feel the need for more information. If this happens, one should be able to find at least something in this Handbook; and hopefully enough to judge if it is worthwile to pursue the quest. In addition to the primary information, references to relevant articles, books or lecture notes should help the reader to complete his understanding. To make this possible, we have provided an index which is more extensive than usual and not limited to definitions, theorems and the like.

For the purpose of this Handbook, algebra has been defined, more or less arbitrarily as the union of the following areas of the Mathematics Subject Classification Scheme:

- 20 (Group theory)
- 19 (K-theory; will be treated at an intermediate level; a separate Handbook of K-theory which goes into far more detail than the section planned for this Handbook of Algebra is under consideration)
- 18 (Category theory and homological algebra; including some of the uses of categories in computer science, often classified somewhere in section 68)
- 17 (Nonassociative rings and algebras; especially Lie algebras)
- 16 (Associative rings and algebras)
- 15 (Linear and multilinear algebra, Matrix theory)
- 13 (Commutative rings and algebras; here there is a fine line to tread between commutative algebras and algebraic geometry; algebraic geometry is not a topic that will be dealt with in this Handbook; a separate Handbook on that topic is under consideration)
- 12 (Field theory and polynomials)
- 11 (As far as it used to be classified under old 12 (Algebraic number theory))
- 08 (General algebraic systems)
- 06 (Certain parts; but not topics specific to Boolean algebras as there is a separate three-volume Handbook of Boolean Algebras)

Planning

Originally, we hoped to cover the whole field in a systematic way. Volume 1 would be devoted to what we now call Section 1 (see below), Volume 2 to Section 2 and so on. A detailed and comprehensive plan was made in terms of topics which needed to be covered and authors to be invited. That turned out to be an inefficient approach. Different authors have different priorities and to wait for the last contribution to a volume, as planned originally, would have resulted in long delays. Therefore, we have opted for a dynamically evolving plan. This means that articles are published as they arrive and that the reader will find in this second volume articles from five different sections. The advantages of this scheme are two-fold: accepted articles will be published quickly and the outline of the series can be allowed to evolve as the various volumes are published. Suggestions from readers both as to topics to be covered and authors to be invited are most welcome and will be taken into serious consideration.

The list of the sections now looks as follows:

Section 1: Linear algebra. Fields. Algebraic number theory
Section 2: Category theory. Homological and homotopical algebra. Methods from logic
Section 3: Commutative and associative rings and algebras
Section 4: Other algebraic structures. Nonassociative rings and algebras. Commutative and associative rings and algebras with extra structure
Section 5: Groups and semigroups
Section 6: Representations and invariant theory
Section 7: Machine computation. Algorithms. Tables
Section 8: Applied algebra
Section 9: History of algebra

For a more detailed plan, the reader is referred to the Outline of the Series following the Preface.

The individual chapters

It is not the intention that the handbook as a whole can also be a substitute undergraduate or even graduate, textbook. The treatment of the various topics will be much too dense and professional for that. Basically, the level is graduate and up, and such material as can be found in P.M. Cohn's three-volume textbook "Algebra" (Wiley) will, as a rule, be assumed. An important function of the articles in this Handbook is to provide professional mathematicians working in a different area with sufficient information on the topic in question if and when it is needed.

Each chapter combines some of the features of both a graduate-level textbook and a research-level survey. Not all of the ingredients mentioned below will be appropriate in each case, but authors have been asked to include the following:

- Introduction (including motivation and historical remarks)
- Outline of the chapter

- Basic concepts, definitions, and results (proofs or ideas/sketches of the proofs are given when space permits)
- Comments on the relevance of the results, relations to other results, and applications
- Review of the relevant literature; possibly supplemented with the opinion of the author on recent developments and future directions
- Extensive bibliography (several hundred items will not be exceptional)

The present

Volume 1 appeared in December 1995 (copyright 1996). Volume 3 is scheduled for 2000. Thereafter, we aim at one volume per year.

The future

Of course, ideally, a comprehensive series of books like this should be interactive and have a hypertext structure to make finding material and navigation through it immediate and intuitive. It should also incorporate the various algorithms in implemented form as well as permit a certain amount of dialogue with the reader. Plans for such an interactive, hypertext, CD-Rom-based version certainly exist but the realization is still a nontrivial number of years in the future.

Bussum, June 1999 Michiel Hazewinkel

Kaum nennt man die Dinge beim richtigen Namen,
so verlieren sie ihren gefährlichen Zauber

(You have but to know an object by its proper name
for it to lose its dangerous magic)

E. Canetti

This Page Intentionally Left Blank

Outline of the Series

(as of June 1999)

Philosophy and principles of the Handbook of Algebra

Compared to the outline in Volume 1 this version differs in several aspects.

First, there is a major shift in emphasis away from completeness as far as more elementary material is concerned and towards more emphasis on recent developments and active areas.

Second, the plan is now more dynamic in that there is no longer a fixed list of topics to be covered, determined long in advance. Instead there is a more flexible nonrigid list that can change in response to new developments and availability of authors.

The new policy is therefore to work with a dynamic list of topics that should be covered, to arrange these in sections and larger groups according to the major divisions into which algebra falls, and to publish collections of contributions as they become available from the invited authors.

The coding by style below is as follows.

- **Author(s) in bold,** followed by the article title: chapters (articles) that have been received and are published or ready for publication;
- *italic*: chapters (articles) that are being written.
- plain text: topics that should be covered but for which no author has yet been definitely contracted.
- chapters that are included in volume 1 or volume 2 have a (1; xx pp.) or (2; xx pp.) after them, where xx is the number of pages.

Compared to the earlier outline the section on "Representation and invariant theory" has been thoroughly revised.

Section 1. Linear algebra. Fields. Algebraic number theory

A. Linear Algebra

G.P. Egorychev, Van der Waerden conjecture and applications (1; 22 pp.)
V.L. Girko, Random matrices (1; 52 pp.)
A. N. Malyshev, Matrix equations. Factorization of matrices (1; 38 pp.)
L. Rodman, Matrix functions (1; 38 pp.)
Linear inequalities (also involving matrices)

Orderings (partial and total) on vectors and matrices
Positive matrices
Special kinds of matrices such as Toeplitz and Hankel
Integral matrices. Matrices over other rings and fields

B. Linear (In)dependence

J.P.S. Kung, Matroids (1; 28 pp.)

C. Algebras Arising from Vector Spaces

Clifford algebras, related algebras, and applications

D. Fields, Galois Theory, and Algebraic Number Theory

(There is also an article on ordered fields in Section 4)
J.K. Deveney and J.N. Mordeson, Higher derivation Galois theory of inseparable field extensions (1; 34 pp.)
I. Fesenko, Complete discrete valuation fields. Abelian local class field theories (1; 48 pp.)
M. Jarden, Infinite Galois theory (1; 52 pp.)
R. Lidl and H. Niederreiter, Finite fields and their applications (1; 44 pp.)
W. Narkiewicz, Global class field theory (1; 30 pp.)
H. van Tilborg, Finite fields and error correcting codes (1; 28 pp.)
Skew fields and division rings. Brauer group
Topological and valued fields. Valuation theory
Zeta and L-functions of fields and related topics
Structure of Galois modules
Constructive Galois theory (realizations of groups as Galois groups)

E. Nonabelian Class Field Theory and the Langlands Program

(To be arranged in several chapters by Y. Ihara)

F. Generalizations of Fields and Related Objects

U. Hebisch and H.J. Weinert, Semi-rings and semi-fields (1; 38 pp.)
G. Pilz, Near rings and near fields (1; 36 pp.)

Section 2. Category theory. Homological and homotopical algebra. Methods from logic

A. Category Theory

S. MacLane and I. Moerdijk, Topos theory (1; 28 pp.)
R. Street, Categorical structures (1; 50 pp.)
Algebraic theories
B. Plotkin, Algebra, categories and databases (2; 68 pp.)
P.J. Scott, Some aspects of categories in computer science (2; 73 pp.)

B. Homological Algebra. Cohomology. Cohomological Methods in Algebra. Homotopical Algebra

J.F. Carlson, The cohomology of groups (1; 30 pp.)
A. Generalov, Relative homological algebra. Cohomology of categories, posets, and coalgebras (1; 28 pp.)
J.F. Jardine, Homotopy and homotopical algebra (1; 32 pp.)
B. Keller, Derived categories and their uses (1; 32 pp.)
A.Ya. Helemskii, Homology for the algebras of analysis (2; 122 pp.)
Galois cohomology
Cohomology of commutative and associative algebras
Cohomology of Lie algebras
Cohomology of group schemes

C. Algebraic K-theory

Algebraic K-theory: the classical functors K_0, K_1, K_2
Algebraic K-theory: the higher K-functors
Grothendieck groups
K_2 *and symbols*
KK-theory and EXT
Hilbert C^**-modules*
Index theory for elliptic operators over C^* *algebras*
Algebraic K*-theory (including the higher* K_n*)*
Simplicial algebraic K*-theory*
Chern character in algebraic K*-theory*
Noncommutative differential geometry
K-theory of noncommutative rings
Algebraic L-theory
Cyclic cohomology

D. Model Theoretic Algebra

Methods of logic in algebra (general)
Logical properties of fields and applications
Recursive algebras
Logical properties of Boolean algebras
F. Wagner, Stable groups (2; 40 pp.)

E. Rings up to Homotopy

Rings up to homotopy

Section 3. Commutative and associative rings and algebras

A. Commutative Rings and Algebras

J.P. Lafon, Ideals and modules (1; 24 pp.)

General theory. Radicals, prime ideals etc. Local rings (general). Finiteness and chain conditions

Extensions. Galois theory of rings

Modules with quadratic form

Homological algebra and commutative rings. Ext, Tor, etc. Special properties (p.i.d., factorial, Gorenstein, Cohen–Macauley, Bezout, Fatou, Japanese, excellent, Ore, Prüfer, Dedekind, ... and their interrelations)

D. Popescu, Artin approximation (2; 34 pp.)

Finite commutative rings and algebras. (See also Section 3B)

Localization. Local–global theory

Rings associated to combinatorial and partial order structures (straightening laws, Hodge algebras, shellability, ...)

Witt rings, real spectra

B. Associative Rings and Algebras

P.M. Cohn, Polynomial and power series rings. Free algebras, firs and semifirs (1; 30 pp.)

Classification of Artinian algebras and rings

V.K. Kharchenko, Simple, prime, and semi-prime rings (1; 52 pp.)

A. van den Essen, Algebraic microlocalization and modules with regular singularities over filtered rings (1; 28 pp.)

F. Van Oystaeyen, Separable algebras (2; 43 pp.)

K. Yamagata, Frobenius rings (1; 48 pp.)

V.K. Kharchenko, Fixed rings and noncommutative invariant theory (2; 38 pp.)

General theory of associative rings and algebras

Rings of quotients. Noncommutative localization. Torsion theories

von Neumann regular rings

Semi-regular and pi-regular rings

Lattices of submodules

A.A. Tuganbaev, Modules with distributive submodule lattice (2; 16 pp.)

A.A. Tuganbaev, Serial and distributive modules and rings (2; 19 pp.)

PI rings

Generalized identities

Endomorphism rings, rings of linear transformations, matrix rings

Homological classification of (noncommutative) rings

Group rings and algebras

Dimension theory

Duality. Morita-duality

Commutants of differential operators

Rings of differential operators

Graded and filtered rings and modules (also commutative)

Goldie's theorem, Noetherian rings and related rings

Sheaves in ring theory

A.A. Tuganbaev, Modules with the exchange property and exchange rings (2; 19 pp.)

Finite associative rings (see also Section 3A)

C. Co-algebras

D. Deformation Theory of Rings and Algebras (Including Lie Algebras)

Deformation theory of rings and algebras (general)
Yu. Khakimdjanov, Varieties of Lie algebras (2; 31 pp.)

Section 4. Other algebraic structures. Nonassociative rings and algebras. Commutative and associative algebras with extra structure

A. Lattices and Partially Ordered Sets

Lattices and partially ordered sets
Frames, locales, quantales

B. Boolean Algebras

C. Universal Algebra

D. Varieties of Algebras, Groups, ... (See also Section 3D)

V.A. Artamonov, Varieties of algebras (2; 29 pp.)
Varieties of groups
Quasi-varieties
Varieties of semigroups

E. Lie Algebras

Yu.A. Bahturin, A.A. Mikhalev and M. Zaicev, Infinite-dimensional Lie superalgebras (2; 34 pp.)
General structure theory
Free Lie algebras
Classification theory of semisimple Lie algebras over **R** and **C**
The exceptional Lie algebras
M. Goze and Yu. Khakimdjanov, Nilpotent and solvable Lie algebras (2; 47 pp.)
Universal envelopping algebras
Modular (ss) Lie algebras (including classification)
Infinite-dimensional Lie algebras (general)
Kac–Moody Lie algebras

F. Jordan Algebras (Finite and infinite dimensional and including their cohomology theory)

G. Other Nonassociative Algebras (Malcev, alternative, Lie admissable, ...)

Mal'tsev algebras
Alternative algebras

H. Rings and Algebras with Additional Structure

Graded and super algebras (commutative, associative; for Lie superalgebras, see Section 4E)
Topological rings
Hopf algebras
Quantum groups
Formal groups
λ-rings, γ-rings, ...
Ordered and lattice-ordered groups, rings and algebras
Rings and algebras with involution. C^*-algebras
Difference and differential algebra. Abstract (and p-adic) differential equations. Differential extensions
Ordered fields

I. Witt Vectors

Witt vectors and symmetric functions. Leibniz Hopf algebra and quasi-symmetic functions

Section 5. Groups and semigroups

A. Groups

A.V. Mikhalev and A.P. Mishina, Infinite Abelian groups: Methods and results (2; 36 pp.)
Simple groups, sporadic groups
Abstract (finite) groups. Structure theory. Special subgroups. Extensions and decompositions
Solvable groups, nilpotent groups, p-groups
Infinite soluble groups
Word problems
Burnside problem
Combinatorial group theory
Free groups (including actions on trees)
Formations
Infinite groups. Local properties
Algebraic groups. The classical groups. Chevalley groups
Chevalley groups over rings
The infinite dimensional classical groups
Other groups of matrices. Discrete subgroups
Reflection groups. Coxeter groups
Groups with BN-pair, Tits buildings, ...
Groups and (finite combinatorial) geometry
“Additive” group theory
Probabilistic techniques and results in group theory
Braid groups

B. Semigroups

Semigroup theory. Ideals, radicals, structure theory
Semigroups and automata theory and linguistics

C. Algebraic Formal Language Theory. Combinatorics of Words

D. Loops, Quasigroups, Heaps, ...

E. Combinatorial Group Theory and Topology

Section 6. Representation and invariant theory

A. Representation Theory. General

Representation theory of rings, groups, algebras (general)
Modular representation theory (general)
Representations of Lie groups and Lie algebras. General

B. Representation Theory of Finite and Discrete Groups and Algebras

Representation theory of finite groups in characteristic zero
Modular representation theory of finite groups. Blocks
Representation theory of the symmetric groups (both in characteristic zero and modular)
Representation theory of the finite Chevalley groups (both in characteristic zero and modular
Modular representation theory of Lie algebras

C. Representation Theory of 'Continuous Groups' (Linear Algebraic Groups, Lie Groups, Loop Groups, ...) and the Corresponding Algebras

Representation theory of compact topolgical groups
Representation theory of locally compact topological groups
Representation theory of $SL_2(\mathbf{R})$, ...
Representation theory of the classical groups. Classical invariant theory
Classical and transcendental invariant theory
Reductive groups and their representation theory
Unitary representation theory of Lie groups
Finite-dimensional representation theory of the ss Lie algebras (in characteristic zero); structure theory of semi-simple Lie algebras
Infinite dimensional representation theory of ss Lie algebras. Verma modules
Representation of Lie algebras. Analytic methods
Representations of solvable and nilpotent Lie algebras. The Kirillov orbit method
Orbit method, Dixmier map, ... for ss Lie algebras
Representation theory of the exceptional Lie groups and Lie algebras
Representation theory of 'classical' quantum groups
A.U. Klimyk, Infinite-dimensional representations of quantum algebras (2; 27 pp.)
Duality in representation theory

Representation theory of loop groups and higher dimensional analogues, gauge groups, and current algebras
Representation theory of Kac–Moody algebras
Invariants of nonlinear representations of Lie groups
Representation theory of infinite-dimensional groups like GL_∞
Metaplectic representation theory

D. Representation Theory of Algebras

Representations of rings and algebras by sections of sheafs
Representation theory of algebras (Quivers, Auslander–Reiten sequences, almost split sequences, ...)

E. Abstract and Functorial Representation Theory

Abstract representation theory
S. Bouc, Burnside rings (2; 64 pp.)
P. Webb, A guide to Mackey functors (2; 30 pp.)

F. Representation Theory and Combinatorics

G. Representations of Semigroups

Representation of discrete semigroups
Representations of Lie semigroups

Section 7. Machine computation. Algorithms. Tables

Some notes on this volume: Besides some general article(s) on machine computation in algebra, this volume should contain specific articles on the computational aspects of the various larger topics occurring in the main volume, as well as the basic corresponding tables. There should also be a general survey on the various available symbolic algebra computation packages.

The CoCoA computer algebra system

Section 8. Applied algebra

Section 9. History of algebra

Contents

List of Contributors

Artamonov, V.A., *Moscow State University, Moscow*
Bahturin, Yu., *Memorial University of Newfoundland, St. John's, NF, and Moscow State University, Moscow*
Bouc, S., *Université Paris 7-Denis Diderot, Paris*
Goze, M., *Université de Haute Alsace, Mulhouse*
Helemskii, A.Ya., *Moscow State University, Moscow*
Khakimdjanov, Yu., *Université de Haute Alsace, Mulhouse*
Kharchenko, V.K., *Universidad Nacional Autónoma de México, México, and Sobolev Institute of Mathematics, Novosibirsk*
Klimyk, A.U., *Institute for Theoretical Physics, Kiev*
Mikhalev, A.A., *The University of Hong Kong, Pokfulam Road*
Mikhalev, A.V., *Moscow State University, Moscow*
Mishina, A.P., *Moscow State University, Moscow*
Plotkin, B., *Hebrew University, Jerusalem*
Popescu, D., *University of Bucharest, Bucharest*
Scott, P.J., *University of Ottawa, Ottawa, ON*
Tuganbaev, A.A., *Moscow State University, Moscow*
Van Oystaeyen, F., *University of Antwerp, Wilrijk*
Wagner, F., *University of Oxford, Oxford*
Webb, P., *University of Minnesota, Minneapolis, MN*
Zaicev, M., *Moscow State University, Moscow*

This Page Intentionally Left Blank

Section 2A
Category Theory

This Page Intentionally Left Blank

Some Aspects of Categories in Computer Science

P.J. Scott

Department of Mathematics, University of Ottawa, Ottawa, Ontario, Canada

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

1. Introduction

Over the past 25 years, category theory has become an increasingly significant conceptual and practical tool in many areas of computer science. There are major conferences and journals devoted wholly or partially to applying categorical methods to computing. At the same time, the close connections of computer science to logic have seen categorical logic (developed in the 1970's) fruitfully applied in significant ways in both theory and practice.

Given the rapid and enormous development of the subject and the availability of suitable graduate texts and specialized survey articles, we shall only examine a few of the areas that appear to the author to have conceptual and mathematical interest to the readers of this Handbook. Along with the many references in the text, the reader is urged to examine the final section (Literature Notes) where we reference omitted important areas, as well as the Bibliography.

We shall begin by discussing the close connections of certain closed categories with typed lambda calculi on the one hand, and with the proof theory of various logics on the other. It cannot be overemphasized that modern computer science heavily uses formal syntax but we shall try to tread lightly. The so-called Curry–Howard isomorphism (which identifies formal proofs with lambda terms, hence with arrows in certain free categories) is the cornerstone of modern programming language semantics and simply cannot be overlooked.

NOTATION. We often elide composition symbols, writing $gf : A \to C$ for $g \circ f : A \to C$, whenever $f : A \to B$ and $g : B \to C$. To save some space, we have omitted large numbers of routine diagrams, which the reader can find in the sources referenced.

2. Categories, lambda calculi, and formulas-as-types

2.1. *Cartesian closed categories*

Cartesian closed categories (ccc's) were developed in the 1960's by F.W. Lawvere [Law66, Law69]. Both Lawvere and Lambek [L74] stressed their connections to Church's lambda calculus, as well as to intuitionistic proof theory. In the 1970's, work of Dana Scott and Gordon Plotkin established their fundamental role in the semantics of programming languages. A precise equivalence between these three notions (ccc's, typed lambda calculi, and intuitionistic proof theory) was published in Lambek and Scott [LS86]. We recall the appropriate definitions:

DEFINITION 2.1.

(i) A *Cartesian category* $\mathcal{C}$ is a category with distinguished finite products (equivalently, binary products and a terminal object $\mathbf{1}$). This says there are isomorphisms (natural in A, B, C)

$$Hom_{\mathcal{C}}(A, \mathbf{1}) \cong \{*\}, \tag{1}$$

$$Hom_{\mathcal{C}}(C, A \times B) \cong Hom_{\mathcal{C}}(C, A) \times Hom_{\mathcal{C}}(C, B). \tag{2}$$

(ii) A *Cartesian closed category* $\mathcal{C}$ is a Cartesian category $\mathcal{C}$ such that, for each object $A \in \mathcal{C}$, the functor $(-) \times A : \mathcal{C} \to \mathcal{C}$ has a specified right adjoint, denoted $(-)^A$. That is, there is an isomorphism (natural in B and C)

$$Hom_{\mathcal{C}}(C \times A, B) \cong Hom_{\mathcal{C}}\big(C, B^A\big). \tag{3}$$

For many purposes in computer science, it is often useful to have categories with explicitly given *strict* structure along with *strict* functors that preserve everything on the nose. We may present such ccc's equationally, in the spirit of multisorted universal algebra. The arrows and equations are summarized in Figure 1. These equations determine the isomorphisms (1), (2), and (3). In this presentation we say the structure is *strict*, meaning there is only one object representing each of the above constructs $\mathbf{1}$, $A \times B$, B^A. The exponential object B^A is often called the *function space* of A and B. In the computer science literature, the function space is often denoted $A \Rightarrow B$, while the arrow $C \xrightarrow{f^*} B^A$ is often called *currying* of f.

REMARK 2.2. Following most categorical logic and computer science literature, we do not assume ccc's have finite limits [Law69,LS86,AC98,Mit96]), in order to keep the correspondence with simply typed lambda calculi, cf. Theorem 2.20 below. Earlier books (cf. [Mac71]) do not always follow this convention.

Let us list some useful examples of Cartesian closed categories: for details see [LS86, Mit96,Mac71]

Objects	Distinguished Arrow(s)	Equations
Terminal $\mathbf{1}$	$A \xrightarrow{!_A} \mathbf{1}$	$!_A = f$, $f : A \to \mathbf{1}$
Products $A \times B$	$\pi_1^{A,B} : A \times B \to A$ $\pi_2^{A,B} : A \times B \to B$ $\dfrac{C \xrightarrow{f} A \quad C \xrightarrow{g} B}{C \xrightarrow{\langle f,g \rangle} A \times B}$	$\pi_1 \circ \langle f, g\rangle = f$ $\pi_2 \circ \langle f, g\rangle = g$ $\langle \pi_1 \circ h, \pi_2 \circ h\rangle = h$, $h : C \to A \times B$
Exponentials B^A	$ev_{A,B} : B^A \times A \to B$ $\dfrac{C \times A \xrightarrow{f} B}{C \xrightarrow{f^*} B^A}$	$ev \circ \langle f^* \circ \pi_1, \pi_2\rangle = f$ $(ev \circ \langle g \circ \pi_1, \pi_2\rangle)^* = g$, $g : C \to B^A$

Fig. 1. CCC's equationally.

EXAMPLE 2.3. The category **Set** of sets and functions. Here $A \times B$ is a chosen Cartesian product and B^A is the set of functions from A to B. The map $B^A \times A \xrightarrow{ev} B$ is the usual evaluation map, while currying $C \xrightarrow{f^*} B^A$ is the map $c \mapsto (a \mapsto f(c, a))$.

An important subfamily of examples are *Henkin models* which are ccc's in which the terminal object **1** is a generator ([Mit96], Theorem 7.2.41). More concretely, for a lambda calculus signature with freely generated types (cf. Section 2.13 below), a *Henkin model* $\mathcal{A}$ is a type-indexed family of sets $\mathcal{A} = \{A_\sigma \mid \sigma \text{ a type}\}$ where $A_1 = \{*\}$, $A_{\sigma\times\tau} = A_\sigma \times A_\tau$, $A_{\sigma\Rightarrow\tau} \subseteq A_\tau^{A_\sigma}$ which forms a ccc with respect to restriction of the usual ccc structure of **Set**. In the case of atomic base sorts b, $\mathcal{A}_b$ is some fixed but arbitrary set. A *full type hierarchy* is a Henkin model with full function spaces, i.e. $A_{\sigma\Rightarrow\tau} = A_\tau^{A_\sigma}$.

EXAMPLE 2.4. More generally, the functor category $\mathbf{Set}^{\mathcal{C}^{op}}$ of presheaves on $\mathcal{C}$ is Cartesian closed. Its objects are (contravariant) functors from $\mathcal{C}$ to **Set**, and its arrows are natural transformations between them. We sketch the ccc structure: given $F, G \in \mathbf{Set}^{\mathcal{C}^{op}}$, define $F \times G$ pointwise on objects and arrows. Motivated by Yoneda's Lemma, define $G^F(A) = Nat(h^A \times F, G)$, where $h^A = Hom(A, -)$. This easily extends to a functor. Finally if $H \times F \xrightarrow{\theta} G$, define $H \xrightarrow{\theta^*} G^F$ by: $\theta^*_A(a)_C(h, c) = \theta_C(H(h)(a), c)$.

Functor categories have been used in studying problematic semantical issues in Algol-like languages [Rey81,Ol85,OHT92,Ten94], as well as recently in concurrency theory and models of π-calculus [CSW,CaWi]. Special cases of presheaves have been studied extensively [Mit96,LS86]:

- Let $\mathcal{C}$ be a poset (qua trivial category). Then $\mathbf{Set}^{\mathcal{C}^{op}}$, the category of Kripke models over $\mathcal{C}$, may be identified with sets indexed (or graded) by the poset $\mathcal{C}$. Such models are fundamental in intuitionistic logic [LS86,TrvD88] and also arise in Kripke Logical Relations, an important tool in semantics of programming languages [Mit96,OHT93, OHRi].
- Let $\mathcal{C} = \mathcal{O}(X)$, the poset of opens of the topological space X. The subcategory $Sh(X)$ of sheaves on X is Cartesian closed.
- Let $\mathcal{C}$ be a monoid M (qua category with one object). Then $\mathbf{Set}^{\mathcal{C}^{op}}$ is the category of M-*sets*, i.e. sets X equipped with a left action; equivalently, a monoid homomorphism $M \to End(X)$, where $End(X)$ is the monoid of endomaps of X. Morphisms of M-sets X and Y are equivariant maps (i.e. functions commuting with the action.) A special case of this is when M is actually a group G (qua category with one object, where all maps are isos). In that case $\mathbf{Set}^{\mathcal{C}^{op}}$ is the category of G-sets, the category of permutational representations of G. Its objects are sets X equipped with left actions $G \to Sym(X)$ and whose morphisms are equivariant maps. We shall return to these examples when we speak of Laüchli semantics and Full Completeness, Section 5.2

EXAMPLE 2.5. ω-**CPO**. Objects are posets P such that countable ascending chains $a_0 \leqslant a_1 \leqslant a_2 \leqslant \cdots$ have suprema. Morphisms are maps which preserve suprema of countable ascending chains (in particular, are order preserving). This category is a ccc, with products $P \times Q$ ordered pointwise and $Q^P = Hom(P, Q)$, ordered pointwise. In this case, the categories are ω-**CPO**-enriched – i.e. the hom-sets themselves form an ω-**CPO**, compatible with composition. An important subccc is ω-**CPO**$_\perp$, in which all objects have a distinguished minimal element $\perp$ (but morphisms need not preserve it).

The category ω-**CPO** is the most basic example in a vast research area, *domain theory*, which has arisen since 1970. This area concerns the denotational semantics of programming languages and models of untyped lambda calculi (cf. Section 2.5 below). See also the survey article [AbJu94].

EXAMPLE 2.6. *Coherent spaces and stable maps.* A *Coherent Space* $\mathcal{A}$ is a family of sets satisfying: (i) $a \in \mathcal{A}$ and $b \subseteq a$ implies $b \in \mathcal{A}$, and (ii) if $B \subseteq \mathcal{A}$ and if $\forall c, c' \in B(c \cup c' \in \mathcal{A})$ then $\bigcup B \in \mathcal{A}$. In particular, $\emptyset \in \mathcal{A}$. Morphisms are *stable maps*, i.e. monotone maps preserving pullbacks and filtered colimits. That is, $f : \mathcal{A} \to \mathcal{B}$ is a stable map if

(i) $b \subseteq a \in \mathcal{A}$ implies $f(b) \subseteq f(a)$,
(ii) $f(\bigcup_{i\in I} a_i) = \bigcup_{i\in I} f(a_i)$, for I directed, and
(iii) $a \cup b \in \mathcal{A}$ implies $f(a \cap b) = f(a) \cap f(b)$.

This gives a category **Stab**. Every coherent space $\mathcal{A}$ yields a reflexive-symmetric (undirected) graph $(|\mathcal{A}|, \frown\!\!\!\!\smile)$ where $|\mathcal{A}| = \{a \mid \{a\} \in \mathcal{A}\}$ and $a \frown\!\!\!\!\smile b$ iff $\{a, b\} \in \mathcal{A}$. Moreover, there is a bijective correspondence between such graphs and coherent spaces. Given two coherent spaces $\mathcal{A}, \mathcal{B}$ their product $\mathcal{A} \times \mathcal{B}$ is defined via the associated graphs as follows: $(|\mathcal{A} \times \mathcal{B}|, \frown\!\!\!\!\smile_{\mathcal{A}\times\mathcal{B}})$, with $|\mathcal{A} \times \mathcal{B}| = |\mathcal{A}| \uplus |\mathcal{B}| = (\{1\} \times |\mathcal{A}|) \cup (\{2\} \times |\mathcal{B}|)$ where $(1, a) \frown\!\!\!\!\smile_{\mathcal{A}\times\mathcal{B}} (1, a')$ iff $a \frown\!\!\!\!\smile_{\mathcal{A}} a'$, $(2, b) \frown\!\!\!\!\smile_{\mathcal{A}\times\mathcal{B}} (2, b')$ iff $b \frown\!\!\!\!\smile_{\mathcal{B}} b'$, and $(1, a) \frown\!\!\!\!\smile_{\mathcal{A}\times\mathcal{B}} (2, b)$ for all $a \in |\mathcal{A}|$, $b \in |\mathcal{B}|$. The function space $\mathcal{B}^{\mathcal{A}} = Stab(\mathcal{A}, \mathcal{B})$ of stable maps can be given the structure of a coherent space, ordered by Berry's order: $f \preceq g$ iff for all $a, a' \in \mathcal{A}$, $a' \subseteq a$ implies $f(a') = f(a) \cap g(a')$. For details, see [GLT,Tr92]. This class of domains led to the discovery of linear logic (Section 4.2).

EXAMPLE 2.7. *Per models.* A *partial equivalence relation* (per) is a symmetric, transitive relation $\sim_A \subseteq A^2$. Thus $\sim_A$ is an equivalence relation on the subset $Dom_A = \{x \in A \mid x \sim_A x\}$. A $\mathcal{P}$-set is a pair $(A, \sim_A)$ where A is a set and $\sim_A$ is a per on A. Given two $\mathcal{P}$-sets $(A, \sim_A)$ and $(B, \sim_B)$ a *morphism of* $\mathcal{P}$*-sets* is a function $f : A \to B$ such that $a \sim_A a'$ implies $f(a) \sim_B f(a')$ for all $a, a' \in A$. That is, f induces a map of quotients $Dom_A / \sim_A \to Dom_B / \sim_B$ which preserves the associated partitions.

$\mathcal{P}$**Set**, the category of $\mathcal{P}$-sets and morphisms is a ccc, with structure induced from **Set**: we define $(A \times B, \sim_{A\times B})$, where $(a, b) \sim_{A\times B} (a', b')$ iff $a \sim_A a'$ and $b \sim_B b'$ and $(B^A, \sim_{B^A})$, where $f \sim_{B^A} g$ iff for all $a, a' \in A$, $a \sim_A a'$ implies $f(a) \sim_B g(a')$. We shall discuss variants of the ccc structure of $\mathcal{P}$**Set** in Section 2.7 below, with respect to reduction-free normalization.

Other classes of *Per* models are obtained by considering pers on a fixed (functionally complete) partial combinatory algebra, for example built over a model of untyped lambda calculus (cf. Section 2.5 below). The prototypical example is the following category *Per*(**N**) of pers on the natural numbers. The objects are pers on **N**. Morphisms $R \xrightarrow{f} S$ are (equivalence classes of) partial recursive functions (= Turing-machine computable partial functions) $\mathbf{N} \rightharpoonup \mathbf{N}$ which induce a total map on the induced partitions, i.e. for all $m, n \in \mathbf{N}$, mRn implies $f(m), f(n)$ are defined and $f(m)Sf(n)$. Here we define equivalence of maps $f, g : R \to S$ by: $f \sim g$ iff $\forall m, n$, mRn implies $f(m), g(n)$ are defined and $f(m)Sg(n)$. The fact that *Per*(**N**) is a ccc uses some elementary recursion theory [BFSS90,Mit96,AL91]. (See also Section 2.4.1.)

EXAMPLE 2.8. *Free CCC's.* Given a set of basic objects $\mathcal{X}$, we can form $\mathcal{F}_{\mathcal{X}}$, the *free ccc generated by* $\mathcal{X}$. Its objects are freely generated from $\mathcal{X}$ and $\mathbf{1}$ using $\times$ and $(-)^{(-)}$, its arrows are freely generated using identities and composition plus the structure in Figure 1, and we impose the minimal equations required to have a ccc. More generally, we may build $\mathcal{F}_{\mathcal{G}}$, the free ccc generated by a directed multigraph (or even a small category) $\mathcal{G}$, by freely generating from the vertices (resp. objects) and edges (resp. arrows) of $\mathcal{G}$ and then – in the case of categories $\mathcal{G}$ – imposing the appropriate equations. The sense that this is free is related to Definition 2.9 and discussed in Example 2.23.

Cartesian closed categories can themselves be made into a category in many ways. This depends, to some extent, on how much 2- , bi-, enriched- , etc. structure one wishes to impose. The following elementary notions have proved useful. We shall mention a comparison between strict and nonstrict ccc's with coproducts in Remark 2.28. More general notions of monoidal functors, etc. will be mentioned in Section 4.1.

DEFINITION 2.9. $\mathbf{CART}_{st}$ is the category of strictly structured Cartesian closed categories with functors that preserve the structure on the nose. **2-CART**$_{st}$ is the 2-category whose 0-cells are Cartesian closed categories, whose 1-cells are strict Cartesian closed functors, and whose 2-cells are natural isomorphisms [Cu93].

As pointed out by Lambek [L74,LS86], given a ccc $\mathcal{A}$, we may adjoin an indeterminate arrow $\mathbf{1} \xrightarrow{x} A$ to $\mathcal{A}$ to form a *polynomial* Cartesian closed category $\mathcal{A}[x]$ over $\mathcal{A}$, with the expected universal property in $\mathbf{CART}_{st}$. The objects of $\mathcal{A}[x]$ are the same as those of $\mathcal{A}$, while the arrows are "polynomials", i.e. formal expressions built from the symbol x using the arrow-forming operations of $\mathcal{A}$. The key fact about such polynomial expressions is a normal form theorem, stated here for ccc's, although it applies more generally (see [LS86], p. 61):

PROPOSITION 2.10 (Functional completeness). *For every polynomial $\varphi(x)$ in an indeterminate $\mathbf{1} \xrightarrow{x} A$ over a ccc $\mathcal{A}$, there is a unique arrow $\mathbf{1} \xrightarrow{h} C^A \in \mathcal{A}$ such that $ev \circ \langle h, x\rangle \underset{x}{=} \varphi(x)$, where $\underset{x}{=}$ is equality in $\mathcal{A}[x]$.*

Looking ahead to lambda calculus notation in the next section, we write $h \equiv \lambda_{x:A}.\varphi(x)$, so the equation above becomes $ev \circ \langle \lambda_{x:A}.\varphi(x), x\rangle \underset{x}{=} \varphi(x)$. The universal property of polynomial algebras guarantees a notion of *substitution of constants* $\mathbf{1} \xrightarrow{a} A \in \mathcal{A}$ for indeterminates x in $\varphi(x)$. We obtain the following:

COROLLARY 2.11 (The β rule). *In the situation above, for any arrow $\mathbf{1} \xrightarrow{a} A \in \mathcal{A}$*

$$ev \circ \langle \lambda_{x:A} \cdot \varphi(x), a\rangle = \varphi(a) \tag{4}$$

holds in $\mathcal{A}$.

The β-rule is the foundation of the *lambda calculus*, fundamental in programming language theory. It says the following: we think of $\lambda_{x:A}.\varphi(x)$ as the function $x \mapsto \varphi(x)$. Equation 4 says: evaluating the function $\lambda_{x:A}.\varphi(x)$ at argument a is just substitution of the constant a for each occurrence of x in $\varphi(x)$. However this process is far more sophisticated than simple polynomial substitution in algebra. In our situation, the argument a may itself be a lambda term, which in turn may contain other lambda terms applied to various arguments, etc. After substitution, the right hand side $\varphi(a)$ of Eq. (4) may be far more complex than the left hand side, with many new possibilities for evaluations created by the substitution. Thus, if we think of *computation* as oriented rewriting from the LHS to the RHS, it is not at all obvious the process ever halts. The fact that it does is a basic theorem in the so-called Operational Semantics of typed lambda calculus. Indeed, the Strong Normalization Theorem (cf. [LS86], p. 81) says *every* sequence of ordered rewrites (from left to right) eventually halts at an irreducible term (cf. Remark 2.49 and Section 2.7 below).

REMARK 2.12. We may also form polynomial ccc's $\mathcal{A}[x_1, \ldots, x_n]$ by adjoining a finite set of indeterminates $\mathbf{1} \xrightarrow{x_i} A_i$. Using product types, one may show $\mathcal{A}[x_1, \ldots, x_n] \cong \mathcal{A}[z]$, for an indeterminate $\mathbf{1} \xrightarrow{z} A_1 \times \cdots \times A_n$.

Polynomial Cartesian or Cartesian closed categories $\mathcal{A}[x]$ may be constructed directly, showing they are the Kleisli category of an appropriate comonad on $\mathcal{A}$ (see [LS86], p. 56). Extensions of this technique to allow adjoining indeterminates to fibrations, using 2-categorical machinery are considered in [HJ95].

2.2. *Simply typed lambda calculi*

Lambda Calculus is an abstract theory of functions developed by Alonzo Church in the 1930's. Originally arising in the foundations of logic and computability theory, more recently it has become an essential tool in the mathematical foundations of programming languages [Mit96]. The calculus itself, to be described below, encompasses the process of building functions from variables and constants, using application and functional abstraction.

Actually, there are many "lambda calculi" – typed and untyped – with various elaborate structures of types, terms, and equations. Let us give the basic typed one. We shall follow an algebraic syntax as in [LS86].

DEFINITION 2.13 (*Typed λ-calculus*). Let Sorts be a set of sorts (or atomic types). The *typed λ-calculus generated by* Sorts is a formal system consisting of three classes: Types, Terms and Equations between terms. We write $a : A$ for "a is a term of type A".

Types: This is the set obtained from the set of Sorts using the following rules: Sorts are types, $\mathbf{1}$ is a type, and if A and B are types then so are $A \times B$ and B^A. We allow the possibility of other types or type-forming operations and possible identifications between types. (Set theorists may even use "classes" instead of "sets".)

Terms: To every type A we assign a denumerable set of typed variables $x_i^A : A$, $i = 0, 1, 2, \ldots$. We write $x : A$ or x^A for a typical variable x of type A. Terms are freely generated from variables, constants, and term-forming operations. We require at least the following distinguished generators:

(1) $* : \mathbf{1}$,
(2) If $a : A,\ b : B,\ c : A \times B$, then $\langle a, b\rangle : A \times B,\ \pi_1^{A,B}(c) : A,\ \pi_2^{A,B}(c) : B$,
(3) If $a : A,\ f : B^A,\ \varphi : B$ then $ev_{A,B}(f, a) : B,\ \lambda_{x:A}.\varphi : B^A$.

There may be additional constants and term-forming operations besides those specified.

We shall abbreviate $ev_{A,B}(f, a)$ by $f`a$, read "f *of* a", omitting types when clear. Intuitively, $ev_{A,B}$ denotes evaluation, $\langle -, -\rangle$ denotes pairing, and $\lambda_{x:A}.\varphi$ denotes the function $x \mapsto \varphi$, where φ is some term expression possibly containing x. The operator $\lambda_{x:A}$ acts like a quantifier, so the variable x in $\lambda_{x:A} \cdot \varphi$ is a bound (or dummy) variable, just like the x in $\forall_{x:A}\varphi$ or in $\int f(x)\,dx$. We inductively define the sets of free and bound variables in a term t, denoted $FV(t), BV(t)$, resp. (cf. [Bar84], p. 24). We shall always identify terms up to renaming of bound variables. The expression $\varphi[a/x]$ denotes the result of substituting the term $a : A$ for each occurrence of $x : A$ in φ, if necessary renaming bound variables in φ so that no clashes occur (cf. [Bar84]). Terms without free variables are called *closed*; otherwise, *open*.

Equations between terms: A *context* Γ is a finite set of (typed) variables. An *equation in context* Γ is an expression $a \underset{\Gamma}{=} a'$, where a, a' are terms of the same type A whose free variables are contained in Γ.

The equality relation between terms (in context) of the same type is generated using (at least) the following axioms and closure under the following rules:

(i) $\underset{\Gamma}{=}$ is an equivalence relation.

(ii) $\dfrac{a \underset{\Gamma}{=} b}{a \underset{\Delta}{=} b}$, whenever $\Gamma \subseteq \Delta$.

(iii) $\underset{\Gamma}{=}$ must be a congruence relation with respect to all term-forming operations. It suffices to consider closure under the following two rules (cf. [LS86])

$$\frac{a \underset{\Gamma}{=} b}{f`a \underset{\Gamma}{=} f`b} \qquad \frac{\varphi \underset{\Gamma\cup\{x^A\}}{=} \varphi'}{\lambda_{x:A} \cdot \varphi \underset{\Gamma}{=} \lambda_{x:A} \cdot \varphi'}.$$

(iv) The following specific axioms (we omit subscripts on terms, when the types are obvious):

Products

(a) $a \underset{\Gamma}{=} *$ for all $a : \mathbf{1}$,
(b) $\pi_1(\langle a, b\rangle) \underset{\Gamma}{=} a$ for all $a : A,\ b : B$,
(c) $\pi_2(\langle a, b\rangle) \underset{\Gamma}{=} b$ for all $a : A,\ b : B$,
(d) $\langle \pi_1(c), \pi_2(c)\rangle \underset{\Gamma}{=} c$ for all $c : C$,

Lambda Calculus
β-Rule $(\lambda_{x:A} \cdot \varphi)`a \underset{\Gamma}{=} \varphi[a/x]$,
η-Rule $\lambda_{x:A} \cdot (f`x) \underset{\Gamma}{=} f$, where $f : B^A$ and x is not a free variable of f.

REMARK 2.14. There may be additional types, terms, or equations. Following standard conventions, we equate terms which only differ by change of bound variables – this is called α-conversion in the literature [Bar84]. Equations are *in context* – i.e. occur within a declared set of free variables. This allows the possibility of *empty types*, i.e. types without closed terms (of that type). This view is fundamental in recent approaches to functional languages [Mit96] and necessary for interpreting such theories in presheaf categories, for example. However, if there happen to be closed terms $a : A$ of each type, we may omit the subscript Γ on equations, because of the following derivable rule (cf. [LS86], Prop. 10.1, p. 75): for $x \notin \Gamma$ and if all free variables of a are contained in Γ,

$$\frac{\varphi(x) \underset{\Gamma\cup\{x\}}{=} \psi(x)}{\varphi[a/x] \underset{\Gamma}{=} \psi[a/x]}.$$

EXAMPLE 2.15. *Freely generated simply typed lambda calculi.* These are freely generated from specified sorts, terms, and/or equations. In the minimal case (no additional assumptions) we obtain the simply typed lambda calculus with finite products freely generated by Sorts. Typically, however, we assume that among the Sorts are distinguished *datatypes* and associated terms, possibly with specified equations. For example, basic universal algebra would be modelled by sorts A with distinguished n-ary operations given by terms $t : A^n \Rightarrow A$ and constants $c : \mathbf{1} \to A$. Any specified term equations are added to the theory as (nonlogical) axioms.

EXAMPLE 2.16. *The internal language of a ccc $\mathcal{A}$.* Here the types are the objects of $\mathcal{A}$, where $\times, (-)^{(-)}, \mathbf{1}$ have the obvious meanings. Terms with free variables $x_1 : A_1, \ldots, x_n : A_n$ are polynomials in $\mathcal{A}[x_1, \ldots, x_n]$, where $\mathbf{1} \xrightarrow{x_i} A_i$ is an indeterminate, lambda abstraction is given by functional completeness, as in Proposition 2.10, and we define $a \underset{X}{=} b$ to hold iff $a \underset{X}{=} b$ as polynomials in $\mathcal{A}[X]$, where $X = \{x_1, \ldots, x_n\}$.

REMARK 2.17.

(i) Historically, typed lambda calculi were often presented with *only* exponential types B^A (no products) and the associated machinery [Bar84,Bar92]. This permits certain simplifications in inductive arguments, athough it is categorically less "natural" (cf. also Remark 2.24).

(ii) It is a fundamental property that lambda calculus is a *higher-order* functional language: terms of type B^A can use an arbitrary term of type A as an argument, and A and B themselves may be very complex. Thus, typed lambda calculus is often referred to as a theory of *functionals of higher type*.

2.3. *Formulas-as-types: The fundamental equivalence*

Let us describe the third component of the trio: Cartesian closed categories, typed lambda calculi, and formulas-as-types. The *Formulas-as-Types* view, sometimes called the Curry–Howard isomorphism, is playing an increasingly influential role in the logical foundations of computing, especially in the foundations of functional programming languages. Its historical roots lie in the so-called Brouwer–Heyting–Kolmogorov (BHK) interpretation of intuitionistic logic from the 1920's [GLT,TrvD88]. The idea is based on modelling proofs (which are programs) by functions, i.e. lambda terms. Since proofs can be modelled by lambda terms and the latter are themselves arrows in certain free categories, it follows that functional programs can be modelled categorically.

In modern guise, the Curry–Howard analysis says the following. Proofs in a constructive logic $\mathcal{L}$ may be identified as terms of an appropriate typed lambda calculus $\lambda_{\mathcal{L}}$, where:

- types = formulas of $\mathcal{L}$,
- lambda terms = proofs (i.e. annotations of Natural Deduction proof trees),
- provable equality of lambda terms corresponds to the equivalence relation on proofs generated by Gentzen's normalization algorithm.

Often researchers impose additional equations between lambda terms, motivated from categorical considerations (e.g., to force traditional datatypes to have a strong universal mapping property).

REMARK 2.18 (*formulas = specifications*). More generally, the Curry–Howard view identifies types of a programming language with formulas of some logic, and programs of type A as proofs within the logic of formula A. Constructing proofs of formula A may then be interpreted as building programs that meet the specification A.

For example, consider the intuitionistic $\{\top, \wedge, \Rightarrow\}$-fragment of propositional calculus, as in Figure 2. This logic closely follows the presentation of ccc's in Definition 2.1 and Figure 1. We now identify (= Formulas-as-Types) the propositional symbols $\top, \wedge, \Rightarrow$ with the type constructors $\mathbf{1}, \times, \Rightarrow$, respectively. We assign lambda terms inductively. To a proof

Formulas	$A ::= \top \mid Atoms \mid A_1 \wedge A_2 \mid A_1 \Rightarrow A_2$
Provability	$\vdash$ is a reflexive, transitive relation such that, for arbitrary formulas A, B, C
	$A \vdash \top, \quad A \wedge B \vdash A, \quad A \wedge B \vdash B$
	$C \vdash A \wedge B$ iff $C \vdash A$ and $C \vdash B$
	$C \wedge A \vdash B$ iff $C \vdash A \Rightarrow B$

Fig. 2. Intuitionistic $\top, \wedge, \Rightarrow$ logic.

of $A \vdash B$ we assign λ-terms $x : A \vdash t(x) : B$, where $t(x)$ is a term of type B with at most the free variable $x : A$ (i.e. in context $\{x : A\}$) as follows:

$$x : A \vdash x : A, \quad \frac{x : A \vdash s(x) : B \;\; y : B \vdash t(y) : C}{x : A \vdash t[s(x)/y] : C},$$

$$x : A \vdash * : \top, \quad x : A \wedge B \vdash \pi_1(x) : A, \quad x : A \wedge B \vdash \pi_2(x) : B,$$

$$\frac{x : C \vdash a : A \;\; x : C \vdash b : B}{x : C \vdash \langle a, b\rangle : A \wedge B}, \quad \frac{z : C \wedge A \vdash t(z) : B}{y : C \vdash \lambda_{x:A} \cdot t[\langle y, x\rangle/z] : A \Rightarrow B},$$

$$\frac{y : C \vdash t(y) : A \Rightarrow B}{z : C \wedge A \vdash t[\pi_1(z)/y]`\pi_2(z) : B}.$$

We can now refer to entire proof trees by the associated lambda terms. We wish to put an equivalence relation on proofs, according to the equations of typed lambda calculus. Given two proofs of an entailment $A \vdash B$, say $x : A \vdash s(x) : B$ and $x : A \vdash t(x) : B$, we say they are *equivalent* if we can derive $s \underset{\{x\}}{=} t$ in the appropriate typed lambda calculus.

DEFINITION 2.19. Let λ-**Calc** denote the category whose objects are typed lambda calculi and whose morphisms are *translations*, i.e. maps Φ which send types to types, terms to terms (including mapping the ith variable of type A to the ith variable of type $\Phi(A)$), preserve all the specified operations on types and terms on the nose, and preserve equations.

THEOREM 2.20. *There are a pair of functors* $\mathbf{C} : \lambda$-**Calc** $\to$ $\mathbf{Cart}_{st}$ *and* $\mathbf{L} : \mathbf{Cart}_{st} \to$ λ-**Calc** *which set up an equivalence of categories* $\mathbf{Cart}_{st} \cong$ λ-**Calc**.

The functor L associates to ccc $\mathcal{A}$ its internal language, while the functor $\mathbf{C}$ associates to any lambda calculus $\mathcal{L}$, a syntactically generated ccc $\mathbf{C}(\mathcal{L})$, whose objects are types of $\mathcal{L}$ and whose arrows $A \to B$ are denoted by (equivalence classes of) lambda terms $t(x)$ representing proofs $x : A \vdash t(x) : B$ as above (see [LS86]).

This leads to a kind of Soundness Theorem for diagrammatic reasoning which is important in categorical logic.

COROLLARY 2.21. *Verifying that a diagram commutes in a ccc* $\mathcal{C}$ *is equivalent to proving an equation in the internal language of* $\mathcal{C}$.

The above result includes allowing algebraic theories modelled in the Cartesian fragment [Mac82,Cr93], as well as extensions with categorical data types (like weak natural numbers objects, see Section 2.3.1). Theorem 2.20 also leads to concrete syntactic presentations of free ccc's [LS86,Tay98]. Let **Graph** be the category of directed multi-graphs [ST96].

COROLLARY 2.22. *The forgetful functor* $\mathcal{U} : \mathbf{Cart}_{st} \to$ **Graph** *has a left adjoint* $\mathcal{F} : \mathbf{Graph} \to \mathbf{Cart}_{st}$. *Let* $\mathcal{F}_\mathcal{G}$ *denote the image of graph* $\mathcal{G}$ *under* $\mathcal{F}$. *We call* $\mathcal{F}_\mathcal{G}$ *the free ccc generated by* $\mathcal{G}$.

EXAMPLE 2.23. Given a discrete graph $\mathcal{G}_0$ considered as a a set, $\mathcal{F}_{\mathcal{G}_0}$ = the free ccc generated by the set of sorts $\mathcal{G}_0$. It has the following universal property: for any ccc $\mathcal{C}$ and for any graph morphism $F : \mathcal{G}_0 \to \mathcal{C}$, there is a unique extension to a (strict) ccc-functor $[\![-]\!]_F : \mathcal{F}_{\mathcal{G}_0} \to \mathcal{C}$.

$$
\begin{tikzcd}
\mathcal{F}_{\mathcal{G}_0} \arrow[r, "{[\![-]\!]_F}"] & \mathcal{C} \\
\mathcal{G}_0 \arrow[u] \arrow[ur, "F"'] &
\end{tikzcd}
$$

This says: given any interpretation F of basic atomic types (= nodes of $\mathcal{G}_0$) as objects of $\mathcal{C}$, there is a unique extension to an interpretation $[\![-]\!]_F$ in $\mathcal{C}$ of the entire simply typed lambda calculus generated by $\mathcal{G}_0$ (identifying the free ccc $\mathcal{F}_{\mathcal{G}_0}$ with this lambda calculus).

REMARK 2.24. A Pitts [Pi9?] has shown how to construct free ccc's syntactically, using lambda calculi without product types. The idea is to take objects to be *sequences* of types and arrows to be sequences of terms. The terminal object is the empty sequence, while products are given by concatenation of sequences. For a full discussion, see [CDS97]. This is useful in reduction-free normalization (see Section 2.7 below).

REMARK 2.25. There are more advanced 2- and bi-categorical versions of the above results. We shall mention more structure in the case of Cartesian closed categories with coproducts, in the next section.

2.3.1. *Some datatypes.* Computing requires datatypes, for example natural numbers, lists, arrays, etc. The categorical development of such datatypes is an old and established area. The reader is referred to any of the standard texts for discussion of the basics, e.g., [MA86, BW95,Mit96,Ten94]. General categorical treatments of abstract datatypes abound in the literature. The standard treatment is to use initial T-algebras (cf. Section 2.4.2 below) or final T-coalgebras for "definable" or "polynomial" endofunctors T. There are interesting common generalizations to lambda calculi with functorial type constructors [Ha87, Wr89], categories with datatypes determined by strong monads [Mo91,CSp91], and using enriched categorical structures [K82]. There is recent discussion of datatypes in distributive categories [Co93,W92], and the use of the categorical theory of sketches [BW95,Bor94].

We shall merely illustrate a few elementary algebraic structures commonly added to a Cartesian or Cartesian closed category (or the associated term calculi).

DEFINITION 2.26. A category $\mathcal{C}$ *has finite coproducts* (equivalently, binary coproducts and an initial object $\mathbf{0}$) if for every $A, B \in \mathcal{C}$ there is a distinguished object $A + B$, together with isomorphisms (natural in $A, B, C \in \mathcal{C}$)

$$Hom_{\mathcal{C}}(\mathbf{0}, A) \cong \{*\}, \tag{5}$$

$$Hom_{\mathcal{C}}(A + B, C) \cong Hom_{\mathcal{C}}(A, C) \times Hom_{\mathcal{C}}(B, C). \tag{6}$$

Objects	Distinguished Arrow(s)	Equations
Initial $\mathbf{0}$	$\mathbf{0} \xrightarrow{O_A} A$	$O_A = f$, $f : \mathbf{0} \to A$
Coproducts $A + B$	$in_1^{A,B} : A \to A + B$ $in_2^{A,B} : B \to A + B$ $\dfrac{A \xrightarrow{f} C \quad B \xrightarrow{g} C}{A+B \xrightarrow{[f,g]} C}$	$[f, g] \circ in_1 = f$ $[f, g] \circ in_2 = g$ $[h \circ in_1, h \circ in_2] = h$, $h : A + B \to C$

Fig. 3. Coproducts.

We say C is *bi-Cartesian closed* (= biccc) if it is a ccc with finite coproducts.[1]

Just as in the case of products (cf. Figure 1), we may present coproducts equationally, as in Figure 3, and speak of *strict* structure, etc. In programming language semantics, coproducts correspond to *variant types*, set-theoretically they correspond to disjoint union, while from the logical viewpoint coproducts correspond to disjunction. Thus a biccc corresponds to intuitionistic $\{\perp, \top, \wedge, \vee, \Rightarrow\}$-logic. We add to the logic of Figure 2 formulas $\perp$ and $A_1 \vee A_2$, together with the rules

$$\begin{array}{ccc} & \perp \vdash A & \\ A \vee B \vdash C & \text{iff} & A \vdash C \text{ and } B \vdash C \end{array}$$

corresponding to Eqs. (5), (6). The associated typed lambda calculus with coproducts is rather subtle to formulate [Mit96,GLT]. The problem is with the copairing operator

$$A + B \xrightarrow{[f,g]} C$$

which in **Sets** corresponds to a *definition-by-cases* operator:

$$[f, g](x) = \begin{cases} f(x) & \text{if } x \in A, \\ g(x) & \text{if } x \in B. \end{cases}$$

The correct lambda calculus formalism for coproduct types corresponds to the logicians' natural deduction rules for strong sums. The issue is not trivial, since the word problem for free biccc's (and the associated type isomorphism problem [DiCo95]) is among the most difficult of this type of question, and – at least for the current state of the art – depends heavily on technical subtleties of syntax for its solution (see [Gh96]).

Just as for ccc's, we may introduce various 2-categories of biccc's (cf. [Cu93]). For example

[1] Not to be confused with *bicategories*, cf. [Bor94].

DEFINITION 2.27. The 2-category $\mathbf{2\text{-}BiCART}_{st}$ has 0-cells strict bi-Cartesian closed categories, 1-cells functors preserving the structure on the nose, and 2-cells natural isomorphisms.

One may similarly define a non-strict version **2-BiCART**.

REMARK 2.28. Every bi-Cartesian closed category is equivalent to a strict one. Indeed, this is part of a general 2-categorical adjointness between the above 2-categories, from a theorem of Blackwell, Kelly, and Power. (See Čubrić [Cu93] for applications to lambda calculi.)

DEFINITION 2.29. In a biccc, define $\mathbf{Boole} = \mathbf{1} + \mathbf{1}$, the type of Booleans.

Boole's most salient feature is that it has two distinguished global elements (Boolean values) $\mathbf{T}, \mathbf{F} : \mathbf{1} \to \mathbf{Boole}$, corresponding to the two injections in_1, in_2, together with the universal property of coproducts. In **Set** we interpret **Boole** as a set of cardinality 2; similarly, in typed lambda calculus, it corresponds to a type with two distinguished constants $\mathbf{T}, \mathbf{F} : \mathbf{Boole}$ and an appropriate notion of definition by cases. In any biccc, we can define all of the classical n-ary propositional logic connectives as arrows $\mathbf{Boole}^n \to \mathbf{Boole}$ (see [LS86], I.8). A weaker notion of Booleans in the category ω-CPO$_\perp$ is illustrated in Figure 4.

DEFINITION 2.30. A *natural numbers object* in a ccc $\mathcal{C}$ is an object $\mathbf{N}$ with arrows $\mathbf{1} \xrightarrow{0} \mathbf{N} \xrightarrow{S} \mathbf{N}$ which is initial among diagrams of that shape. That is, for any object A and arrows $\mathbf{1} \xrightarrow{a} A \xrightarrow{h} A$, there is a unique *iterator* $\mathcal{I}_{ah} : \mathbf{N} \to A$ making the following diagram commute:

$$\begin{tikzcd}
\mathbf{1} \arrow[r, "0"] \arrow[dr, "a"'] & \mathbf{N} \arrow[r, "S"] \arrow[d, "\mathcal{I}_{ah}"] & \mathbf{N} \arrow[d, "\mathcal{I}_{ah}"] \\
 & A \arrow[r, "h"] & A
\end{tikzcd}$$

A *weak natural numbers object* is defined as above, but just assuming existence and not necessarily uniqueness of $\mathcal{I}_{ah}$.

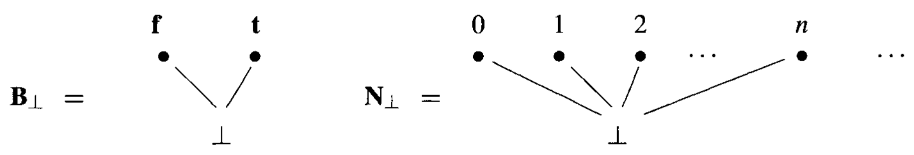

Fig. 4. Flat datatypes in ω-CPO$_\perp$.

In the category **Set**, the natural numbers $(\mathbf{N}, 0, S)$ is a natural numbers object, where $Sn = n + 1$.

In functor categories $\mathbf{Set}^{\mathcal{C}}$, a natural numbers object is given by the constant functor $K_{\mathbf{N}}$, where $K_{\mathbf{N}}(A) = \mathbf{N}$, and $K_{\mathbf{N}}(f) = id_{\mathbf{N}}$, with obvious natural transformations $\mathbf{1} \xrightarrow{0} K_{\mathbf{N}} \xrightarrow{S} K_{\mathbf{N}}$. In ω-**CPO** there are numerous *weak* natural numbers objects: for example the *flat pointed* natural numbers $\mathbf{N}_\perp = \mathbf{N} \uplus \{\perp\}$, ordered as follows: $a \leqslant b$ iff $a = b$ or $a = \perp$, where $S(n) = n + 1$ and $S(\perp) = \perp$, pictured in Figure 4.

Natural numbers objects – when they exist – are unique up to isomorphism; however weak ones are far from unique. Typical programming languages and typed lambda calculi in logic assume only weak natural numbers objects.

If a ccc $\mathcal{C}$ has a natural numbers object $\mathbf{N}$, we can construct parametrized versions of iteration, using products and exponentiation in $\mathcal{C}$ [LS86,FrSc]. For example, in **Set**: given functions $g : A \to B$ and $f : \mathbf{N} \times A \times B \to B$, there exists a unique *primitive recursor* $\mathcal{R}_{gf} : \mathbf{N} \times A \to B$ satisfying: (i) $\mathcal{R}_{gf}(0, a) = g(a)$ and (ii) $\mathcal{R}_{gf}(Sn, a) = f(n, a, \mathcal{R}_{gf}(n, a))$. These equations are easily represented in any ccc with $\mathbf{N}$, or in the associated typed lambda calculus (e.g., the number $n \in \mathbf{N}$ being identified with $S^n 0$). In the case $\mathcal{C}$ has only a weak natural numbers object, we may prove the existence but not necessarily the uniqueness of $\mathcal{R}_{gf}$.

An important datatype in Computer Science is the type of finite lists of elements of some type A. This is defined analogously to (weak) natural numbers objects:

DEFINITION 2.31. Given an object A in a ccc $\mathcal{C}$, we define the object $\ell ist(A)$ of *finite lists on* A with the following distinguished structure: arrows $\mathsf{nil} : \mathbf{1} \to \ell ist(A)$, $\mathsf{cons} : A \times \ell ist(A) \to \ell ist(A)$ satisfying the following (weak) universal property: for any object B and arrows $b : \mathbf{1} \to B$ and $h : A \times B \to B$, there exists an "iterator" $\mathcal{I}_{bh} : \ell ist(A) \to B$ satisfying (in the internal language):

$$\mathcal{I}_{bh}\, \mathsf{nil} = b, \qquad \mathcal{I}_{bh}\, \mathsf{cons}\langle a, w\rangle = h\langle a, \mathcal{I}_{bh} w\rangle.$$

Here nil corresponds to the empty list, and cons takes an element of A and a list and concatenates the element onto the head of the list.

Analogously to (weak) natural numbers objects $\mathbf{N}$, we can use product types and exponentiation to extend iteration on $\ell ist(A)$ to *primitive recursion with parameters* (cf. [GLT], p. 92).

What n-ary numerical **Set** functions are represented by arrows $\mathbf{N}^n \to \mathbf{N}$ in a ccc? The answer, of course, depends on the ccc. In general, the best we could expect is the following (cf. [LS86], Part III, Section 2):

PROPOSITION 2.32. *Let $\mathcal{F}_{\mathbf{N}}$ be the free ccc with weak natural numbers object. The class of numerical total functions representable therein is properly contained between the primitive recursive and the Turing-machine computable functions.*

In general, such fast-growing functions as the Ackermann function are representable in any ccc with weak natural numbers object (see [LS86]). Analogous results hold for symmetric monoidal and monoidal closed categories, [PR89].

The question of *strong* versus *weak* datatypes is of some interest. For example, although we can define addition $+ : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ by primitive recursion on a weak natural numbers type, *commutativity* of addition follows from having a strong natural numbers object; a weak parametrized primitive recursor would only allow us to derive $x + n = n + x$ for each closed numeral n but we cannot then extend this to *variables* (cf. Gödel's incompleteness theorem, cf. [LS86], p. 263). Notice that, on the face of it, the definition of a natural numbers object appears not to be equational: informally, uniqueness of the arrow $\mathcal{I}_{ah}$ requires an *implication*: *for all* $f : \mathbf{N} \to A$ (*if* $f0 = a$ *and* $fS = hf$) *then* $f = \mathcal{I}_{ah}$.

Here we remark on a curious observation of Lambek [L88]. Let us recall from universal algebra that a Mal'cev operator on an algebra A is a function $m_A : A^3 \to A$ satisfying $m_A xxz = z$ and $m_A xzz = x$. For example, if A were a group, $m_A = xy^{-1}z$ is such an operator. Similarly, the definition of a Mal'cev operator on an object A makes sense in any ccc (e.g., as an arrow $A^3 \xrightarrow{m_A} A$ satisfying some diagrams) or, equivalently, in any typed lambda calculus (e.g., as a closed term $m_A : A^3 \Rightarrow A$ satisfying some equations).

THEOREM 2.33 (Lambek). *Let $\mathcal{C}$ be a ccc with weak natural numbers* $(\mathbf{N}, 0, S)$ *in which each object A has a Mal'cev operator m_A. Then the fact that* $(\mathbf{N}, 0, S)$ *is a natural numbers object is equationally definable using the family* $\{m_A \mid A \in \mathcal{C}\}$. *In particular, if $\mathcal{C} = \mathcal{F}_{\mathbf{N}}$, the free ccc with weak natural numbers object, there are a finite number of additional equations* (*as schema*) *that, when added to the original data, guarantee that every type has a Mal'cev operator and* $\mathbf{N}$ *is a natural numbers object.*

2.4. *Polymorphism*

"The perplexing subject of polymorphism."
C. Darwin, *Life & Lett*, 1887

Although Darwin was speaking of biology, he might very well have been discussing computer science 100 years later. Christopher Strachey in the 1960's introduced various notions of polymorphism into programming language design (see [Rey83,Mit96]). Perhaps the most influential was his notion of *parametric polymorphism*. Intuitively, a parametric polymorphic function is one which has a *uniformly given algorithm at all types*. Imagine a "generic" algorithm capable of being instantiated at any arbitrary type, but which is the "same algorithm" at each type instance. It is this idea of the "plurality of form" which inspired the biological metaphor.

EXAMPLE 2.34 (*Reverse*). Consider a simple algorithm that takes a finite list and reverses it. Here "lists" could mean: lists of natural numbers, lists of reals, lists of arrays, indeed lists of lists of .... The point is, the types do not matter: we have a uniform algorithm for all types. Let $list(\alpha)$ denote the type of finite lists of entities of type α. We thus might type this algorithm

$$rev_\alpha : list(\alpha) \Rightarrow list(\alpha) \quad \text{where } rev_\alpha(a_1, \dots, a_n) = (a_n, \dots, a_1).$$

A second example, discussed by Strachey, is

EXAMPLE 2.35 (*Map-list*). This algorithm begins with a function of type $\alpha \Rightarrow \beta$ and a finite α-list, applies the function to each element of the list, and then makes a β-list of the subsequent values. We might represent it as:

$$map_{\alpha,\beta} : (\alpha \Rightarrow \beta) \Rightarrow (list(\alpha) \Rightarrow list(\beta))$$

where $map_{\alpha,\beta}(f)(a_1, \ldots, a_n) = (f(a_1), \ldots, f(a_n))$.

Many recent programming languages (e.g., ML, Ada) support sophisticated uses of generic types and polymorphism. The mathematical foundations of such languages were a major challenge in the past decade and category theory played a fundamental role. We shall briefly recall the issues.

2.4.1. *Polymorphic lambda calculi.* The logician J.-Y. Girard [Gi71,Gi72] in a series of important works examined higher-order logic from the Curry–Howard viewpoint. He developed formal calculi of variable types, the so-called polymorphic lambda calculi, which correspond to proofs in higher-order logics. At the same time he developed the proof theory of such systems. J. Reynolds [Rey74] independently discovered the second-order fragment of Girard's system, and proposed it as a syntax representing Strachey's parametric polymorphism.

Let us briefly examine Girard's *System* $\mathcal{F}$, second order polymorphic lambda calculus. The underlying logical system is intuitionistic second order propositional calculus. The latter theory is similar to ordinary propositional calculus, except we can universally quantify over propositional variables.

The syntax of second order propositional calculus is presented in Figure 5. The usual notions of free and bound variables in formulas are assumed. For example, in $\forall\alpha(\alpha \Rightarrow \beta)$, α is a bound variable, while β is free. $A[B/\alpha]$ denotes A with formula B substituted for free α, changing bound variables if necessary to avoid clashes. Notice in the quantifier rules that when we instantiate a universally quantified formula to obtain, say, $\Gamma \vdash A[B/\alpha]$,

Formulas	$A ::= vbl \mid A_1 \Rightarrow A_2 \mid \forall\alpha.A$
Provability	$\vdash$ is a relation between finite sets of formulas and formulas
	$\Gamma \vdash A$ if $A \in \Gamma$
	$\dfrac{\Gamma \cup \{A\} \vdash B}{\Gamma \vdash A \Rightarrow B}, \quad \dfrac{\Gamma \vdash A \quad \Delta \vdash A \Rightarrow B}{\Gamma \cup \Delta \vdash B}$
	$\dfrac{\Gamma \vdash A(\alpha)}{\Gamma \vdash \forall\alpha A(\alpha)}, \quad \dfrac{\Gamma \vdash \forall\alpha A(\alpha)}{\Gamma \vdash A[B/\alpha]}$
	where $\alpha \notin FV(\Gamma)$ for any formula B.

Fig. 5. Second order intuitionistic propositional calculus.

the formula B may be of arbitrary logical complexity. Thus inductive proof techniques based on the complexity of subformulas are not available in higher-order logic. This is the essence of the problem of *impredicativity* in polymorphism.

We now introduce Girard's second order lambda calculus. We use the notation $FV(t)$ and $BV(t)$ for the set of free and bound variables of term t, respectively. We write $FTV(A)$ and $BTV(A)$ for the set of free type variables and bound type variables of formula A, respectively.

DEFINITION 2.36 (*Girard's System $\mathcal{F}$*).

Types: Freely generated from type variables $\alpha, \beta, \ldots$ by the rules: if A, B are types, so are $A \Rightarrow B$ and $\forall\alpha.A$.

Terms: Freely generated from variables x_i^A of every type A by

(1) First-order lambda calculus rules: if $f : A \Rightarrow B$, $a : A$, $\varphi : B$ then $f`a : B$ and $\lambda_{x:A}.\varphi : A \Rightarrow B$.

(2) Specifically second-order rules:

(a) If $t : A(\alpha)$, then $\Lambda\alpha.t : \forall\alpha A(\alpha)$ where $\alpha \notin FTV(FV(t))$,

(b) If $t : \forall\alpha A(\alpha)$ then $t[B] : A[B/\alpha]$ for any type B.

Equations: Equality is the smallest congruence relation closed under β and η for both lambdas, that is:

(3) $(\lambda_{x:A}.\varphi)`a =_{\beta^1} \varphi[a/x]$ and $\lambda_{x:A}(f`x) =_{\eta^1} f$, where $x \notin FV(f)$.

(4) $(\Lambda\alpha.\psi)[B] =_{\beta^2} \psi[B/\alpha]$ and $\Lambda\alpha.t[\alpha] =_{\eta^2} t$, where $\alpha \notin FTV(t)$.

Eqs. (3) are the first order $\beta\eta$ equations, while Eqs. (4) are second order $\beta\eta$.

From the Curry–Howard viewpoint, the types of $\mathcal{F}$ are precisely the formulas of second order propositional calculus (Fig. 5), while terms denote proofs. For example, to annotate second order rules we have:

$$\frac{\vec{x} : \Gamma \vdash t : A(\alpha)}{\vec{x} : \Gamma \vdash \Lambda\alpha \cdot t : \forall\alpha A(\alpha)}, \qquad \frac{\vec{x} : \Gamma \vdash t : \forall\alpha A(\alpha)}{\vec{x} : \Gamma \vdash t[B] : A[B/\alpha]}.$$

The $\beta\eta$ equations of course express equality of proof trees.

What about polymorphism? Suppose we think of a term $t : \forall\alpha A(\alpha)$ as an algorithm of type $A(\alpha)$ varying uniformly over all types α. Then $t[B] : A[B/\alpha]$ is the instantiation of t at the specific type B. Moreover, B may be arbitrarily complex. Thus the type variable acts as a parameter.

In System $\mathcal{F}$ we can internally represent common inductive data types within the syntax as *weak* T-algebras, for covariant definable functors T. Weakness refers to the categorical fact that these structures satisfy existence but not uniqueness of the mediating arrow in the universal mapping property. Thus, for any types A, B we are able to define the types $\mathbf{1}$, $\mathbf{Nat}$, $List(A)$, $A \times B$, $A + B$, $\exists\alpha \cdot A$, etc. (see [GLT] for a full treatment).

Let us give two examples and at the same time illustrate polymorphic instantiation.

EXAMPLE 2.37. The type of Booleans is given by

$$\mathbf{Boole} = \forall\alpha.(\alpha \Rightarrow (\alpha \Rightarrow \alpha)).$$

It has two distinguished elements $\mathbf{T}, \mathbf{F}$: **Boole** given by $\mathbf{T} = \Lambda\alpha.\lambda_{x:\alpha}.\lambda_{y:\alpha}.x$ and $\mathbf{F} = \Lambda\alpha.\lambda_{x:\alpha}.\lambda_{y:\alpha}.y$, together with a *Definition by Cases* operator (for each type A) $D_A : A \Rightarrow (A \Rightarrow (\mathbf{Boole} \Rightarrow A))$ defined by $D_A uvt = (t[A]`u)`v$ where $u, v : A$, $t : \mathbf{Boole}$. One may easily verify that $D_A uv\mathbf{T} =_\beta u$ and $D_A uv\mathbf{F} =_\beta v$ (where β stands for $\beta^1 \cup \beta^2$).

EXAMPLE 2.38. The type of (Church) numerals

$$\mathbf{Nat} = \forall\alpha((\alpha \Rightarrow \alpha) \Rightarrow (\alpha \Rightarrow \alpha)).$$

The numeral $\mathbf{n} : \mathbf{Nat}$ corresponds at each α to n-fold composition $f \mapsto f^n$, where $f^n = f \circ f \circ \cdots \circ f$ (n times) and $f^0 = Id_\alpha = \lambda x_\alpha \cdot x$. Formally, it is the closed term $\mathbf{n} = \Lambda\alpha.\lambda_{f:\alpha\Rightarrow\alpha}.f^n : \mathbf{Nat}$. Thus for *any* type B we have a uniform algorithm: $\mathbf{n}[B] = \lambda_{f:B\Rightarrow B}.f^n : (B \Rightarrow B) \Rightarrow (B \Rightarrow B)$. Successor is given by $\mathbf{S} = \lambda_{n:\mathbf{Nat}}.n + 1$, where $n + 1 = \Lambda\alpha.\lambda_{f:\alpha\Rightarrow\alpha}.f^{n+1} = \Lambda\alpha.\lambda_{f:\alpha\Rightarrow\alpha}.f \circ f^n = \Lambda\alpha.\lambda_{f:\alpha\Rightarrow\alpha} f \circ (n[\alpha]`f)$. Finally, iteration is given by: if $a : A$, $h : A \Rightarrow A$, $\mathcal{I}_{ah} = \lambda_{x:\mathbf{Nat}}.(x[A]`h)`a$: $\mathbf{Nat} \Rightarrow A$. The reader may easily calculate that $\mathcal{I}_{ah}\mathbf{0} =_\beta a$ and $\mathcal{I}_{ah}(\mathbf{n} + \mathbf{1}) =_\beta h`(\mathcal{I}_{ah}\mathbf{n})$ for numerals $\mathbf{n}$.

Let us illustrate the power of impredicativity in this situation. See the discussion of Church vs. Curry typing, Section 2.5.3. Notice that for any type B, $\mathbf{n}[B \Rightarrow B]`\mathbf{n}[B]$ makes perfectly good sense. In particular, let $B = \mathbf{Nat}$, the type of $\mathbf{n}$ itself. This is a well-defined term and if we erase all its types we obtain the *untyped* expression $\mathbf{n}`\mathbf{n} = \lambda f.f^n$. This latter untyped term is *not* typable in simply typed lambda calculus.

Formal systems describing far more powerful versions of polymorphism have been developed. For example, Girard's thesis described the typed lambda calculus corresponding to ω-order intuitionist type theory, so-called $\mathcal{F}_\omega$. Programming in the various levels of Girard's theories $\{\mathcal{F}_n\}$, $n = 1, 2, \ldots, \omega$, is described in [PDM89]. Other systems include Coquand–Huet's Calculus of Constructions and its extensions [Luo94]. These theories include not only Girard's $\mathcal{F}_\omega$ but also Martin-Löf's dependent type theories [H97a]. Indeed, these theories are among the most powerful logics known, yet form the basis of various proof-development systems (e.g., LEGO and Coq) [LP92,D$^+$93].

2.4.2. *What is a model of System $\mathcal{F}$?* The problem of finding – and indeed defining precisely – a model of System $\mathcal{F}$ was difficult. Cartesian closedness is not the issue. The problem, of course, is the universal quantifier: clearly in $\forall\alpha.A$ the α is to range over all the objects of the model, and at the same time $\forall$ should be interpreted as some kind of product (over *all* objects). Such "large" products create havoc, as foreshadowed in the following theorem of Freyd (cf. [Mac71], Proposition 3, p. 110).

THEOREM 2.39 (Freyd). *A small category which is small complete is a preorder.*

Cartesian closed preorders (e.g., complete Heyting algebras) are of no interest for modelling proofs; we seek "nontrivial" categories.

Suppose instead we try to define a naive "set-theoretic" model of System $\mathcal{F}$, in which $\times, \Rightarrow$ have their usual meaning, and $\forall\alpha$ is interpreted as a "large" product. Such models are defined in detail in [RP,Pi87]. John Reynolds proved the following

THEOREM 2.40. *There is no* **Set** *model for System* $\mathcal{F}$.

There is an elegant categorical proof in Reynolds and Plotkin [RP]. Let us sketch the proof, which applies to somewhat more general categories than **Set**.

Let C be a category with an endofunctor $T : C \to C$. A *T-algebra* is an object A together with an arrow $TA \xrightarrow{a} A$. A morphism of T-algebras is a commutative square:

$$\begin{tikzcd} TA \arrow[r, "Tf"] \arrow[d, "a"'] & TB \arrow[d, "b"] \\ A \arrow[r, "f"] & B \end{tikzcd}$$

An *initial T-algebra* (resp. *weakly initial T-algebra*) is one for which there exists a unique morphism (resp. there exists a morphism) to any other T-algebra.

We shall be interested in objects and arrows of the model category C which are "definable", i.e. denoted by types and terms of System $\mathcal{F}$. There are simple covariant endofunctors T on C whose action on objects is definable by types and whose actions on arrows is definable by terms (of System $\mathcal{F}$). For example, the identity functor $T(\alpha) = \alpha$ and the functor $T(\alpha) = (\alpha \Rightarrow B) \Rightarrow B$, for any fixed B, have this property.

Now it may be shown (see [RP]) that for any definable functor T, the System $\mathcal{F}$ expression $P = \forall\alpha.(T(\alpha) \Rightarrow \alpha) \Rightarrow \alpha$ is a weakly initial T-algebra. *Suppose the ambient model category C has equalizers of all subsets of arrows* (e.g., **Set** has this property). Essentially by taking a large equalizer (cf. the Solution Set Condition in Freyd's Adjoint Functor Theorem, [Mac71], p. 116) we could then construct a subalgebra of P which is an initial T-algebra. Call this initial T-algebra $\mathcal{I}$. We then use the following important observation of Lambek:

PROPOSITION 2.41 (Lambek). *If* $T(\mathcal{I}) \xrightarrow{f} \mathcal{I}$ *is an initial T-algebra, then f is an isomorphism.*

Applying this to the definable functor $T(\alpha) = (\alpha \Rightarrow B) \Rightarrow B$, we observe that $T(\mathcal{I}) \cong \mathcal{I}$. In particular, let $C =$ **Set** and $B =$ **Boole**, and take the usual **Set** interpretation of $\times$ as Cartesian product and $\Rightarrow$ as the full function space. Notice $card(B) \geqslant 2$ (since there are always the two distinct closed terms **T** and **F**). Hence we obtain a bijection $B^{B^{\mathcal{I}}} \cong \mathcal{I}$, for some set $\mathcal{I}$, which is impossible for cardinality reasons.

The search for models of System $\mathcal{F}$ led to some extraordinary phenomena that had considerable influence in semantics of programming languages. Let us just briefly mention the history. Notice that the Reynolds–Plotkin proof depends on a simple cardinality argument, which itself depends on classical set theory. Similarly, the proof of Freyd's result, Theorem 2.39, depends on using classical (i.e. non-intuitionistic) logical reasoning in the metalanguage. This suggests that it is really the non-constructive nature of the category **Sets** that is at fault; if we were to work within a non-classical universe – say within a model of intuitionistic set theory – there is still a chance that we could escape the above problems but still have a "set-theoretical" model of System $\mathcal{F}$. And, from one point of view, that is exactly what happened.

These ambient categories, called toposes [LS86,MM92], are in general models of intuitionistic higher-order logic (or set-theory), and include such categories as functor categories and sheaves on a topological space, as well as **Sets**. Moggi suggested constructing models of System $\mathcal{F}$ based on an internally complete internal full subcategory of a suitable ambient topos. This ambient topos would serve as our constructive set-theory, and function types would still be interpreted as the full "set-theoretical" space of total functions. M. Hyland [Hy88] proved that the Realizability (or Effective) Topos had (non-trivial) such internal category objects. The difficult development and clarification of these internal models was undertaken by many researchers, e.g., D. Scott, M. Hyland, E. Robinson, P. Rosolini, A. Carboni, P. Freyd, A. Scedrov, A. Pitts et al. (e.g., [HRR,Rob89,Ros90,CPS88,Pi87]).

In a separate development, R. Seely [See87] gave the first general categorical definition of a so-called *external model* of System $\mathcal{F}$, and more generally $\mathcal{F}_\omega$. The definition was based on the theory of *indexed* or *fibred* categories. This view of logic was pioneered by Lawvere [Law69] who emphasized that quantifiers were interpretable as adjoint functors. Pitts [Pi87] clarified the relationship between Seely's models and internal-category models within ambient toposes of presheaves. Moreover, he showed that there are enough such internal models for a Completeness Theorem. It is worth remarking that Pitts' work uses properties of Yoneda embeddings. For general expositions see [AL91]. Extensions of "set-theoretical" models to cases where function spaces include *partial functions* (i.e. non-termination) is in [RR90].

One can *externalize* these internal category models [Hy88,AL91] to obtain ordinary categories. And one such internal category in the Realizability Topos, the *modest sets*, when externalized is precisely the ccc category $Per(\mathbf{N})$ discussed in Section 2.1.

PROPOSITION 2.42. *$Per(\mathbf{N})$ is a model of System $\mathcal{F}$.*

The idea is that in addition to the ccc structure of $Per(\mathbf{N})$, we interpret $\forall$ as a large intersection (the intersection of an arbitrary family of pers is again a per). We shall return to this example in Section 3.2.

Ironically, in essence this model was already in Girard's original PhD thesis [Gi72]. Later, domain-theoretic models of System $\mathcal{F}$ were considered by Girard in [Gi86] and were instrumental in his development of linear logic

2.5. *The untyped world*

The advantages of types in programming languages are familiar and well-documented (e.g., [Mit96]). Nonetheless, there is an underlying *untyped* aspect of computation, already going back to the original work on lambda calculus and combinatory logic in the 1930's, which often underlies concrete machine implementations. In this early view, developed by Church, Curry, and Schönfinkel, functions were understood in the old-fashioned (pre-Cantor) sense of "rules", as a computational process of going from an argument to a value. Such a functional process could take anything, even itself, as an argument. Let us just briefly mention some key directions (see [Bar84,AGM]).

2.5.1. *Models and denotational semantics.* From the viewpoint of ccc's, untypedness amounts to finding a ccc $\mathcal{C}$ with an object $D \not\cong \mathbf{1}$ satisfying the isomorphism

$$D^D \cong D. \tag{7}$$

Thus function spaces and elements are "on the same level". It then makes sense to define formal application $f`g$ for constants $f, g : D^D$ by $f`g = ev(f, \varphi(g))$, where $\varphi : D^D \xrightarrow{\cong} D$ is the isomorphism above. In particular, self-application $f`f$ makes perfectly good sense.

Dana Scott found the first semantical (topological) models of the untyped lambda calculus in 1970 [Sc72]; i.e. non-trivial solutions D to "equations" of the form (7) in various ccc's, perhaps the simplest being in ω-**CPO**. This was part of his general investigations (with Christopher Strachey) into the foundations of programming languages, culminating in the so-called Scott–Strachey approach to the semantics of programming languages. Arguably, this has been one of the major arenas in the use of category theory in Computer Science, with an enormous literature. For an introduction, see [AbJu94,Gun92,Ten94].

More generally, one seeks to find non-trivial domains D satisfying certain so-called "recursive domain equations", of the form

$$D \cong \cdots D \cdots, \tag{8}$$

where $\cdots D \cdots$ is some expression built from type constructors. The difficulty is that the variable D may appear both co- and contravariantly. Such recursive defining "equations" are used to specify the semantics of numerous notions in computer science, from datatypes in functional programming languages, to modelling nondeterminism, concurrency, etc. (cf. also [DiCo95]).

The seminal early paper on categorical solutions of domain equations is the paper of Smyth and Plotkin [SP82]. More recent work has focussed on axiomatic and synthetic domain theory (e.g., [AbJu94,FiPl96,ReSt97]) and use of bisimulations and relation-theoretic methods for reasoning about recursive domains [Pi96a]. These methods rely on fundamental work of Peter Freyd on recursive types (e.g., [Fre92]).

2.5.2. *C-monoids and categorical combinators.* On a more algebraic level, a model of untyped lambda calculus is a ccc with (up to isomorphism) one non-trivial object. That is, a ccc $\mathcal{C}$ with an object $D \not\cong \mathbf{1}$ satisfying the domain equations:

$$D \cong D^D \cong D \times D. \tag{9}$$

An example of such a D in ω-**CPO** is given in [LS86], using the constructions of D. Scott and Smyth–Plotkin mentioned above. An interesting axiomatization of such D's comes from simply considering $Hom_{\mathcal{C}}(D, D) \cong Hom_{\mathcal{C}}(\mathbf{1}, D^D)$ as an abstract monoid. It turns out that the axioms are easy to obtain: take the axioms of a ccc, remove the terminal object, *and erase all the types!* That is (following the treatment in [LS86], p. 93):

DEFINITION 2.43. A *C-monoid* (C for Curry, Church, Combinatory, or CCC) is a monoid $(\mathcal{M}, \circ, id)$ together with extra structure structure $(\pi_1, \pi_2, \varepsilon, (-)^*, \langle -, - \rangle)$ where π_i, ε are elements of $\mathcal{M}$, $(-)^*$ is a unary operation on $\mathcal{M}$, and $\langle -, - \rangle$ is a binary operation on $\mathcal{M}$, satisfying untyped versions of the equations of a ccc (cf. Figure 1):

$$\begin{aligned} \pi_1\langle a, b\rangle &= a, & \varepsilon\langle h^*\pi_1, \pi_2\rangle &= h, \\ \pi_2\langle a, b\rangle &= b, & (\varepsilon\langle k\pi_1, \pi_2\rangle)^* &= k, \\ \langle \pi_1 c, \pi_2 c\rangle &= c \end{aligned}$$

for any $a, b, c, h, k \in \mathcal{M}$ (where we elide the monoid operation $\circ$).

C-monoids were first discovered independently by D. Scott and J. Lambek around 1980. The elementary algebraic theory and connections with untyped lambda calculus were developed in [LS86] (and independently in [Cur93], where they were called *categorical combinators*). Obviously C-monoids form an equational class; thus, just like for general ccc's, we may form free algebras, polynomial algebras, prove Functional Completeness, etc. The associated internal language is *untyped lambda calculus with pairing operators*. As above, this language is obtained from simply typed lambda calculus by omitting the type **1** and erasing all the types from terms.

The rewriting theory of categorical combinators has been discussed by Curien, Hardin et al. (e.g., see [Har93]). Categorical combinators form a particularly efficient mechanism for implementing functional languages; for example, the language CAML is a version of the functional language ML based on categorical combinators (see [Hu90], Part 1).

The deepest mathematical results on *the Cartesian fragment of* C-monoids to date have been obtained by R. Statman [St96]. In particular, Statman characterizes the free Cartesian monoid F (in terms of a representation into certain continuous shift operators on Cantor space), as well as characterizing the finitely generated submonoids of F and the recursively enumerable subsets of F. The latter two results are based on projections of (suitably encoded) unification problems.

2.5.3. *Church vs. Curry typing.* The fundamental feature of the untyped lambda calculus is self-application. The β-rule $\lambda x \cdot \varphi(x)`a = \varphi[a/x]$ is now totally unrestricted with respect to typing constraints. This permits non-halting computations: for example, the term $\Omega =_{def} (\lambda x.x`x)`(\lambda x.x`x)$ has no normal form and only β-reduces to itself, while the fixed point combinator $\mathsf{Y} =_{def} \lambda f.(\lambda x.f`(x`x))`(\lambda x.f`(x`x))$ satisfies $f`(\mathsf{Y}`f) =_{\beta\eta} \mathsf{Y}`f$, hence $\mathsf{Y}`f$ is a fixed point of f, for any f. Hence we immediately obtain: *all terms of the untyped lambda calculus have a fixed point.*

Untyped lambda calculus suggests a different approach to the typed world: "typing" untyped terms. The Church view (which we have adopted here) insists all terms be explicitly typed, starting with the variables. On the other hand Curry, the founder of the related but older subject of Combinatory Logic, had a different view: start with untyped terms, but add "type inference rules" to algorithmically infer appropriate types (when possible). Many modern typed programming languages (e.g., ML) essentially follow this Curry view and use "typing rules" to assign appropriate type schema to untyped terms. This leads to the so-called Type Inference Problem: given an explicitly typed language $\mathcal{L}$ and type erasure function $\mathcal{L} \xrightarrow{Erase} \mathcal{U}$ (where $\mathcal{U}$ is untyped lambda calculus), decide if an untyped term t satisfies $t = Erase(M)$ for some $M \in \mathcal{L}$. It turns out that a problem of type inference is essentially equivalent to a so-called unification problem, familiar from logic programming (cf. [Mit96]). Fortunately in the case of ML and other typed programming languages there are known type inference algorithms; however in general (e.g., for System $\mathcal{F}$, $\mathcal{F}_\omega, \ldots$) the problem is undecidable. To the best of our knowledge, the Church-vs-Curry view of typed languages has not yet been systematically analyzed categorically.

2.6. *Logical relations and logical permutations*

Logical relations play an important role in the recent proof theory and semantics of typed lambda calculi [Mit96,Plo80,St85]. Recall the notion of *Henkin model* (Section 2.1) as a subccc of **Set**.

DEFINITION 2.44. Given two Henkin models $\mathcal{A}$ and $\mathcal{B}$, a *logical relation* from $\mathcal{A}$ to $\mathcal{B}$ is a family of binary relations $\mathcal{R} = \{R_\sigma \subseteq A_\sigma \times B_\sigma \mid \sigma \text{ a type}\}$ satisfying (we write $a R_\sigma b$ for $(a, b) \in R_\sigma$):

(1) $* R_1 *$

(2) $(a, a') R_{\sigma\times\tau} (b, b')$ if and only if $a R_\sigma b$ and $a' R_\tau b'$, for any $(a, a') \in A_{\sigma\times\tau}$, $(b, b') \in B_{\sigma\times\tau}$, *i.e. ordered pairs are related exactly when their components are.*

(3) For any $f \in A_{\sigma\Rightarrow\tau}$, $g \in B_{\sigma\Rightarrow\tau}$, $f R_{\sigma\Rightarrow\tau}\, g$ if and only if for all $a \in A_\sigma$, $b \in B_\sigma$ ($a R_\sigma b$ implies $fa R_\tau gb$), *i.e. functions are related when they map related inputs to related outputs.*

For each (atomic) base type b, fix a binary relation $R_b \subseteq A_b \times B_b$. Then: there is a smallest family of binary relations $\mathcal{R} = \{R_\sigma \subseteq A_\sigma \times B_\sigma \mid \sigma \text{ a type}\}$ defined inductively from the R_b's by (1)–(3) above. That is, any property (relation) at base-types can be inductively lifted to a family $\mathcal{R}$ at all higher types, satisfying (1)–(3) above. We write $a\mathcal{R}b$ to denote $a R_\sigma b$ for some σ. If $\mathcal{A} = \mathcal{B}$ and $\mathcal{R}$ is a logical relation from $\mathcal{A}$ to itself, we say an element $a \in \mathcal{A}$ is *invariant* under $\mathcal{R}$ if $a\mathcal{R}a$.

The fundamental property of logical relations is the Soundness Theorem [Mit96,St85]. Let $\vec{x} : \Gamma \vdash M : \sigma$ denote that M is a term of type σ with free variables $\vec{x}$ in context Γ. Consider Henkin models $\mathcal{A}$ with $\eta_{\mathcal{A}}$ an assignment function, assigning variables to elements in $\mathcal{A}$. Let $[\![M]\!]_{\eta_{\mathcal{A}}}$ denote the meaning of term M in model $\mathcal{A}$ w.r.t. the given variable assignment (following [Mit96], we only consider assignments η such that $\eta_{\mathcal{A}}(x_i) \in A_\sigma$ if $x_i : \sigma \in \Gamma$). The following is proved by induction on the form of M:

THEOREM 2.45 (Soundness). *Let $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{B}$ be a logical relation between Henkin models $\mathcal{A}$, $\mathcal{B}$. Let $\vec{x} : \Gamma \vdash M : \sigma$. Suppose assignments $\eta_{\mathcal{A}}$, $\eta_{\mathcal{B}}$ of the variables are related, i.e. for all x_i, $\mathcal{R}(\eta_{\mathcal{A}}(x_i), \eta_{\mathcal{B}}(x_i))$. Then $\mathcal{R}([\![M]\!]_{\eta_{\mathcal{A}}}, [\![M]\!]_{\eta_{\mathcal{B}}})$.*

In particular, if $\mathcal{A} = \mathcal{B}$ and M is a closed term (i.e. contains no free variables), its meaning $[\![M]\!]$ in a model $\mathcal{A}$ is invariant under all logical relations. This holds also for *languages which have constants at base types, by assuming $[\![c]\!]$ is invariant, for all such constants c.*

This result has been used by Plotkin, Statman, Sieber et al. [Plo80,Sie92,St85] to show certain elements (of models) are *not* lambda definable: it suffices to find some logical relation on $\mathcal{A}$ for which the element in question is not invariant.

There is no reason to restrict ourselves to binary logical relations: one may speak of n-ary logical relations, which relate n Henkin models [St85]. Indeed, since Henkin models are closed under products, it suffices to consider unary logical relations, known as *logical predicates*.

EXAMPLE 2.46 (*Hereditary permutations*). Consider a Henkin model $\mathcal{A}$, with a specified permutation $\pi_b : A_b \to A_b$ at each base type b. We extend π to all types as follows: (i) on product types we extend componentwise: $\pi_{\sigma\times\tau} = \pi_\sigma \times \pi_\tau : A_{\sigma\times\tau} \to A_{\sigma\times\tau}$; (ii) on function spaces, extend by conjugation: $\pi_{\sigma\Rightarrow\tau}(f) = \pi_\tau \circ f \circ \pi_\sigma^{-1}$, where $f \in A_{\sigma\Rightarrow\tau}$. We build a logical relation $\mathcal{R}$ on $\mathcal{A}$ by letting R_σ = the graph of permutation $\pi_\sigma : A_\sigma \to A_\sigma$, i.e. $R_\sigma(a,b) \Leftrightarrow \pi_\sigma(a) = b$. Members of $\mathcal{R}$ will be called hereditary permutations. $\mathcal{R}$-invariant elements $a \in A_\sigma$ are simply fixed points of the permutation: $\pi_\sigma(a) = a$.

Hereditary permutations and invariant elements also arise categorically by interpretation into $\mathbf{Set}^{\mathbf{Z}}$:

PROPOSITION 2.47. *The category* $\mathbf{Set}^{\mathbf{Z}}$ *of (left)* $\mathbf{Z}$*-sets is equivalent to the category whose objects are sets equipped with a permutation and whose maps (= equivariant maps) are functions commuting with the distinguished permutations. Invariant elements of A are arrows* $\mathbf{1} \to A \in \mathbf{Set}^{\mathbf{Z}}$.

Alas, $\mathbf{Set}^{\mathbf{Z}}$ is not a Henkin model ($\mathbf{1}$ is not a generator). In the next section we shall slightly generalize the notion of logical relation to work on a larger class of structures.

2.6.1. *Logical relations and syntax.* Logical predicates (also called *computability predicates*) originally arose in proof theory as a technique for proving normalization and other syntactical properties of typed lambda calculi [LS86,GLT]. Later, Plotkin, Statman, and Mitchell [Plo80,St85,Mit96] constructed logical relations on various kinds of structures more general than Henkin models. Following Statman and Mitchell, we extend the notion of logical relation to certain *applicative typed structures* $\mathcal{A}$ for which (i) appropriate meaning functions on the syntax of typed lambda calculus, $[\![M]\!]_{\eta_\mathcal{A}}$, are well-defined, and (ii) all logical relations $\mathcal{R}$ are (in a suitable sense) congruence relations with respect to the syntax. This guarantees that the meanings of lambda abstraction and application behave appropriately under these logical relations. Following [Mit96,St85] we call them *admissible* logical relations.

The Soundness Theorem still holds in this more general setting, now using admissible logical relations on applicative typed structures (see [Mit96], Lemma 8.2.10).

EXAMPLE 2.48. Let $\mathcal{A}$ be the hereditary permutations in Example 2.46. Consider a free simply typed lambda calculus, without constants. Then as a corollary of Soundness we have: *the meaning of any closed term M is invariant under all hereditary permutations.* This conclusion is itself a consequence of the universal property of free Cartesian closed categories when interpreted in $\mathbf{Set}^{\mathbf{Z}}$ (cf. Corollary 2.22 and Läuchli's Theorem 5.4).

REMARK 2.49. The rewriting theory of lambda calculi is a prototype for Operational Semantics of many programming languages (recall the discussion after the β-rule, Corollary 2.11). See also Section 2.8.1 below on PCF. Logical Predicates (so-called *computability predicates*) were first introduced to prove strong normalization for simply typed lambda calculi (with natural numbers types) by W. Tait in the 1960's. Highly sophisticated computability predicates for polymorphically typed systems like $\mathcal{F}$ and $\mathcal{F}_\omega$ were first introduced by Girard in his thesis [Gi71,Gi72]. For a particularly clear presentation, see his book [GLT]. These techniques were later revisited by Statman and Mitchell – using more general logical relations – to also prove Church–Rosser and a host of other syntactic and semantic results for such calculi (see [Mit96]).

For general categorical treatments of logical relations, see [Mit96,MitSce,MaRey] and references there. Uses of logical relations in operational semantics of typed lambda calculi are covered in [AC98,Mit96]. A categorical theory of logical relations applied to data refinement is in [KOPTT]. Use of operationally-based logical relations in programming language semantics is in [Pi96b,Pi97]. For techniques of categorical rewriting applied to lambda calculus, see [JGh95].

2.7. *Example 1: Reduction-free normalization*

The operational semantics of λ-calculi have traditionally been based on rewriting theory or proof theory, e.g., normalization or cut-elimination, Church–Rosser, etc. More recently, Berger and Schwichtenberg [BS91] gave a model-theoretic extraction of normal forms – a kind of "inverting" of the canonical set-theoretic interpretation used in Friedman's Completeness Theorem (cf. 5.2 below).

In this section we sketch the use of categorical methods (essentially from Yoneda's Lemma, cf. Theorem 5.1) to obtain the Berger–Schwichtenberg analysis. A first version of this technique was developed by Altenkirch, Hoffman, and Streicher [AHS95,AHS96]. The analysis given here comes from the article [CDS97], which also mentions intriguing analogues to the Joyal–Gordon–Power–Street techniques for proving coherence in various structured (bi-)categories. The essential idea common to these coherence theorems is to use a version of Yoneda's lemma to embed into a "stricter" presheaf category.

To actually extract a normalization algorithm from these observations requires us to constructively reinterpret the categorical setting in $\mathcal{P}\mathbf{Set}$, as explained below. This leads to a non-trivial example of program extraction from a structured proof, in a manner advocated by Martin-Löf and his school [ML82,Dy95,H97a,CD97]. The reader is referred to [CDS97] for the fine details of the proof.

In a certain sense, the results sketched below are "dual" to Lambek's original goal of categorical proof theory [L68,L69], in which he used cut-elimination to study categorical coherence problems. Here, we use a method inspired from categorical coherence proofs to normalize simply typed lambda terms (and thus intuitionistic proofs).

2.7.1. *Categorical normal forms.* Let $\mathcal{L}$ be a language, $\mathcal{T}$ the set of $\mathcal{L}$-terms and $\sim$ a congruence relation on $\mathcal{T}$. One way to decide whether two terms are $\sim$-congruent is to find an *abstract normal form function*, i.e. a computable function $\mathbf{nf}: \mathcal{T} \to \mathcal{T}$ satisfying the following conditions for some (finer) congruence relation $\equiv$:

(NF1) $\mathbf{nf}(f) \sim f$,
(NF2) $f \sim g \Rightarrow \mathbf{nf}(f) \equiv \mathbf{nf}(g)$,
(NF3) $\equiv \subseteq \sim$,
(NF4) $\equiv$ is decidable.

From (NF1), (NF2) and (NF3) we see that $f \sim g \Leftrightarrow \mathbf{nf}(f) \equiv \mathbf{nf}(g)$. This clearly permits a decision procedure: to decide if two terms are $\sim$-related, compute **nf** of each one, and see if they are $\equiv$ related, using (NF4). The normal form function **nf** essentially "reduces" the decision problem of $\sim$ to that of $\equiv$. This view is inspired from [CD97].

Here we let $\mathcal{L}$ be typed lambda calculus, $\mathcal{T}$ the set of λ-terms, $\sim$ be $\beta\eta$-conversion, and $\equiv$ be α-congruence. Let us see heuristically how category theory can be used to give simply typed λ-calculus a normal form function **nf**.

Recall 2.22 that λ-terms modulo $\beta\eta$-conversion $\sim_{\beta\eta}$ determine the free ccc $\mathcal{F}_{\mathcal{X}}$ on the set of sorts (atoms) $\mathcal{X}$.[2] By the universal property 2.22, for any ccc $\mathcal{C}$ and any interpretation of the atoms $\mathcal{X}$ in $ob(\mathcal{C})$, there is a unique (up to iso) ccc-functor $[\![-]\!]: \mathcal{F}_{\mathcal{X}} \to \mathcal{C}$ freely extending this interpretation. Let $\mathcal{C}$ be the presheaf category $\mathbf{Set}^{\mathcal{F}_{\mathcal{X}}^{op}}$. There are two obvious ccc-functors: (i) the Yoneda embedding $\mathcal{Y}: \mathcal{F}_{\mathcal{X}} \to \mathbf{Set}^{\mathcal{F}_{\mathcal{X}}^{op}}$ (cf. 5.1) and (ii) if we interpret the atoms by Yoneda, there is also the free extension to the ccc-functor $[\![-]\!]: \mathcal{F}_{\mathcal{X}} \to \mathbf{Set}^{\mathcal{F}_{\mathcal{X}}^{op}}$. By the universal property, there is a natural isomorphism $q: \mathcal{Y} \to [\![-]\!]$. By the Yoneda lemma we shall invert the interpretation $[\![-]\!]$ on each hom-set, according to the following commutative diagram:

$$\begin{tikzcd}
\mathcal{F}_{\mathcal{X}}(A,B) \arrow[rr, "{[\![-]\!]}"] \arrow[dr, shift left, "\mathcal{Y}"] & & \mathbf{Set}^{\mathcal{F}_{\mathcal{X}}^{op}}([\![A]\!],[\![B]\!]) \arrow[dl, "q_B \circ - \circ q_A^{-1}"] \\
& \mathbf{Set}^{\mathcal{F}_{\mathcal{X}}^{op}}(\mathcal{Y}A, \mathcal{Y}B) \arrow[ul, shift left, "-_A(1_A)"] &
\end{tikzcd}$$

That is, for any $f \in \mathcal{F}_{\mathcal{X}}(A,B)$, we obtain natural transformations $\mathcal{Y}A \xrightarrow{q_A^{-1}} [\![A]\!] \xrightarrow{[\![f]\!]} [\![B]\!] \xrightarrow{q_B} \mathcal{Y}B$. Then evaluating these transformations at A gives the **Set** functions: $\mathcal{F}_{\mathcal{X}}(A,A) \xrightarrow{q_{A,A}^{-1}} [\![A]\!]A \xrightarrow{[\![f]\!]_A} [\![B]\!]A \xrightarrow{q_{B,A}} \mathcal{F}_{\mathcal{X}}(A,B)$. Hence starting with $1_A \in \mathcal{F}_{\mathcal{X}}(A,A)$, we can *define* an **nf** function by:

$$\mathbf{nf}(f) =_{def} q_{B,A}\left([\![f]\!]_A\left(q_{A,A}^{-1}(1_A)\right)\right). \tag{10}$$

[2] We actually use the free ccc of sequences of λ-terms as defined by Pitts [Pi9?].

Clearly $\mathbf{nf}(f) \in \mathcal{F}_{\mathcal{X}}(A, B)$. But, alas, by Yoneda's Lemma, $\mathbf{nf}(f) = f$! Indeed, this is just a restatement of part of the Yoneda isomorphism. But all is not lost: recall NF1 says $\mathbf{nf}(f) \sim f$. This suggests we should reinterpret the entire categorical argument, including the use of functor categories and Yoneda's Lemma, in a setting where "=" becomes "$\sim$", a (partial) equivalence relation. Diagrams which previously commuted now should commute "up to $\sim$".

This viewpoint has a long history in constructive mathematics, where it is common to use sets $(X, \sim)$ equipped with explicit equivalence relations in place of quotients $X/\sim$ (because of problems with the Axiom of Choice). Thus, along with specifying elements of a set, one must also say what it means for two elements to be "equal" (see [Bee85]).

2.7.2. *$\mathcal{P}$-category theory and normalization algorithms.* Motivated by enriched category theory [K82,Bor94], this view leads to the setting of *$\mathcal{P}$-category theory* in [CDS97]. In $\mathcal{P}$-category theory

(i) hom-sets are $\mathcal{P}$**Sets**, i.e. sets equipped with a partial equivalence relation (per) $\sim$,
(ii) all operations on arrows are $\mathcal{P}$**Set** maps, i.e. preserve $\sim$,
(iii) functors are $\sim$-versions of enriched functors,
(iv) $\mathcal{P}$-functor categories and $\mathcal{P}$-natural transformations are $\sim$-versions of appropriate enriched structure, etc.

One then proves a $\mathcal{P}$-version of Yoneda's lemma. In essence, $\mathcal{P}$-category theory is the development of ordinary (enriched) category theory in a constructive setting, where equality of arrows is systematically replaced by explicit pers, making sure every operation on arrows is a congruence with respect to the given pers. For an example, see Figure 6.

$\mathcal{P}$-Products

- $c \sim c$ for any constant $c \in \{!_A, \pi_1, \pi_2\}$,
- $f \sim f'$ for any $f, f' : A \to \mathbf{1}$,
- $f_i \sim g_i$ implies $\langle f_1, f_2\rangle \sim \langle g_1, g_2\rangle$, $\pi_i\langle f_1, f_2\rangle \sim f_i$ where $f_i, g_i : C \to A_i$,
- $\langle \pi_1 k, \pi_2 k\rangle \sim k$, for $k : C \to A_1 \times A_2$

$\mathcal{P}$-Exponentials

- $ev \sim ev$ for $B^A \times A \xrightarrow{ev} B$,
- $h \sim h'$ implies $h^* \sim h'^* : C \to B^A$, where $h, h' : C \times A \to B$
- $h \sim h'$ implies $ev\langle h^*\pi_1, \pi_2\rangle \sim h'$,
- $l \sim l'$ implies $(ev\langle l\pi_1, \pi_2\rangle)^* \sim l' : C \to B^A$.

$\mathcal{P}$-Functor

- $f \sim f'$ implies $Ff \sim Ff'$ for all f, f',
- $f \sim f'$, $g \sim g'$ implies $F(gf) \sim Fg'Ff'$ for all composable f, g and f', g',
- $F(id_A) \sim id_{F(A)}$,
- Specified $\mathcal{P}$-isomorphisms $\mathbf{1} \xrightarrow{\cong} F\mathbf{1}$, $FA \times FB \xrightarrow{\cong} F(A \times B)$, $(FB)^{FA} \xrightarrow{\cong} F(B^A)$

Fig. 6. $\mathcal{P}$-ccc's and $\mathcal{P}$-ccc functors.

Now consider the free ccc $(\mathcal{F}_{\mathcal{X}}, \sim)$ as a $\mathcal{P}$-category, where the arrows are actually sequences of λ-terms and the per $\sim$ on arrows is $\beta\eta$-equality $=_{\beta\eta}$. Analogously to the above, freeness in the $\mathcal{P}$-setting yields a unique $\mathcal{P}$-ccc functor

$$[\![-]\!] : (\mathcal{F}_{\mathcal{X}}, \sim) \to \mathcal{P}\mathbf{Set}^{(\mathcal{F}_{\mathcal{X}}, \sim)^{op}},$$

where atoms $X \in \mathcal{X}$ are interpreted by $\mathcal{P}$-Yoneda, i.e. as $Hom_{(\mathcal{F}_{\mathcal{X}}, \sim)}(-, X)$. Just as in the ordinary case, the $\mathcal{P}$-Yoneda functor $\mathcal{Y}$ is a $\mathcal{P}$-ccc functor, so we have a $\mathcal{P}$-natural isomorphism $q : [\![-]\!] \to \mathcal{Y}$ of $\mathcal{P}$-ccc $\mathcal{P}$-functors:

$$\begin{tikzcd} (\mathcal{F}_{\mathcal{X}}, \sim) \arrow[r, shift left=2, "{[\![-]\!]}"] \arrow[r, shift right=2, "\mathcal{Y}"'] \arrow[r, phantom, "q \downarrow\ \uparrow q^{-1}"] & \mathcal{P}\mathbf{Set}^{(\mathcal{F}_{\mathcal{X}}, \sim)^{op}}. \end{tikzcd}$$

In this setting $\mathbf{nf}$ as defined by Eq. (10) will be a per-preserving function *on terms themselves* and not just on $\beta\eta$-equivalence classes of terms (recall that, classically, a free ccc has for its arrows equivalence classes of terms modulo the appropriate equations). Arguing as before, but now using the $\mathcal{P}$-Yoneda isomorphism, it follows immediately that $\mathbf{nf}$ is *an identity $\mathcal{P}$-function*. But this means $\mathbf{nf}(f) \sim f$, which is precisely the statement of NF1. Moreover, the part of $\mathcal{P}$-category theory that we use is constructive in the sense that all functions we construct are algorithms. Therefore $\mathbf{nf}$ is computable.

It remains to prove NF2: $f \sim g \Rightarrow \mathbf{nf}(f) \equiv \mathbf{nf}(g)$. This is the most subtle point. Here too the $\mathcal{P}$-version of a general categorical fact will help us (cf. 2.4): the $\mathcal{P}$-presheaf category $\mathcal{P}\mathbf{Set}^{\mathcal{C}^{op}}$ is a $\mathcal{P}$-ccc for any $\mathcal{P}$-category $\mathcal{C}$. In particular, let $\mathcal{C}$ be the $\mathcal{P}$-category $(\mathcal{F}_{\mathcal{X}}, \equiv)$ of sequences of λ-terms up to "change of bound variable" $\equiv$. This is a trivially decidable equivalence relation on terms (called *α-congruence* in the literature) and obviously $\equiv\ \subseteq\ =_{\beta\eta}$. Note that this $\mathcal{P}$-category has the same objects and arrows as $(\mathcal{F}_{\mathcal{X}}, \sim)$, but the pers on arrows are different.

By the freeness of $(\mathcal{F}_{\mathcal{X}}, \sim)$, we have another interpretation $\mathcal{P}$-functor

$$[\![-]\!]^{\equiv} : (\mathcal{F}_{\mathcal{X}}, \sim) \to \mathcal{P}\mathbf{Set}^{(\mathcal{F}_{\mathcal{X}}, \equiv)^{op}},$$

where we interpret atoms $X \in \mathcal{X}$ by the $\mathcal{P}$-presheaf $Hom_{(\mathcal{F}_{\mathcal{X}}, \equiv)}(-, X)$. The key fact is that this $\mathcal{P}$-functor $[\![-]\!]^{\equiv}$ has exactly the same set-theoretic effect on objects and arrows as $[\![-]\!]$. That is, one proves by induction:

LEMMA 2.50. *For all objects C and arrows f in $(\mathcal{F}_{\mathcal{X}}, \sim)$, $|[\![C]\!]| = |[\![C]\!]^{\equiv}|$ and similarly $|[\![f]\!]| = |[\![f]\!]^{\equiv}|$, where $|-|$ means taking the underlying set-theoretic structure.*

Hence, we can conclude that $f \sim g$ implies $[\![f]\!] \equiv [\![g]\!]$ (here $\equiv$ refers to the per on arrows in $\mathcal{P}\mathbf{Set}^{(\mathcal{F}_{\mathcal{X}}, \equiv)^{op}}$). We can show that q_B and q_B^{-1} are $\equiv$-natural, in particular $q_{B,A}$ and $q_{B,A}^{-1}$ preserve $\equiv$. It then follows that $\mathbf{nf}(f) \equiv \mathbf{nf}(g)$, as desired.

REMARK 2.51.

(i) The normal forms obtained by this method can be shown to coincide with the so-called *long $\beta\eta$ normal forms* used in lambda calculus [CDS97].

(ii) The direct inductive proofs used above correspond more naturally to a more-involved bicategorical definition of freeness ([CDS97], Remark 3.17).

Finally, in [CDS97] it is shown how to apply the method to the word problem for typed λ-calculi with additional axioms and operations, i.e. to freely-generated ccc's modulo certain theories. This employs appropriate free $\mathcal{P}$-ccc's (over a $\mathcal{P}$-category, a $\mathcal{P}$-Cartesian category, etc). These are generated by various notions of λ-theory, which are determined not only by a set of atomic types, but also by a set of basic typed constants as well as a set of equations between terms. Although the Yoneda methods always yield an algorithm **nf** it does not necessarily satisfy NF4 (the decidability of $\equiv$). What is obtained in these cases is a reduction of the word problems for such free ccc's to those of the underlying generating categories.

2.8. *Example 2: PCF*

The language PCF, due to Dana Scott in 1969, has deeply influenced recent programming language theory. Much of this influence arises from seminal work of Gordon Plotkin in the 1970's on operational and denotational semantics for PCF. We shall briefly outline the syntax and basic semantical issues of the language following from Plotkin's work. We follow the treatment in [AC98,Sie92], although the original paper [Plo77] is highly recommended.

2.8.1. *PCF.* The language PCF is an explicitly typed lambda calculus with the following structure:

Types: Generated from *nat*, *boole* by $\Rightarrow$.

Lambda Terms: generated from typed variables using the following specified constants:

$\mathbf{n} : nat$, for each $n \in \mathbf{N}$	$zero? : nat \Rightarrow boole$
$\mathbf{T} : boole$	$cond_{nat} : boole \Rightarrow (nat \Rightarrow (nat \Rightarrow nat))$
$\mathbf{F} : boole$	$cond_{boole} : boole \Rightarrow (boole \Rightarrow (boole \Rightarrow boole))$
$succ : nat \Rightarrow nat$	$\perp_\sigma : \sigma$
$pred : nat \Rightarrow nat$	$\mathsf{Y}_\sigma : (\sigma \Rightarrow \sigma) \Rightarrow \sigma.$

Categorical models. The *standard model* of PCF is defined in the ccc ω-CPO$_\perp$ as follows: interpret the base types as in Figure 4 $[\![nat]\!] = \mathbf{N}_\perp$, $[\![boole]\!] = \mathbf{B}_\perp$, $[\![\sigma \Rightarrow \tau]\!] = [\![\sigma]\!] \Rightarrow [\![\tau]\!]$ (= function space in ω-CPO). Interpret constants as follows (for clarity, we omit writing $[\![-]\!]$): $succ, pred : \mathbf{Nat}_\perp \to \mathbf{Nat}_\perp$, $zero? : \mathbf{Nat}_\perp \to B_\perp$, $cond_\sigma : boole \times \sigma^2 \to$

$$
\begin{array}{rclrcl}
(\lambda x \cdot \varphi)`a & \to & \varphi[a/x] & \mathit{zero?}`(\mathbf{0}) & \to & \mathbf{t} \\
\mathsf{Y}`f & \to & f`(\mathsf{Y}`f) & \mathit{zero?}`(\mathbf{n}+\mathbf{1}) & \to & \mathbf{f} \\
\mathit{succ}`\mathbf{n} & \to & \mathbf{n}+\mathbf{1} & \mathit{cond}\,\mathbf{t}ab & \to & a \\
\mathit{pred}`(\mathbf{n}+\mathbf{1}) & \to & \mathbf{n} & \mathit{cond}\,\mathbf{f}ab & \to & b
\end{array}
$$

$$
\frac{f \to g}{f`a \to g`a} \qquad \frac{f \to g}{u`f \to u`g} \quad \text{where } u \in \{\mathit{succ}, \mathit{pred}, \mathit{zero?}\}
$$

$$
\frac{f \to g}{\mathit{cond}\, f a b \to \mathit{cond}\, g a b} \quad \text{where } \mathit{cond}\, xyz = ((\mathit{cond}`x)`y)`z
$$

Fig. 7. Operational semantics for PCF.

σ, $\sigma \in \{\mathit{nat}, \mathit{boole}\}$ (*cond* is for *conditional*, sometimes called *if then else*), $[\![\mathbf{T}]\!] = \mathbf{t}$, $[\![\mathbf{F}]\!] = \mathbf{f}$, $[\![\mathbf{n}]\!] = n$,

$$
\mathit{succ}(x) = \begin{cases} x+1 & \text{if } x \neq \perp, \\ \perp & \text{if } x = \perp, \end{cases} \qquad \mathit{cond}_\sigma(p, y, z) = \begin{cases} \mathbf{y} & \text{if } p = \mathbf{t}, \\ \mathbf{z} & \text{if } p = \mathbf{f}, \\ \perp & \text{if } p = \perp, \end{cases}
$$

$$
\mathit{pred}(x) = \begin{cases} x-1 & \text{if } x \neq \perp, 0, \\ \perp & \text{else}, \end{cases}
$$

$\perp_\sigma$ is the least element of $[\![\sigma]\!]$ (denoting "non-termination" or "divergent").

$$
\mathit{zero?}(x) = \begin{cases} \mathbf{t} & \text{if } x = 0, \\ \mathbf{f} & \text{if } x \neq 0, \perp, \\ \perp & \text{if } x = \perp, \end{cases}
$$

Y_σ is the least fixed point operator
$\mathsf{Y}_\sigma(f) = \bigvee\{f^n(\perp_\sigma) \mid n \geqslant 0\}$
(see Example 3.5).

More generally, a *standard model of PCF* is an ω-CPO-enriched ccc $\mathcal{C}$, in which each homset $\mathcal{C}(A, B)$ has a smallest element $\perp_{AB} : A \to B$ with the following properties:

(i) pairing and currying are monotonic,

(ii) $\perp \circ f = f$ and $ev \circ \langle \perp, f \rangle = \perp$ for all f of appropriate type,

(iii) there are objects **nat** and **boole** whose sets of global elements satisfy $\mathcal{C}(\mathbf{1}, \mathbf{nat}) \cong \mathbf{N}_\perp$ and $\mathcal{C}(\mathbf{1}, \mathbf{boole}) \cong \mathbf{B}_\perp$ and in which the constants are all interpreted in the internal language of $\mathcal{C}$ as in the standard model above (e.g., interpreting $[\![\mathit{succ}(x)]\!]$ by $ev \circ \langle \mathit{succ}, x \rangle$, etc.).

A model is *order-extensional* if $\mathbf{1}$ is a generator and the order on hom-sets coincides with the pointwise ordering.

The operational semantics of PCF is given by a set of rewriting rules, displayed in Figure 7. This is intended to describe the dynamic evaluation of a term, as a sequence of 1-step transitions. The fixed-point combinator Y guarantees some computations may not terminate. It is important to emphasize that in operational semantics with partially-defined (i.e. possibly non-terminating) computations, different orders of evaluation (e.g., left-most outermost vs innermost) may lead to non-termination in some cases and may also effect efficiency, etc. (see [Mit96], Chapter 2). We have chosen a simple operational semantics for PCF, given by a deterministic evaluation relation, following [AC98].

2.8.2. *Adequacy.* A *PCF program* is a closed term of base type (i.e. either *nat* or *boole*). The *observable behaviour* of a PCF program $P : nat$ is the set $Beh(P) = \{n \in \mathbf{N} \mid P \xrightarrow{*} \mathbf{n}\}$, and similarly for $P : boole$. The set $Beh(P)$ is either empty if P diverges or a singleton if P converges to a (necessarily unique) normal form. The following theorem is proved using a logical relations argument.

THEOREM 2.52 (Computational adequacy). *Let C be any standard model of PCF. Then for all programs $P : nat$ and $n \in \mathbf{N}$,*

$$P \xrightarrow{*} \mathbf{n} \quad \textit{iff} \quad [\![P]\!] = n$$

and similarly for $P : boole$. Hence $[\![P]\!] = [\![Q]\!]$ iff their sets of behaviours are equal.

We are interested in a notion of "observational equivalence" arising from the operational semantics. A program (a closed term of base type) can be observed to converge to a specific numeral or boolean value. More generally, what can we observe about arbitrary terms of any type? The idea is to plug them into arbitrary program code, and observe the behaviour. More precisely, a *program context* $C[-]$ is a program with a hole in it (the hole has a specified type) which is such that if we formally plug a PCF term M into (the hole of) $C[-]$ (we don't care about possible clashes of bound variables) we obtain a program $C[M]$. We are looking at the convergence behaviour (with respect to the operational semantics) of the resulting program (cf. [AC98]).

DEFINITION 2.53. Two PCF terms M, N of the same type are *observationally equivalent* (denoted $M \approx N$) iff $Beh(C[M]) = Beh(C[N])$ for every program context $C[-]$.

That is, $M \approx N$ means that for all program contexts $C[-]$, $C[M] \xrightarrow{*} c$ iff $C[N] \xrightarrow{*} c$ (for c either a Boolean value or a numeral). Thus, by the previous theorem, $M \approx N$ iff $[\![C[M]]\!] = [\![C[N]]\!]$. In order to prove observational equivalence of two PCF terms, R. Milner showed it suffices to pick *applicative contexts*, i.e.

LEMMA 2.54 (Milner). *Two closed PCF expressions*

$$M, N : \sigma_1 \Rightarrow (\sigma_2 \Rightarrow \cdots (\cdots (\sigma_n \Rightarrow nat) \cdots)$$

are observationally equivalent iff $[\![M P_1 \ldots P_n]\!] = [\![N P_1 \ldots P_n]\!]$ for all closed $P_i : \sigma_i$, $1 \leqslant i \leqslant n$.

Finally, the main definition of the subject is:

DEFINITION 2.55 (*Full abstraction*). A model is called *fully abstract* if observational equivalence coincides with denotational equality in the model, i.e. for any two PCF terms M, N

$$M \approx N \quad \text{iff} \quad [\![M]\!] = [\![N]\!].$$

Which models are fully abstract? There are two main theorems, due to Milner and Plotkin. First we introduce the "parallel-or" function $por : \mathbf{B}_\perp \times \mathbf{B}_\perp \to \mathbf{B}_\perp$ on the standard model of PCF:

$$por(a,b) = \begin{cases} \mathbf{t} & \text{if } a = \mathbf{t} \text{ or } b = \mathbf{t}, \\ \mathbf{f} & \text{if } a = b = \mathbf{f}, \\ \perp & \text{else.} \end{cases}$$

THEOREM 2.56 (Plotkin).
- *por is not definable in PCF.*
- *The standard model is not fully abstract.*
- *The standard model is fully abstract for the language PCF + por.*

The proof of the first two parts of the theorem use logical relations ([Sie92,AC98, Gun92]). In 1977, R. Milner proved the following [Mil77]:

THEOREM 2.57 (Milner). *There is a unique (up to isomorphism) fully abstract order-extensional model of PCF.*

Milner's construction is syntactical, so the question became: find a more "mathematical" (i.e. not explicitly syntactical) characterization of the unique fully abstract model. This is related to the Full Completeness Problems discussed in Section 5.2. A satisfying solution to the Full Abstraction Problem for PCF was recently given by S. Abramsky, R. Jagadeesan, and P. Malacaria and also M. Hyland and L. Ong who use various monoidal categories of games. This has recently led to highly active subject of *games semantics* for programming languages (see Section 4.3.2 and the articles mentioned there).

3. Parametricity

What is parametricity in polymorphism? We have already seen such notions as
- Uniformity of algorithms across types.
- Passing types as parameters in programs.

But the problem is that a type like $\forall\alpha(\alpha \Rightarrow \alpha)$, when interpreted in a model as a large product over all types, may contain in Strachey's words unintended *ad hoc* elements. In addition to removing some entities, we may wish to include yet others. For example, should we consider closure of parametric functions under isomorphism of types?

We have already mentioned the idea of types being functors, in Section 2.4.2. Indeed, this suggests an obvious kind of modelling
- Types = functors
- Terms (programs) = natural transformations

all defined over some ccc $\mathcal{C}$. This view of categorical program semantics has had a fruitful history. Reynolds, Oles, and later O'Hearn and Tennent have used functor categories to develop semantics of local variables, block structure, non-interference, etc. in Algol-like languages (see [OHT92,Ten94] and references there).

In the case of polymorphism this is also not such a far-fetched idea. Imagine a term $t : \forall\alpha \cdot \alpha \Rightarrow \alpha$. We know that for each type A, $t[A] : A \Rightarrow A$. Thus, from our Curry–Howard viewpoint, we think of this as an object-indexed family of arrows. Combining this idea with the mild parametricity condition of *naturality* then seems reasonable. In the mid 1980's, Girard gave functor category models of System $\mathcal{F}$ [Gi86]. However to handle the functorial problem of co/contravariance in an expression like $\alpha \Rightarrow \beta$ (or worse, in $\alpha \Rightarrow \alpha$, which is not a functor at all) he introduced categories of embedding-projection pairs (as in domain theory, Section 2.5). Below we shall consider dinaturality, a multivariant notion of naturality which takes into account such problems.

Reynolds [Rey83] also proposed an analysis of parametricity using the notion of *logical relations*, a fundamental tool in the theory of typed lambda calculi. The paper [BFSS90] studied the above two frameworks for parametricity: Reynolds' relational approach and the dinaturality approach. This work was extended and formalized in [ACC93,BAC95, PlAb93].

3.1. *Dinaturality*

One attempt to understand parametric polymorphism is to require certain *naturality* conditions on families interpreting universal types. In this view we begin with some appropriate ccc $\mathcal{C}$ of values and interpret polymorphic type expressions $A(\alpha_1, \ldots, \alpha_n)$, with type variables α_i, as certain kinds of multivariant "definable" functors $F : (\mathcal{C}^{op})^n \times \mathcal{C}^n \to \mathcal{C}$. Terms or programs t are then interpreted as certain multivariant (= *dinatural*) transformations between (the interpretations of) the types. We need to account for naturality not only in positive (covariant) positions, but also in negative (contravariant) ones. As we shall see, the difficulty will be compositionality.

DEFINITION 3.1. Let $\mathcal{C}$ be a category, and $F, G : (\mathcal{C}^{op})^n \times \mathcal{C}^n \to \mathcal{C}$ functors. A *dinatural transformation* $\theta : F \to G$ is a family of $\mathcal{C}$-morphisms $\theta = \{\theta_A : FAA \to GAA \mid A \in \mathcal{C}^n\}$ satisfying (for any n-tuple $f : A \to B \in \mathcal{C}^n$):

$$\begin{tikzcd}
 & FAA \arrow[r, "\theta_A"] & GAA \arrow[dr, "GAf"] & \\
FBA \arrow[ur, "FfA"] \arrow[dr, "FBf"'] & & & GAB \\
 & FBB \arrow[r, "\theta_B"] & GBB \arrow[ur, "GfB"'] &
\end{tikzcd}$$

For a history of this notion, see [Mac71]. Dinatural transformations include ordinary natural transformations as a special case (e.g., construe covariant F, G as bifunctors, dummy in the contravariant variable), as well as transformations between co- and contravariant functors. The parametric aspect of naturality here is that θ_A may be varied along an arbitrary map $f : A \to B$ in both the co- and contravariant positions.

In the following examples, K_A denotes the constant functor with value A (where $K_A(f) = id_A$).

We use set-theoretic notation, but the examples make sense in any ccc (e.g., using the internal language). We follow the treatment in [BFSS90].

EXAMPLE 3.2 (*Polymorphic identity*). Let $F = K_{\mathbf{1}}$, let $G(-,-) = (-)^{(-)}$. Consider the family $I = \{I_A : \mathbf{1} \to A^A \mid A \in \mathcal{C}\}$ where $I_A(*) = \lambda_{x:A} \cdot x = (\mathbf{1} \times A \xrightarrow{\pi_2} A)^*$. Definition 3.1 reduces to the following commuting square:

$$\begin{tikzcd}
 & A^A \arrow[dr, "f^A"] & \\
\mathbf{1} \arrow[ur, "I_A"] \arrow[dr, "I_B"'] & & B^A \\
 & B^B \arrow[ur, "B^f"'] &
\end{tikzcd}$$

which essentially says $f \circ id_A = id_B \circ f$. This equation is true (in **Set** or in the internal language of any ccc) since both sides equal f.

EXAMPLE 3.3 (*Evaluation*). Fix an object $D \in \mathcal{C}$. Let $F(-) = D^{(-)}$ and $G = K_D$. The family $Ev = \{ev_A : (D^A) \times A \to D \mid A \in \mathcal{C}\} : F \to G$ is a dinatural transformation, where ev_A is the usual evaluation in any ccc. Definition 3.1 reduces to the following commuting square, for any $f : A \to B$

$$\begin{tikzcd}
 & D^A \times A \arrow[dr, "ev_A"] & \\
D^B \times A \arrow[ur, "D^f \times A"] \arrow[dr, "D^B \times f"'] & & D \\
 & D^B \times B \arrow[ur, "ev_B"'] &
\end{tikzcd}$$

This says, for any $g : D^B,\ a : A, ev_A(g \circ f, a) = ev_B(g, f(a))$. More informally, $(g \circ f)(a) = g(f(a))$. Again, this is a truism in any ccc.

Extending the above example, *generalized evaluation* $EV = \{ev_{A,A'} : A'^A \times A \to A' \mid A, A' \in \mathcal{C}\}$ determines a dinatural transformation between appropriate functors (cf. [BFSS90]). Dinaturality corresponds to the true equation $f'((g \circ f)(a)) = (f' \circ g)(f(a))$ for $g : A'^B$, $a : A$, and any $f : A \to B,\ f' : A' \to B'$.

EXAMPLE 3.4 (*Church numerals*). Define $\mathbf{n} : (-)^{(-)} \to (-)^{(-)}$ to be the family where $\mathbf{n}_A : A^A \to A^A$ is given by mapping $h \mapsto h^n$, with $h^n = h \circ h \circ h \circ \cdots \circ h$ (n times). Dinaturality corresponds to the diagram (for any $f : A \to B$)

$$\begin{tikzcd}
& A^A \arrow[r, "\mathbf{n}_A"] & A^A \arrow[dr, "f^A"] & \\
A^B \arrow[ur, "A^f"] \arrow[dr, "f^B"'] & & & B^A \\
& B^B \arrow[r, "\mathbf{n}_B"] & B^B \arrow[ur, "B^f"'] &
\end{tikzcd}$$

i.e. if $g : A^B$, $(f \circ g)^n \circ f = f \circ (g \circ f)^n$, an instance of associativity.

We shall see dinaturality again in Section 6.1. Observe that each of the families $\theta = \{\theta_A \mid A \in \mathcal{C}\}$ above – which in essence arise from the syntax of ccc's – have *uniform algorithms* θ_A across all types A. For example, $\mathbf{n}_A = \lambda_{h:A^A} \cdot h^n$, uniform in each type A.

We end with an operator which is fundamental to denotational semantics.

EXAMPLE 3.5 (*Fixedpoint combinators*). In many ccc's $\mathcal{C}$ used in programming language semantics, e.g., certain subcategories of ω-CPO, there is a dinatural fixed point combinator $\mathsf{Y}_{(-)} : (-)^{(-)} \to (-)$. That is, we have a family $\{\mathsf{Y}_A : A^A \to A \mid A \in \mathcal{C}\}$ making the following diagram commute, for any $f : A \to B$:

$$\begin{tikzcd}
& A^A \arrow[r, "\mathsf{Y}_A"] & A \arrow[dr, "f"] & \\
A^B \arrow[ur, "A^f"] \arrow[dr, "f^B"'] & & & B \\
& B^B \arrow[r, "\mathsf{Y}_B"] & B \arrow[ur, "id_B"'] &
\end{tikzcd}$$

This says, using informal set-theoretic notation, if $g : A^B$, $f(\mathsf{Y}_A(g \circ f)) = \mathsf{Y}_B(f \circ g)$. In particular, setting $B = A$ and letting $g = id_B$, we have the fixed-point equation $f(\mathsf{Y}_A(f)) = \mathsf{Y}_A(f)$.

For example, consider ω-CPO$_\perp$ – the subccc of ω-CPO whose objects A have a least element $\perp_A$ but the morphisms need not preserve it. It may be shown that the family given by $\mathsf{Y}_A(f) =$ *the least fixed-point of* $f = \bigvee\{f^n(\perp_A) \mid n \geqslant 0\}$ is dinatural (see [BFSS90, Mul91,Si93]).

There is a calculus of multivariant functors $F, G : (\mathcal{C}^{op})^n \times \mathcal{C}^n \to \mathcal{C}$ functors. For example basic type constructors may be defined (using products and exponentials in $\mathcal{C}$) by setting

$$(F \times G)AB = FAB \times GAB, \tag{11}$$

$$(G^F)AB = GAB^{FBA}. \tag{12}$$

Here A is the list of n contravariant and B the list of n covariant arguments. Note the twist of the arguments in the definition of exponentiation. Much of the structure of Cartesian closedness (e.g., evaluation maps, currying, projections, pairing, etc.) exists within the world of dinatural transformations and there is a kind of abstract functorial calculus (cf. [BFSS90], Appendix A6, [Fre93]).

Unfortunately, there is a serious problem: in general, dinaturals do not compose. That is, given dinatural families $\{FAA \xrightarrow{\theta_A} GAA \mid A \in \mathcal{C}\}$ and $\{GAA \xrightarrow{\psi_A} HAA \mid A \in \mathcal{C}\}$, the composite $\{FAA \xrightarrow{\psi_A \circ \theta_A} HAA \mid A \in \mathcal{C}\}$ does not always make the appropriate hexagon commute. However, with respect to the original question of closure of parametric functions under isomorphisms of types, we note that families dinatural with respect to isomorphisms f do in fact compose. But this class is too weak for a general modelling. Detailed studies of such phenomena have been done in [BFSS90,FRRa,FRRb].

Remarkably, there are certain categories $\mathcal{C}$ over which there are large classes of multivariant functors and dinatural transformation which provide a compositional semantics:

- In [BFSS90] it is shown that if $\mathcal{C} = Per(\mathbf{N})$, that so-called *realizable* dinatural transformations between *realizable* functors compose. Realizable functors include almost any functors that arise in practice (e.g., those definable from the syntax of System $\mathcal{F}$) while realizable dinatural transformations are families of per morphisms whose action is given uniformly by a single Turing machine. This semantics also has a kind of universal quantifier modelling System $\mathcal{F}$(see below).
- In [GSS91] it is shown that the syntax of simply typed lambda calculus with type variables – i.e. $\mathcal{C}$ = a free ccc – admits a compositional dinatural semantics (between logically definable functors and dinatural families). This uses the cut-elimination theorem from proof theory. This work was extended to Linear Logic by R. Blute [Blu93].
- In [BS96] there is a compositional dinatural semantics for the multiplicative fragment of linear logic (generated by atoms). Here $\mathcal{C} = \mathcal{RTVEC}$, a category of reflexive topological vector spaces first studied by Lefschetz [Lef63], with functors being syntactically definable. In [BS96b,BS98] this was extended to a compositional dinatural semantics for Yetter's noncommutative *cyclic* linear logic. In both cases, one demands certain uniformity conditions on dinatural families, involving equivariance w.r.t. continuous group (respectively Hopf-algebra) actions induced from actions on the atoms (see also Section 5.2 below). (For the non-cyclic fragment this is automatic.)

Associated with dinaturality is a kind of "parametric" universal quantifier first described by Yoneda and which plays a fundamental role in modern category theory [Mac71].

DEFINITION 3.6. An *end* of a multivariant functor G on a category $\mathcal{C}$ is an object $E = \int_A GAA$ and a dinatural transformation $K_E \to G$, universal among all such dinatural transformations.

In more elementary terms, there is a family of arrows

$$\left\{ \int_A GAA \xrightarrow{\theta_X} GXX \mid X \in \mathcal{C} \right\}$$

making the main square in the following diagram commute, for any $X \xrightarrow{f} Y$, and such that given any other family $u = \{u_X \mid X \in \mathcal{C}\}$ such that $GXf \circ u_X = GfY \circ u_Y$, there is a unique Au making the appropriate triangles commute:

$$\begin{tikzcd}
D \arrow[drr, "u_X"] \arrow[dddrr, "u_Y"'] \arrow[dd, dotted, "Au"'] & & & \\
& & GXX \arrow[dr, "GXf"] & \\
\int_A GAA \arrow[urr, "\theta_X"] \arrow[drr, "\theta_Y"'] & & & GXY \\
& & GYY \arrow[ur, "GfY"'] &
\end{tikzcd}$$

One may think of $\int_A GAA$ as a subset of $\Pi_A GAA$ (note, this is a "large" product over all $A \in \mathcal{C}$, so $\mathcal{C}$ must have appropriate limits for this to exist)

$$\begin{aligned}\int_A GAA = \{&g \in \Pi_A GAA \mid GXf \circ \theta_X = GfY \circ \theta_Y, \\ &\text{for all } X, Y,\ f : X \to Y \in \mathcal{C}\}.\end{aligned}$$

In [BFSS90] versions of such ends over *Per*(**N**) are discussed with respect to parametric modelling of System $\mathcal{F}$.

In a somewhat different direction, *co-ends* (dual to ends) are a kind of sum or existential quantifier. Their use in categorical computer science was strongly emphasized in early work of Bainbridge [B72,B76] on duality theory for machines in categories. A useful observation is that we may consider functors $R : \mathcal{C} \times \mathcal{D} \to \mathbf{Set}$ and $S : \mathcal{D} \times \mathcal{E} \to \mathbf{Set}$ as generalized relations, with relational composition being determined by the coend formula $R; S(C, E) = \int^D (R(C, D) \times S(D, E))$. This view has recently been applied to relational semantics of dataflow languages in [HPW].

We should mention that dinaturality is also intimately connected with categorical coherence theorems and geometrical properties of proofs [Blu93,So87]. It also seems to be hidden in deeper aspects of Cut-Elimination [GSS91], although here there is still much to understand. We shall meet dinaturality again in several places (e.g., in Traced Monoidal Categories, Section 6.1).

3.2. *Reynolds parametricity*

Reynolds [Rey83] analyzed Strachey's notion of parametric polymorphism using a relational model of types. Although his original idea of using a **Set**-based model was later shown by Reynolds himself to be untenable, the framework has greatly influenced subsequent studies. As a concrete illustration, following [BFSS90] we shall sketch a relational model over *Per*(**N**). Related results in more general frameworks were obtained by Hasegawa [RHas94,RHas95] and Ma and Reynolds [MaRey]. Although originally Reynolds' work was semantical, general logics for reasoning about formal parametricity, supported by such *Per* models, were developed in [BAC95,PlAb93,ACC93].

Given pers $A, A' \in Per(\mathbf{N})$, a *saturated relation* $R : A \nrightarrow A'$ is a relation $R \subset dom_A \times dom_{A'}$ satisfying $R = A; R; A'$, where ; denotes relational composition. For all pers A, A', B, B', saturated relations $R : A \nrightarrow A'$ and $S : B \nrightarrow B'$ and elements $a \in dom_A$, $a' \in dom_{A'}$, $b \in dom_B$, $b' \in dom_{B'}$ we define *a relational* System $\mathcal{F}$ type structure as follows:

- $R \times S : (A \times B) \nrightarrow (A' \times B')$, given componentwise by $(a, b) R \times S(a', b')$ iff aRa' and bSb'.
- $R \Rightarrow S : (A \Rightarrow B) \nrightarrow (A' \Rightarrow B')$, where $f(R \Rightarrow S)g$ iff f, g are (codes of) Turing computable functions satisfying aRa' implies $faSga'$ for any a, a' as above.
- $\forall\alpha \cdot \tau(\alpha, S) : \forall\alpha \cdot \tau(\alpha, B) \nrightarrow \forall\alpha \cdot \tau(\alpha, B')$ is defined by a simultaneous inductive definition based on the formation of type expression $\tau(\alpha, \beta)$. We shall omit the technical construction (see [BFSS90], p. 49) but the key idea is to redefine the *Per*-interpretation of $\forall\alpha \cdot \tau(\alpha)$ by triming down the intersection $\bigcap_A \tau(A)$ to only those elements invariant under all saturated relations, while $\forall\alpha \cdot \tau(\alpha, S) = \bigcap_R \tau(R, S)$, the intersection being over all pers A, A' and saturated $R : A \nrightarrow A'$.

The somewhat involved construction of $\forall\alpha \cdot \tau(\alpha)$ ensures the type expressions $\tau(-)$ act like functors with respect to saturated relations. More precisely, Reynolds' parametricity entails:

- If $R : A \nrightarrow A'$ is a saturated relation, then for any polymorphic type τ, $\tau(R) : \tau(A) \nrightarrow \tau(A')$ is a saturated relation.
- *Identity Extension Lemma:* τ preserves identity relations, i.e. $\tau(id_A) = id_{\tau(A)}$, as saturated relations $\tau(A) \nrightarrow \tau(A)$ (and similarly for $\tau(id_{A_1}, \ldots, id_{A_n})$).

One obtains a Soundness Theorem, essentially the interpretation of the free term model of System $\mathcal{F}$ into the relational *Per* model above. For simplicity, consider terms with only one variable. Let $\sigma = \sigma(\alpha_1, \ldots, \alpha_n)$ and $\tau = \tau(\alpha_1, \ldots, \alpha_n)$ denote polymorphic types with free type variables $\subseteq \{\alpha_1, \ldots, \alpha_n\}$. Let $x : \sigma \vdash t : \tau$ denote term $t : \sigma$ with free variable $x : \sigma$. Associated to every System $\mathcal{F}$ term t is a Turing computable numerical function f_t, obtained by essentially erasing all types and considering the result as an untyped lambda term, qua computable partial function (see [BFSS90], Appendix A.1).

THEOREM 3.7 (Soundness). *Let A_i, A'_i be pers and $R_i : A_i \nrightarrow A'_i$ saturated relations. Then if $m\sigma(\vec{R})m'$ then $f_t(m)\tau(\vec{R})f_t(m')$. Also, if $t = t'$ in System $\mathcal{F}$, then $f_t = f_{t'}$ as Per*(**N**) *maps $\sigma(\vec{A}) \to \tau(\vec{A})$.*

Thus terms (programs) become "relation transformers" $\sigma(\vec{R}) \to \tau(\vec{R})$ (cf. [MitSce, Fre93]) of the form

$$\begin{tikzcd}
 & \sigma(\vec{A}) \arrow[rr, "f_t"] & & \tau(\vec{A}) \\
\sigma(\vec{R}) \arrow[ur] \arrow[dr] \arrow[rr, dotted, "\exists"] & & \tau(\vec{R}) \arrow[ur] \arrow[dr] & \\
 & \sigma(\vec{A}') \arrow[rr, "f_t"'] & & \tau(\vec{A}')
\end{tikzcd}$$

In particular this exemplifies Reynolds' interpretation of Strachey's parametricity: if one instantiates an element of polymorphic type at two related types, then the two values obtained must be related themselves.

Reynolds parametricity has the following interesting consequence [BFSS90,RHas94]. Recall the category of T-algebras, for definable functors T (cf. the proof of Theorem 2.40).

THEOREM 3.8. *Let T be a System $\mathcal{F}$-definable covariant functor. Then in the parametric Per model, $\forall\alpha \cdot ((T\alpha \Rightarrow \alpha) \Rightarrow \alpha)$ is the initial T-algebra.*

This property becomes a general theorem in the formal logics of parametricity (e.g., [PlAb93,RHas95]), and hence would be true in any appropriate parametric model. Thus, although the syntax of second-order logic in general only guarantees *weakly initial* data types as in [GLT], in parametric models of System $\mathcal{F}$ the usual definitions actually yield strong data types.

The reader might rightly enquire: do relational parametricity and dinaturality have anything in common? This is exactly the kind of question that requires a logic for reasoning *about* parametricity. Plotkin and Abadi's logic [PlAb93] extends the equational theory of System $\mathcal{F}$ with quantification over functions and relations, together with a schema expressing Reynolds' relational parametricity. The dinaturality hexagon in Definition 3.1, for definable functors and families, is expressible as a quantified equation in this logic.

PROPOSITION 3.9. *In the formal system above, relational parametricity implies dinaturality.*

Reynolds' work on parametricity continues to inspire fundamental research directions in programming language theory, even beyond polymorphism. For example, O'Hearn and Tennent [OHT92,OHT93] use relational parametricity to examine difficult problems in local-variable declarations in Algol-like languages. Their framework is particularly interesting. They use ccc's of functor categories and natural transformations, à la Oles and Reynolds, but *internal to the category of reflexive directed multigraphs*. The same framework, somewhat generalized, was then used by A. Pitts [Pi96a] in a general relational approach to reasoning about properties of recursively defined domains. Pitts work has led to new approaches to induction and co-induction, etc. (see Section 2.5). The reader is referred to Pitts [Pi96a], p. 74, and O'Hearn and Tennent [OHT93] for many examples of these so-called *relational structures over categories* $\mathcal{C}$.

4. Linear logic

4.1. *Monoidal categories*

We briefly recall the relevant definitions. For details, the reader is referred to [Mac71, Bor94].

DEFINITION 4.1. A *monoidal category* is a category $\mathcal{C}$ equipped with a functor $\otimes : \mathcal{C} \times \mathcal{C} \to \mathcal{C}$, an object I, and specified natural isomorphisms:

$$a_{ABC} : (A \otimes B) \otimes C \xrightarrow{\cong} A \otimes (B \otimes C)$$

$$\ell_A : I \otimes A \xrightarrow{\cong} A \quad \text{and} \quad r_A : A \otimes I \xrightarrow{\cong} A$$

satisfying coherence equations: associativity coherence (Mac Lane's Pentagon) and the unit coherence.

A *symmetric monoidal category* is a monoidal category with a natural symmetry isomorphism $s_{AB} : A \otimes B \xrightarrow{\cong} B \otimes A$ satisfying: $s_{B,A}s_{A,B} = id_{A\otimes B}$, for all A, B, and (omitting subscripts) $r = ls$, $asa = (1 \otimes s)a(s \otimes 1)$.

Symmetric monoidal categories include Cartesian categories (with $\otimes = \times$) and co-Cartesian categories (with $\otimes = +$). However in the two latter cases, the structure is uniquely determined (up to isomorphism) – and similarly for the coherence isomorphisms – by the universal property of products (resp. coproducts). This is not true in the general case – there may be many symmetric monoidal structures on the same category.

We now introduce the monoidal analog of ccc's:

DEFINITION 4.2. A *symmetric monoidal closed category* (= smcc) $(\mathcal{C}, \otimes, I, \multimap)$ is a symmetric monoidal category such that for each object $A \in \mathcal{C}$, the functor $- \otimes A : \mathcal{C} \to \mathcal{C}$ has a specified right adjoint $A \multimap -$, i.e. for each A there is an isomorphism, natural in B, C:

$$Hom_{\mathcal{C}}(C \otimes A, B) \cong Hom_{\mathcal{C}}(C, A \multimap B). \tag{13}$$

As a consequence, in any smcc there are "evaluation" and "coevaluation" maps $(A \multimap B) \otimes A \xrightarrow{ev_{AB}} B$ and $C \longrightarrow (A \multimap (C \otimes A))$ determined by the adjointness (13). We shall try to keep close to our ccc notation, Section 2.1. In particular the analog of Currying arising from (13) is denoted $C \xrightarrow{f^*} (A \multimap B)$. Moreover, this data actually determines a (bi)functor $- \multimap - : \mathcal{C}^{op} \times \mathcal{C} \to \mathcal{C}$. No special coherences have to be supposed for $- \multimap -$: they follow from coherence for $\otimes$ and adjointness.

For the purposes of studying linear logic below, we need (among other things) a notion of an smcc, equipped with an involutive negation or duality, reminiscent of finite dimensional vector spaces. The general theory of such categories, due to M. Barr [Barr79], was developed in the mid 1970's, some ten years before linear logic.

Consider an smcc $\mathcal{C}$, with a distinguished object $\perp$. Consider the map $ev_{A\perp} \circ s_{A\perp} : A \otimes (A \multimap \perp) \to \perp$. By (13) this corresponds to a map $\mu_A : A \to (A \multimap \perp) \multimap \perp$. Let us write $A^\perp$ for $A \multimap \perp$. Thus we have a morphism $\mu_A : A \to A^{\perp\perp}$. Objects A for which μ_A is an isomorphism are called *reflexive*, or more precisely reflexive with respect to $\perp$.

DEFINITION 4.3. A *$*$-autonomous category* $(\mathcal{C}, \otimes, I, \multimap, \perp)$ is an smcc $\mathcal{C}$ with a distinguished object $\perp$ such that all objects are reflexive, i.e. the canonical map $\mu_A : A \to A^{\perp\perp}$ is an isomorphism for all $A \in \mathcal{C}$. The object $\perp$ is called the *dualizing object*.

It may be shown that a $*$-autonomous category $\mathcal{C}$ has a contravariant dualizing functor $(-)^\perp : \mathcal{C}^{op} \to \mathcal{C}$, defined on objects by $A \mapsto A^\perp$. There is a natural isomorphism: $Hom_{\mathcal{C}}(A, B) \cong Hom_{\mathcal{C}}(B^\perp, A^\perp)$.

In any $*$-autonomous category $\mathcal{C}$ there are isomorphisms

$$\begin{aligned} (A \multimap B)^\perp &\cong A \otimes B^\perp, \\ I &\cong \perp^\perp. \end{aligned}$$

The reader is referred to [Barr79] for many examples. Let us mention the obvious one:

EXAMPLE 4.4. The category Vec_{fd} of finite-dimensional vector spaces over a field $\mathbf{k}$ is $*$-autonomous. Here $A \multimap B = Lin(A, B)$, the space of linear maps from A to B and the dualizing object $\perp = \mathbf{k}$. In particular $A^\perp = A^*$ is the usual dual space. More generally, within the smcc category Vec of $\mathbf{k}$-vector spaces (with $\perp = \mathbf{k}$), an object is reflexive iff it is finite-dimensional.

In a $*$-autonomous category, we may define the *cotensor* $\parr$ by de Morgan duality: $A \parr B = (A^\perp \otimes B^\perp)^\perp$. The above example Vec_{fd} is somewhat "degenerate" since $\otimes$ and $\parr$ are identified (see Definition 4.10 of *compact* category). In a typical $*$-autonomous category this is not the case; indeed in linear logic one does not want to identify tensor and cotensor.

To obtain more general $*$-autonomous categories of vector spaces, we add a topological structure, due to Lefschetz [Lef63]. The following discussion is primarily based on work of M. Barr [Barr79], following the treatment in Blute [Blu96].

DEFINITION 4.5. Let V be a vector space. A topology τ on V is *linear* if it satisfies the following three properties:

- Addition and scalar multiplication are continuous, when the field $\mathbf{k}$ is given the discrete topology.
- τ is Hausdorff.
- $0 \in V$ has a neighborhood basis of open linear subspaces.

Let $\mathcal{TVEC}$ denote the category whose objects are vector spaces equipped with linear topologies, and whose morphisms are linear continuous maps.

Barr showed that $\mathcal{TVEC}$ is a symmetric monoidal closed category, when $V \multimap W$ is defined to be the vector space of linear continuous maps, topologized with the topology of pointwise convergence. (It is shown in [Barr96] that the forgetful functor $\mathcal{TVEC} \to \mathcal{VEC}$ is tensor-preserving.) Let $V^{\perp}$ denote $V \multimap \mathbf{k}$. Lefschetz proved that the embedding $V \to V^{\perp\perp}$ is always a bijection, but need not be an isomorphism. This is analogous to Dana Scott's method of solving domain equations in denotational semantics, using the topology to cut down the size of the function spaces.

THEOREM 4.6 (Barr). *$\mathcal{RTVEC}$, the full subcategory of reflexive objects in $\mathcal{TVEC}$, is a complete, cocomplete $*$-autonomous category, with $I^{\perp} = I = \mathbf{k}$ the dualizing object.*

Moreover, in $\mathcal{RTVEC}$, $\otimes$ and $⅋$ are not equated. More generally, other classes of $*$-autonomous categories arise by taking a linear analog of G-sets, namely categories of group representations.

DEFINITION 4.7. Let G be a group. A *continuous G-module* is a linear action of G on a space V in $\mathcal{TVEC}$, such that for all $g \in G$, the induced map $g \cdot (\ ) : V \to V$ is continuous. Let $\mathcal{TMOD}(G)$ denote the category of continuous G-modules and continuous equivariant maps. Let $\mathcal{RTMOD}(G)$ denote the full subcategory of reflexive objects.

We have the following result, which in fact holds in the more general context of Hopf algebras (see below).

THEOREM 4.8. *The category $\mathcal{TMOD}(G)$ is symmetric monoidal closed. The category $\mathcal{RTMOD}(G)$ is $*$-autonomous, and a reflective subcategory of $\mathcal{TMOD}(G)$ via the functor $(\ )^{\perp\perp}$. Furthermore the forgetful functor $|\ | : \mathcal{RTMOD}(G) \to \mathcal{RTVEC}$ preserves the $*$-autonomous structure.*

Still more general classes of $*$-autonomous categories may be obtained from categories of modules of cocommutative Hopf algebras. Given a Hopf algebra H, a *module* over H is a linear action $\rho : \mathsf{H} \otimes V \to V$ satisfying the appropriate diagrams, analogous to the notion of G-module. Let $\mathcal{MOD}(\mathsf{H})$ denote the category of H-modules and equivariant maps. Similarly, $\mathcal{TMOD}(\mathsf{H})$, the category of continuous H-modules, is the linearly topologized version of $\mathcal{MOD}(\mathsf{H})$ where H is given the discrete topology and all vector spaces and maps are in $\mathcal{TVEC}$.

PROPOSITION 4.9. *If H is a cocommutative Hopf algebra, $\mathcal{MOD}(\mathsf{H})$ and $\mathcal{TMOD}(\mathsf{H})$ are symmetric monoidal categories.*

We then obtain precisely the same statement as Theorem 4.8 by replacing the group G by a cocommutative Hopf algebra H. Later we shall mention noncommutative Hopf algebra models for linear logic, with respect to full completeness theorems, Section 5.2.

The case where we do identify $\otimes$ and $⅋$ is an important class of monoidal categories:

DEFINITION 4.10. A $*$-autonomous category is *compact* if $(A \otimes B)^{\perp} \cong A^{\perp} \otimes B^{\perp}$ (i.e. equivalently if $A \multimap B \cong A^{\perp} \otimes B$).

In addition to the obvious example of Vec_{fd}, there are compact categories of relations, which have considerable importance in computer science. One such is:

EXAMPLE 4.11. $\mathbf{Rel}_{\times}$ is the category whose objects are sets and whose maps $R : X \to Y$ are (binary) relations $R \subseteq X \times Y$. Composition is relational composition, etc. This is a compact category, with $X \otimes Y = X \multimap Y =_{def} X \times Y$, the ordinary set-theoretic Cartesian product. Define $\perp = \{*\}$, a one-element set; hence $X^{\perp} = X \multimap \perp = X$. On maps we have $R^{\perp} = R^{op}$, where $yR^{op}x$ iff xRy.

Given two smcc's $\mathcal{C}$ and $\mathcal{D}$ (we shall not distinguish the structure) what are the morphisms between them?

DEFINITION 4.12. A *symmetric monoidal functor* is a functor $F : \mathcal{C} \to \mathcal{D}$ together with two natural transformations $m_I : I \to F(I)$ and $m_{UV} : F(U) \otimes F(V) \to F(U \otimes V)$ such that three coherence diagrams commute. In the case of the closed structure, we can define another natural transformation $\hat{m}_{UV} : F(U \multimap V) \to (FU \multimap FV)$ by $\hat{m}_{UV} = (F(ev_{U \multimap V,U}) \circ m_{U \multimap V,U})^{*}$. A symmetric monoidal functor is *strong* (resp. *strict*) if m_I and m_{UV} are natural isomorphisms (resp. identities) for all U, V. A symmetric monoidal functor is *strong closed* (resp. *strict closed*) if m_I and $\hat{m}_{UV}$ are natural isomorphisms (resp. identities) for all U, V. Similarly, one defines $*$*-autonomous* functors.

Finally, we need an appropriate notion of natural transformation for monoidal functors.

DEFINITION 4.13. A natural transformation $\alpha : F \to G$ is *monoidal* if it is compatible with both m_I and m_{UV}, for all U, V, in the sense that the following equations hold:

$$\begin{aligned} \alpha_I \circ m_I &= m_I, \\ m_{UV} \circ (\alpha_U \otimes \alpha_V) &= \alpha_{U \otimes V} \circ m_{UV}. \end{aligned}$$

4.2. *Gentzen's proof theory*

Gentzen's proof theory [GLT], especially his *sequent calculi* and his fundamental theorem on Cut-Elimination, have had a profound influence not only in logic, but in category theory and computer science as well.

In the case of category theory, J. Lambek [L68,L69] introduced Gentzen's techniques to study *coherence theorems* in various free monoidal and residuated categories. This logical approach to coherence for such categories was greatly extended by G. Mints, S. Soloviev, B. Jay et al. [Min81,So87,So95,J90]. For a comparison of Mints' work with more traditional Kelly–Mac Lane coherence theory see [Mac82]. More recently, coherence for large classes of structured monoidal categories arising in linear logic has been established in a series of papers by Blute, Cockett, Seely et al. This is based on Girard's extensions of Gentzen's methods. (See [BCST96,BCS96,BCS97,CS91,CS96b].)

Recent coherence theorems of Gordon–Power–Street, Joyal–Street et al. [GPS96,JS91, JS93] have made extensive use of higher dimensional category theory techniques and Yoneda methods, rather than logical methods. Related Yoneda techniques are now being introduced, in the reverse direction into proof theory, as we outlined in Section 2.7 above.

In computer science, entire research areas: proof search (AI, Logic Programming), operational semantics, type inference algorithms, logical frameworks, etc. are a testimonial to Gentzen's work. Indeed, Gentzen's Natural Deduction and Sequent Calculi are fundamental methodological as well as mathematical tools.

A profound and exciting analysis of Gentzen's work has arisen recently in the rapidly growing area of Linear Logic (= LL), developed by J.-Y. Girard in 1986. While classical logic is about universal truth, and intuitionistic logic is about constructive proofs, LL is a logic of resources and their management and reuse (e.g., see [Gi87,Gi89,GLR95,Sc93, Sc95,Tr92]).

4.2.1. *Gentzen sequents.* Gentzen's analysis of Hilbert's proof theory begins with a fundamental reformulation of the syntax. We follow the presentation in [GLT].

A *sequent* for a logical language $\mathcal{L}$ is an expression

$$A_1, A_2, \ldots, A_m \vdash B_1, B_2, \ldots, B_n, \tag{14}$$

where $A_1, A_2, \ldots, A_m$ and $B_1, B_2, \ldots, B_n$ are finite lists (possibly empty) of formulas of $\mathcal{L}$. Sequents are denoted $\Gamma \vdash \Delta$, for lists of formulas Γ and Δ. Gentzen introduced formal rules for generating sequents, the so-called *derivable* sequents. Gentzen's rules analyze the deep structure and implicit symmetries hidden in logical syntax. Computation in this setting arises from one of two methods:

- The traditional method is Gentzen's Cut-Elimination Algorithm, which allows us to take a formal sequent calculus proof and reduce it to cut-free form. This is closely related to both normalization of lambda terms (cf. Sections 2.7) as well as the operational semantics of such programming languages as PROLOG.
- More recent is the proof search paradigm, which is the bottom-up, goal-directed view of building sequent proofs and is the basis of the discipline of Logic Programming [MNPS, HM94,Mill].

Categorically, the cut elimination algorithm is at the heart of the proof-theoretic approach to coherence theorems previously mentioned. On the other hand, Logic Programming and the proof-search paradigm have only recently attracted the attention of categorists (cf. [FiFrL,PK96]).

Lambek pointed out that Gentzen's sequent calculus was analogous to Bourbaki's method of bilinear maps. For example, given sequences $\Gamma = A_1 \cdots A_m$ and $\Delta = B_1 B_2 \cdots B_n$ of R-R bimodules of a given ring R, there is a natural isomorphism

$$\mathit{Mult}(\Gamma A B \Delta, C) \cong \mathit{Mult}(\Gamma A \otimes B \Delta, D) \tag{15}$$

between $m+n+2$-linear and $m+n+1$-linear maps. Bourbaki derived many aspects of tensor products just from this universal property. Such a formal bijection is at the heart of Linear Logic (e.g., [L93]).

Traditional logicians think of sequent (14) as saying: *the conjunction of the* A_i *entails the disjunction of the* B_j. More generally, following Lambek and Lawvere (cf. Section 2.1), categorists interpret such sequents (modulo equivalence of proofs) as *arrows* in appropriate categories. For example, in the case of logics similar to linear logic [CS91], the sequent (14) determines an arrow of the form

$$A_1 \otimes A_2 \otimes \cdots \otimes A_m \longrightarrow B_1 ⅋ B_2 \cdots ⅋ B_n \tag{16}$$

in a symmetric monoidal category with a "cotensor" ⅋ (see Section 4.1 above).

4.2.2. *Girard's analysis of the structural rules.* Gentzen broke down the manipulations of logic into two classes of rules applied to sequents: *structural rules* and *logical rules*. All rules come in pairs (left/right) applying to the left (resp. right) side of a sequent.

GENTZEN'S STRUCTURAL RULES (LEFT/RIGHT)

Permutation	$\frac{\Gamma\vdash\Delta}{\sigma(\Gamma)\vdash\Delta}$	$\frac{\Gamma\vdash\Delta}{\Gamma\vdash\tau(\Delta)}$	σ, τ permutations.
Contraction	$\frac{\Gamma,A,A\vdash\Delta}{\Gamma,A\vdash\Delta}$	$\frac{\Gamma\vdash\Delta,B,B}{\Gamma\vdash\Delta,B}$.	
Weakening	$\frac{\Gamma\vdash\Delta}{\Gamma,A\vdash\Delta}$	$\frac{\Gamma\vdash\Delta}{\Gamma\vdash\Delta,B}$.	

For simplicity, consider *intuitionistic* sequents, i.e. those of the form $A_1, A_2, \ldots, A_m \vdash B$ with one conclusion. So the right rules disappear and we consider the *left* rules above. We can give a Curry–Howard-style analysis to Gentzen's intuitionistic sequents (cf. Section 2.3), assigning lambda terms (qua functions) to sequents, e.g., $x_1 : A_1, \ldots, x_m : A_m \vdash t(\vec{x}) : B$. The structural rules say the following: *Permutation* says that the class of functions is closed under permutations of arguments; *Contraction* says that the class of functions is closed under duplicating arguments – i.e. setting two input variables equal; and *Weakening* says the class of functions is closed under adding dummy arguments. In the absence of such rules, we obtain the so called *linear* lambda terms, terms where all variables occur exactly once (see [GSS91,Abr93,L89]).

By removing these traditional structural rules, logic takes on a completely different character[3] (see Figure 8). Previously equivalent notions now split into subtle variants based on resource allocation. For example, the rules for *Multiplicative* connectives simply concatenate their input hypotheses Γ and Γ', whereas the rules for *Additive* connectives merge two input hypotheses Γ into one. The situation is analogous for outputs Δ and Δ' (see Figure 8). The resultant logical connectives can represent linguistic distinctions related to resource use which are simply impossible to formulate in traditional logic (see [Gi86, Abr93,Sc93,Sc95]).

[3] Formulas of LL are generated from literals $p, q, r, \ldots, p^\perp, q^\perp, r^\perp, \ldots$ and constants $I, \perp, \mathbf{1}, \mathbf{0}$ using binary operations $\otimes, ⅋, \times, \oplus$ and unary $!, ?$. Negation $(-)^\perp$ is defined inductively: $I^\perp = \perp$, $\perp^\perp = I$, $\mathbf{1}^\perp = \mathbf{0}$, $\mathbf{0}^\perp = \mathbf{1}$, $p^{\perp\perp} = p$, $(A \otimes B)^\perp = A^\perp ⅋ B^\perp$, $(A ⅋ B)^\perp = A^\perp \otimes B^\perp$, $(A \times B)^\perp = (A^\perp \oplus B^\perp)$, $(A \oplus B)^\perp = A^\perp \times B^\perp$, $(!A)^\perp = ?(A^\perp)$, $(?A)^\perp = !(A^\perp)$. Also we define $A \multimap B = A^\perp ⅋ B$.

Structural	*Perm*	$\frac{\Gamma\vdash\Delta}{\sigma(\Gamma)\vdash\tau(\Delta)}$	σ, τ permutations.
Axiom&Cut	*Axiom*	$A\vdash A$	
	Cut	$\frac{\Gamma\vdash A,\Delta \quad \Gamma',A\vdash\Delta'}{\Gamma,\Gamma'\vdash\Delta,\Delta'}$	
Negation		$\frac{\Gamma\vdash A,\Delta}{\Gamma,A^{\perp}\vdash\Delta}$	$\frac{\Gamma,A\vdash\Delta}{\Gamma\vdash A^{\perp},\Delta}$
Multiplicatives	*Tensor*	$\frac{\Gamma,A,B\vdash\Delta}{\Gamma,A\otimes B\vdash\Delta}$	$\frac{\Gamma\vdash A,\Delta \quad \Gamma'\vdash B,\Delta'}{\Gamma,\Gamma'\vdash A\otimes B,\Delta,\Delta'}$
	Par	$\frac{\Gamma,A\vdash\Delta \quad \Gamma',B\vdash\Delta'}{\Gamma,\Gamma',A\parr B\vdash\Delta,\Delta'}$	$\frac{\Gamma\vdash A,B,\Delta}{\Gamma\vdash A\parr B,\Delta}$
	Units	$\frac{\Gamma\vdash\Delta}{\Gamma,I\vdash\Delta}$	$\vdash I$
		$\bot\vdash$	$\frac{\Gamma\vdash\Delta}{\Gamma\vdash\bot,\Delta}$
	Implication	$\frac{\Gamma\vdash A,\Delta \quad \Gamma',B\vdash\Delta'}{\Gamma,\Gamma',A\multimap B\vdash\Delta,\Delta'}$	$\frac{\Gamma,A\vdash B,\Delta}{\Gamma\vdash A\multimap B,\Delta}$
Additives	*Product*	$\frac{\Gamma,A\vdash\Delta}{\Gamma,A\times B\vdash\Delta} \quad \frac{\Gamma,B\vdash\Delta}{\Gamma,A\times B\vdash\Delta}$	$\frac{\Gamma\vdash A,\Delta \quad \Gamma\vdash B,\Delta}{\Gamma\vdash A\times B,\Delta}$
	Coproduct	$\frac{\Gamma,A\vdash\Delta \quad \Gamma,B\vdash\Delta}{\Gamma,A+B\vdash\Delta}$	$\frac{\Gamma\vdash A,\Delta}{\Gamma\vdash A+B,\Delta} \quad \frac{\Gamma\vdash B,\Delta}{\Gamma\vdash A+B,\Delta}$
	Units	$\Gamma,\mathbf{0}\vdash\Delta$	$\Gamma\vdash\mathbf{1},\Delta$
Exponentials	*Weakening*	$\frac{\Gamma\vdash\Delta}{\Gamma,!A\vdash\Delta}$	*Contraction* $\frac{\Gamma,!A,!A\vdash\Delta}{\Gamma,!A\vdash\Delta}$
	Storage	$\frac{!\Gamma\vdash A}{!\Gamma\vdash !A}$	*Dereliction* $\frac{\Gamma,A\vdash\Delta}{\Gamma,!A\vdash\Delta}$

Fig. 8. Rules for classical propositional LL.

REMARK 4.14. First we should remark that on the controversial subject of notation in LL, we have chosen a reasonable categorical notation, somewhere between [Gi87] and [See89]. Observe that in CLL, two-sided sequents can be replaced by one-sided sequents, since $\Gamma\vdash\Delta$ is equivalent to $\vdash\Gamma^{\perp},\Delta$, with $\Gamma^{\perp}$ the list $A_1^{\perp},\ldots,A_n^{\perp}$, where Γ is $A_1,\ldots,A_n$.

Thus the key aspect of linear logic proofs is their resource sensitivity. We think of a linear entailment $A_1,\ldots,A_m\vdash B$ not as an ordinary function, but as an *action* – a kind of process that in a single step consumes the inputs A_i and produces output B. For example, this permits representing in a natural manner the step-by-step behaviour of various abstract machines, certain models of concurrency like Petri Nets, etc. Thus, linear logic permits us to describe the instantaneous state of a system, and its step-wise evolution, intrinsically within the logic itself (e.g., with no need for explicit time parameters, etc.).

But linear logic is not about simply removing Gentzen's structural rules, but rather modulating their use. To this end, Girard introduces a new connective $!A$, which indicates that contraction and weakening may be applied to formula A. This yields the *Exponential* connectives in Figure 8. From a resource viewpoint, an hypothesis $!A$ is one which can be reused arbitrarily. Moreover, this permits decomposing "$\Rightarrow$" (categorically, the ccc function space) into more basic notions:

$$A \Rightarrow B = (!A) \multimap B.$$

Finally, nothing is lost: classical (as well as intuitionistic) logic can be faithfully translated into CLL ([Gi87,Tr92]).

4.2.3. *Fragments and exotic extensions.* The richness of LL permits many natural subtheories (cf. [Gi87,Gi95a]). For a survey of the surprisingly intricate complexity-theoretic structure of many of the fragments of LL see Lincoln [Li95]. These results often involve direct and natural simulation of various kinds of abstract computing machines within the logic [Sc95,Ka95]. Of course, there are specific fragments corresponding to various subcategories of categorical models, in the next section. There are also fragments directly connected with classifying complexity classes in computing [GSS92,Gi97] but these latter have not been the object of categorical analysis.

More exotic "noncommutative" fragments of LL are obtained by eliminating or modifying the permutation rule; i.e. one no longer assumes $\otimes$ is symmetric. One such precursor to LL is the work of J. Lambek in the 1950's on categorial grammars in mathematical linguistics (for recent surveys, see [L93,L95]). Here the language becomes yet more involved, since there are two implications $\multimapinv$ and $\multimap$ and two negations $A^{\perp}$ and ${}^{\perp}A$. It has been proven by Pentus [Pen93] that Lambek grammars are equivalent to context-free grammars. In [L89] there is a formulation of Lambek grammars using the notion of multicategory, an idea currently of some interest in higher-dimensional category theory and higher dimensional rewriting theory [HMP98].

D. Yetter [Y90] considered *cyclic* linear logic, a version of LL in which the permutation rule is modified to only allow *cyclic* permutations. This will be discussed briefly below in Section 5.2 with respect to Full Completeness. A proposed classification of different fragments of LL, including braided versions, based on Hopf-algebraic models is in Blute [Blu96], see also Section 5.2.

4.2.4. *Topology of proofs.* Let us briefly mention one of the main novelties of linear logic. Traditional Gentzen proof theory writes proofs as trees. In order to give a Curry–Howard isomorphism to arbitrary sequents $\Gamma \vdash \Delta$, Girard introduced multiple-conclusion graphical networks to interpret proofs. These *proof nets* use graph rewriting moves for their operational semantics. It is here that one sees the dynamic aspects of cut-elimination. In essense these networks are the "lambda terms" of linear logic. There are known mathematical criteria to classify which (among arbitrary) networks arise from Gentzen sequent proofs, i.e. in a sense which of these parallel networks are "sequentializable" into a Gentzen proof tree. Homological aspects of proof nets are studied in [Mét94]

The technology of proof nets has grown into an intricate subject. In addition to their uses in linear logic, proof nets are now used in category theory, as a technical tool in graphical approaches to coherence theorems for structured monoidal categories (e.g., [Blu93, BCST96,CS91,CS96b]). There are proof net theories for numerous non-commutative, cyclic, and braided linear logics, e.g., [Abru91,Blu93,Fl96,FR94,Y90].

The method of proof nets has been extended by Y. Lafont [Laf95] to a general graphical language of computation, his *interaction nets*. These latter provide a simple model of parallel computation with, at the same time, new insights into sequential computation.

4.3. *What is a categorical model of LL?*

4.3.1. *General models.* As in Section 2.1, we are interested in finding the categories appropriate to modelling linear logic proofs (just as Cartesian closed categories modelled intuitionistic $\wedge, \Rightarrow, \top$ proofs). The basic equations we postulate arise from the operational semantics – that is *normalization of proofs*. In the case of sequent calculi, this is Gentzen's Cut-Elimination process [GLT]. However, there are sometimes natural categorical equations which are not decided by traditional proof theory. The problem is further compounded in linear logic (and monoidal categories) in that there may be several (non-canonical) candidates for appropriate monoidal structure.

The first categorical semantics of LL is in Seely's paper [See89], which is still perhaps the most readable account. Subsequent development of appropriate term calculi [Abr93, Bie95,BBPH,W94,BCST96,CS91,CS96b] have modified and enlarged the scope, but not essentially changed the original analysis for the case of classical linear logic (= CLL). We impose the following equations between CLL proofs, in order to form a category $\mathcal{C}$, where sequents are interpreted as (equivalence classes of) arrows according to formula (16), based on the rules in Figure 8.

- $\mathcal{C}$ is a *symmetric monoidal closed category with products, coproducts, and units* (from the rules: Axiom, Cut, Perm, the Multiplicatives and the Additives).
- $\mathcal{C}$ is $*$-*autonomous* (from the Negation rule) with $\otimes$ and $\invamp$ related by de Morgan duality.
- $! : \mathcal{C} \to \mathcal{C}$ is an endofunctor, with associated monoidal transformations $\varepsilon : ! \to id_{\mathcal{C}}$ and $\delta : ! \to !!$ satisfying:
  (1) $(!, \delta, \varepsilon)$ forms a *monoidal* comonad on $\mathcal{C}$.
  (2) There are natural isomorphisms
  $$I \cong !\mathbf{1} \quad \text{and} \quad !A \otimes !B \cong !(A \times B).$$
  making $! : (\mathcal{C}, \times, \mathbf{1}) \to (\mathcal{C}, \otimes, I)$ a symmetric monoidal functor.
  (3) In particular, $I \xleftarrow{e_A} !A \xrightarrow{d_A} !A \otimes !A$ is a cocommutative comonoid, for all A in $\mathcal{C}$ and the coalgebra maps $\varepsilon_A : !A \to A$ and $\delta_A : !A \to !!A$ are comonoid maps. In fact, these conditions are a consequence of (2), but are required explicitly in weaker settings.

For modifications appropriate to more general situations (e.g., various fragments of LL without products, linearly distributive categories, etc.) see [Bie95,BCS96].

The essence of Girard's translation of intuitionistic logic into LL is the following easy result (cf. [See89,Bie95]).

PROPOSITION 4.15. *If $\mathcal{C}$ is a categorical model of CLL, as above, then the Kleisli category $\mathcal{K}_{\mathcal{C}}$ of the comonad $(!, \delta, \varepsilon)$ is a ccc. Moreover finite products in $\mathcal{K}_{\mathcal{C}}$ and $\mathcal{C}$ coincide, while exponentials in $\mathcal{K}_{\mathcal{C}}$ are given by: $A \Rightarrow B = (!A) \multimap B$.*

We should mention one formal rule, MIX, which appears frequently in the literature. To express it, we use one-sided sequents:

$$Mix\frac{\vdash \Gamma \quad \vdash \Delta}{\vdash \Gamma, \Delta}.$$

Categorically, MIX entails there is a map $A \otimes B \to A \parr B$. This rule seems to be valid in most models, certainly so in ones based on $\mathcal{RTVEC}$.

REMARK 4.16. The categorical comonad approach to models of linear logic has been put to use by Asperti in clarifying optimal graph reduction techniques in the untyped lambda calculus [Asp] (see also [GAL92]).

4.3.2. *Concrete models.* There are by now many categorical models of LL and its interesting fragments. Let us just mention a few [Gi95a,Tr92]:

- *Posetal Models* or *Girard's Phase semantics.* These are $*$-autonomous posets with additional structure. This gives an algebraic semantics analogous to Boolean or Heyting algebras for classical (resp. intuitionistic) logic. As categories they are trivial (each hom set has at most one element); hence they do not model proofs but rather provability. There is associated a traditional Tarski semantics, with Soundness and Completeness Theorems. Recently, these models have been applied in Linear Concurrent Constraint Programming, for proving "safety" properties of programs [FRS98].
- *Domain-Theoretic Models.* The category LIN = coherent spaces and linear maps[4] gave the first non-trivial model of LL proofs. This model arose from Girard's analysis of the ccc STAB, realizing that there were many other logical operations available. Indeed, STAB is the Kleisli category of an approproate comonad $(!, \delta, \varepsilon)$ on LIN (cf. Proposition 4.15). In the model LIN, $!A$ is a minimal solution of the domain equation $!A \cong I \times A \times (!A \otimes !A)$, indeed is a cofree comonoid.
- *Relational Models.* As discussed in Barr [Barr91], many compact categories are complete enough to interpret

  $$!A \cong \perp \times A \times E_2(A) \times \cdots \times E_n(A) \times \cdots,$$

  where $E_n(A)$ is the equalizer of the $n!$ permutations of the nth tensor power $A^{\otimes n}$, for $n \geqslant 2$. For example, Barr proves $\mathbf{Rel}_\times$ has that property. More generally, Barr [Barr91] constructs models based on the *Chu-space* construction in [Barr79]. Chu spaces are themselves an interesting class of models of LL and have been the subject of intensive investigation by Michael Barr [Barr91,Barr95] and by Vaughan Pratt (e.g., [Pra95]).
- *Games Models.* Categories of Games now provide some of the most exciting new semantics for LL and Programming Languages. This so-called *intensional semantics* provides a finer-grained analysis of computation than traditional (categorical) models, taking into account the dynamic or interactive aspects of computation. For example, such games can be used to model interactions between a System and its Environment and provided the first syntax-free fully abstract models of PCF, answering a long-standing open

[4] A stable map is *linear* if it commutes with arbitrary unions.

problem. Games categories have been extended to handle programming languages with many additional properties, e.g., control features, call-by-value languages, languages with side-effects and store, etc. as well as modern logics like LL, System $\mathcal{F}$, etc. For basic introductions, see [Abr97,Hy97]. For a small sample of more recent work, see [Mc97,AHMc98,AMc98,BDER97].

- *GoI and Functional Analytic Models*: The Geometry of Interaction Program (see, e.g., [Gi88,Gi90,Gi95b,DR95]) aims to model the dynamics of cut-elimination by interpreting proofs as objects of a certain C^* algebra, with logical rules corresponding to certain $*$-isomorphisms. The essence of Gentzen's cut-elimination theorem is summarized by the so-called *execution formula*. We shall look at an abstract form of the GoI program (in traced monoidal categories) in Section 6.1. The GoI program itself has influenced both game semantics and work on optimal reduction.
- Finally, as the name suggests, linear logic is roughly inspired from linear algebra. Thus $!A$ is analogous to the Grassmann algebra. Indeed, in categories of Hilbert or Banach spaces, one is reminded of the symmetric and antisymmetric Fock space construction [Ge85]. For a (non-categorical) Banach space interpretation of LL, see Girard [Gi96].

5. Full completeness

5.1. *Representation theorems*

The most basic representation theorem of all is the Yoneda embedding:

THEOREM 5.1 (Yoneda). *If $\mathcal{A}$ is locally small, the Yoneda functor $\mathcal{Y}:\mathcal{A}\to\mathbf{Set}^{\mathcal{A}^{op}}$, where $\mathcal{Y}(A)=Hom_{\mathcal{A}}(-,A)$, is a fully faithful embedding.*

Indeed, Yoneda preserves limits as well as Cartesian closedness.

We seek mathematical models which describe the behaviour of programs. From the viewpoint of the Curry–Howard isomorphism (which identifies proofs with programs) we seek representation theorems for proofs – i.e. mathematical models which fully and faithfully represent proofs. From the viewpoint of a logician, these are Completeness Theorems, but now at the level of proofs rather than provability.

One of the first such theorems was proved by H. Friedman [Frie73]. Friedman showed completeness of typed lambda calculus with respect to ordinary set-theoretic reasoning. Consider the pure typed lambda calculus $\mathcal{L}^{\Rightarrow}$, whose types are generated from some base sorts by $\Rightarrow$ only. We interpret $\mathcal{L}^{\Rightarrow}$ set-theoretically in a full type hierarchy $\mathcal{A}$ (see Example 2.3).

THEOREM 5.2 (Friedman). *Let $\mathcal{A}$ be a full type hierarchy with base sorts interpreted as infinite sets. Then for any pure typed lambda terms M, N, $M \underset{\Gamma}{=} N$ is true in $\mathcal{A}$ iff $M \underset{\Gamma}{=} N$ is provable using the rules of typed lambda calculus.*

Similar results but using instead full type hierarchies over ω-CPO or Per-based models have been given by Plotkin and by Mitchell using logical relations (see [Mit96]). Friedman's original **Set**-based theorem has been extended by Čubrić to the entire ccc language $\Rightarrow, \times, \mathbf{1}$ [Cu93] to yield the following

THEOREM 5.3 (Čubrić). *Let $\mathcal{C}$ be a free ccc generated by a graph. Then there exists a faithful ccc functor $F : \mathcal{C} \to$ **Set**.*

Alas this representation is not full.

Let $\mathcal{B}_{\mathcal{G}}$ = the free ccc with binary coproducts generated by discrete graph $\mathcal{G}$, given syntactically by types and terms of typed lambda calculus. For any group G, the functor category $\mathbf{Set}^G$ is a ccc with coproducts. So according to the universal property, if F is an initial assignment of G-sets to atomic types then a proof of formula σ, qua closed term $M : \sigma$, qua $\mathcal{F}_{\mathcal{G}}$-arrow $M : \mathbf{1} \to \sigma$, corresponds to a G-set (= equivariant) map $[\![M_F]\!] : \mathbf{1} \to [\![\sigma_F]\!]$. Such maps are fixed points under the action. In particular, letting $G = \mathbf{Z}$, we obtain the easy half of the following theorem, due to Läuchli [Lau]:

THEOREM 5.4 (Läuchli). *A $\{\top, \wedge, \Rightarrow, \vee\}$-formula σ of intuitionistic propositional calculus is provable if and only if for every interpretation F of the base types, its $Set^{\mathbf{Z}}$-interpretation $[\![\sigma_F]\!]$ has an invariant element.*

Indeed, Harnik and Makkai extend Läuchli's theorem to a representation theorem. Recall, a functor Φ is *weakly full* if $Hom(A, B) = \emptyset$ implies $Hom(\Phi(A), \Phi(B)) = \emptyset$.

THEOREM 5.5 (Harnik, Makkai [HM92]). *Let $\mathcal{B}$ be a countable free ccc with binary coproducts. There is a weakly full representation Φ of $\mathcal{B}$ into a countable power of $Set^{\mathbf{Z}}$. If in addition the terminal object $\mathbf{1}$ is indecomposable, then there is a weakly full representation into $Set^{\mathbf{Z}}$.*

A weakly full representation of $\mathcal{B}$ corresponds to *completeness with respect to provability*: i.e. $Hom_{Set^{\mathbf{Z}}}(\mathbf{1}, \Phi(B)) \neq \emptyset$ implies $Hom_{\mathcal{B}}(\mathbf{1}, B) \neq \emptyset$, so B is provable. We shall give stronger representation theorems still based on the idea of invariant elements.

5.2. *Full completeness theorems*

A recent topic of considerable interest is *full completeness* theorems. Suppose we have a free category $\mathcal{F}$. We shall say that a model category $\mathcal{M}$ is *fully complete for $\mathcal{F}$* or that we have *full completeness of $\mathcal{F}$ with respect to $\mathcal{M}$* if the unique free functor (with respect to any interpretation of the generators) $[\![-]\!] : \mathcal{F} \to \mathcal{M}$ is full. It is even better to demand that $[\![-]\!]$ is a fully faithful representation.

For example, suppose $\mathcal{F}$ is a free ccc generated by the typed lambda calculus (cf. Example 2.22). To say a ccc $\mathcal{M}$ is fully complete for $\mathcal{F}$ means the following: given any interpretation of the generators as objects of $\mathcal{M}$, any arrow $[\![A]\!] \to [\![B]\!] \in \mathcal{M}$ between definable objects is itself definable, i.e. of the form $[\![f]\!]$ for $f : A \to B$. If the representation is fully faithful, f is unique. Thus, by Curry–Howard, any morphism in the model between definable objects is itself the image of a proof (or program); indeed of a unique proof if the representation is fully faithful. Thus, such models $\mathcal{M}$, while being semantical, really capture aspects of the syntax of the language.

Such results are mainly of interest when the models $\mathcal{M}$ are "genuine" mathematical models not apparently connected to the syntax. In that case Full Completeness results are more surprising (and interesting). For example, an explicit use of the Yoneda embedding $\mathcal{Y}:\mathcal{C} \to \mathbf{Set}^{\mathcal{C}^{op}}$ is not what we want, since $\mathbf{Set}^{\mathcal{C}^{op}}$ depends too much on $\mathcal{C}$.

Probably the first full completeness results for free ccc's were by Plotkin [Plo80], using categories of logical relations. In the case of simply typed lambda calculus generated from a fixed base type (= the free ccc on one object), Plotkin proved the following result. Consider the Henkin model $\mathbf{T}_B$ = *the full type hierarchy over a set* B, i.e. the full sub-ccc of Sets generated by some set B. The Soundness Theorem for logical relations says that if a term f is lambda definable, it is invariant under all logical relations. We ask for the converse.

The *rank* of a type is defined inductively: $\mathrm{rank}(b) = 0$, where b is a base type, $\mathrm{rank}(\sigma \Rightarrow \tau) = \max\{\mathrm{rank}(\sigma) + 1, \mathrm{rank}(\tau)\}$, $\mathrm{rank}(\sigma \times \tau) = \max\{\mathrm{rank}(\sigma), \mathrm{rank}(\tau)\}$. The rank of an element $f \in B_\sigma$ in $\mathbf{T}_B$ is the rank of the type σ.

THEOREM 5.6 (Plotkin [Plo80]). *In the full type hierarchy* $\mathbf{T}_B$ *over an infinite set* B, *all elements* f *of rank* $\leqslant 2$ *satisfy: if* f *is invariant under all logical relations, then* f *is lambda definable.*

This result has been extended by Statman [St85], but the same question for terms of arbitrary rank is still open. However Plotkin [Plo80] did prove the above result for lambda terms of arbitrary rank, by moving to *Kripke Logical Relations* rather than **Set**-based logical relations. Kripke relations occur essentially by replacing **Set** by a functor category $\mathbf{Set}^{P^{op}}$, P a poset, i.e. by looking at P-indexed families of sets and relations. Extensions, with new characterizations of lambda definability, are in work of Jung and Tiuryn [JT93]. A clear categorical treatment of their work, and logical-relations-based full completeness theorems, is in Alimohamed [Ali95] (cf. also [Mit96]).

The name "Full Completeness" first arose in Game Semantics, where the fundamental paper of Abramsky and Jagadeesan [AJ94b] proved full completeness for multiplicative linear logic (+ the Mix rule), using categories of games with history-free winning strategies as morphisms. It is shown there that "uniform" history-free winning strategies are the denotations of unique proof nets. A more involved notion of game, developed by Hyland and Ong (see [Hy97]), permits eliminating the Mix rule in proofs of full completeness for the multiplicatives. These results paved the way for the most spectacular application of these game-theoretic methods: the solution of the Full Abstraction problem for PCF, by Abramsky, Jagadeesan, and Malacaria and by Hyland and Ong, referred to in Section 4.3.2.

In [BS96,BS96b,BS98] Full Completeness for MLL + Mix and for Yetter's Cyclic Linear Logic were also proved using dinaturality and a generalization of Laüchli semantics. Let us briefly recall that view.

5.2.1. *Linear Läuchli semantics.* Let $\mathcal{C}$ be a $*$-autonomous category. Given an *MLL* formula $\varphi(\alpha_1, \ldots, \alpha_n)$ built from $\otimes$, $\multimap$, $(\ )^\perp$, with type variables $\alpha_1, \ldots, \alpha_n$, we inductively define its *functorial interpretation* $[\![\varphi(\alpha_1, \ldots, \alpha_n)]\!] : (\mathcal{C}^{op})^n \times \mathcal{C}^n \to \mathcal{C}$ as follows:

- $[\![\varphi]\!](\mathbf{AB}) = \begin{cases} B_i & \text{if } \varphi(\alpha_1,\ldots,\alpha_n) \equiv \alpha_i, \\ A_i^{\perp} & \text{if } \varphi(\alpha_1,\ldots,\alpha_n) \equiv \alpha_i^{\perp}. \end{cases}$
- $[\![\varphi_1 \otimes \varphi_2]\!](\mathbf{AB}) = [\![\varphi_1]\!](\mathbf{AB}) \otimes [\![\varphi_2]\!](\mathbf{AB})$.
- $[\![\varphi_1 \multimap \varphi_2]\!](\mathbf{AB}) = [\![\varphi_1]\!](\mathbf{BA}) \multimap [\![\varphi_2]\!](\mathbf{AB})$.

The last two clauses correspond to Eqs. (11) and (12) (following Example 3.5 in Section 3.1). It is readily verified that $[\![\varphi^{\perp}]\!] = [\![\varphi]\!]^{\perp}$. Also recall that in *MLL*, $A \multimap B$ is defined as $A^{\perp} \mathbin{⅋} B$. From now on, let $\mathcal{C} = \mathcal{RTVEC}$.

The set $Dinat(F, G)$ of dinatural transformations from F to G is a vector space, under pointwise operations. We call it the *space of proofs* associated to the sequent $F \vdash G$ (where we identify formulas with definable functors.) If $\vdash \Gamma$ is a one-sided sequent, then $Dinat(\Gamma)$ denotes the set of dinaturals from $\mathbf{k}$ to $[\![⅋\Gamma]\!]$. In such sequents, we sometimes abbreviate $[\![⅋\Gamma]\!]$ to $[\![\Gamma]\!]$.

The following is proved in [BS96,BS98]. A binary sequent is one where each atom appears exactly twice, with opposite variances.

THEOREM 5.7 (Full completeness for binary sequents). *Let F and G be formulas in multiplicative linear logic, interpreted as definable multivariant functors on $\mathcal{RTVEC}$. Given a binary sequent $F \vdash G$, then $Dinat(F, G)$ is zero or 1-dimensional, depending on whether or not $F \vdash G$ is provable. If it is provable, every dinatural is a scalar multiple of the denotation of the unique cut-free proof (qua Girard proof-net).*

A *diadditive dinatural transformation* is one which is a linear combination of substitution instances of binary dinaturals. Under the same hypotheses as above we obtain:

THEOREM 5.8 (Full completeness). *The proof space $Dinat(F, G)$ of diadditive dinatural transformations has as basis the denotations of cut-free proofs in the theory MLL + MIX.*

EXAMPLE 5.9. The proof space of the sequent

$$\alpha_1, \alpha_1 \multimap \alpha_2, \alpha_2 \multimap \alpha_3, \ldots, \alpha_{n-1} \multimap \alpha_n \vdash \alpha_n$$

has dimension 1, generated by the evaluation dinatural.

The proofs of the above results actually yield a fully faithful representation theorem for a free $*$-autonomous category with MIX, canonically enriched over vector spaces ([BS98]).

In [BS98], a similar Full Completeness Theorem and fully faithful representation theorem is given for Yetter's Cyclic Linear Logic. In this case one employs the category $\mathcal{RTMOD}(\mathsf{H})$, for a Hopf algebra H.[5] This is based on the following observation [Blu96]:

PROPOSITION 5.10. *If H is a Hopf algebra with an involutive antipode, i.e. $S^2 = id$ then $\mathcal{RTMOD}(\mathsf{H})$ is a cyclic $*$-autonomous category, i.e. a model of Yetter's cyclic linear logic.*

[5] Recently, M. Hamano [Ham99] eliminated MIX in the dinatural framework, using Pontrjagin Duality. He also generalized [BS96,BS98] to a non-dinatural semantics.

The particular Hopf algebra used is the shuffle Hopf algebra, described in [Ben,Haz, BS98]. Once again we consider formulas as multivariant functors on $\mathcal{RTVEC}$, but restrict the dinaturals to so-called *uniform dinaturals* $\theta_{|V_1|,\ldots,|V_n|}$, i.e. those which are equivariant with respect to the H-action induced from the atoms, for H-modules $V_i \in \mathcal{RTMOD}(\mathsf{H})$. This is completely analogous to the techniques used in logical relations.

Related results using Chu spaces are in [Pra97].

6. Feedback and trace

6.1. *Traced monoidal categories*

This new class of categories, introduced by Joyal, Street, and Verity [JSV96], have shown surprising connections to models of computation and iteration. The original versions were very general, including braided and tortile categories arising in several branches of mathematics. At the moment, most of the applications to computing omit any braided structure. But even at the abstract level of [JSV96], the authors illustrate a computational, geometric calculus somewhat akin to Girard's proof nets in linear logic [Gi87], and indeed some precise connections have been made [BCS98]. Moreover, the main construction in [JSV96] has been shown by Abramsky [Abr96] to have fascinating connections with Girard's GoI program, as already hinted by Joyal, Street, and Verity.

We now give a version of traced symmetric monoidal categories. For ease of readability and without loss of generality, we consider strict monoidal categories (recall, from Mac Lane's coherence theorem, that every monoidal category is equivalent to a strict one).

DEFINITION 6.1. A *traced symmetric monoidal category* (= *tmc*) is a symmetric monoidal category $(\mathcal{C}, \otimes, I, s)$ (where $s_{X,Y} : X \otimes Y \to Y \otimes X$ is the symmetry morphism) with a family of functions $Tr^U_{X,Y} : \mathcal{C}(X \otimes U, Y \otimes U) \to \mathcal{C}(X, Y)$, called a *trace*, subject to the following conditions:

- **Natural** in X, $Tr^U_{X,Y}(f)g = Tr^U_{X',Y}(f(g \otimes 1_U))$, where $f : X \otimes U \to Y \otimes U$, $g : X' \to X$,
- **Natural** in Y, $gTr^U_{X,Y}(f) = Tr^U_{X,Y'}((g \otimes 1_U)f)$, where $f : X \otimes U \to Y \otimes U$, $g : Y \to Y'$,
- **Dinatural** in U, $Tr^U_{X,Y}((1_Y \otimes g)f) = Tr^{U'}_{X,Y}(f(1_X \otimes g))$, where $f : X \otimes U \to Y \otimes U'$, $g : U' \to U$,
- **Vanishing**, $Tr^I_{X,Y}(f) = f$ and $Tr^{U \otimes V}_{X,Y}(g) = Tr^U_{X,Y}(Tr^V_{X \otimes U, Y \otimes U}(g))$ for $f : X \otimes I \to Y \otimes I$ and $g : X \otimes U \otimes V \to Y \otimes U \otimes V$.
- **Superposing**, $g \otimes Tr^U_{X,Y}(f) = Tr^U_{W \otimes X, Z \otimes Y}(g \otimes f)$
- **Yanking**, $Tr^U_{U,U}(s_{U,U}) = 1_U$.

From a computer science viewpoint, the essential feature is to think of $Tr^U_{X,Y}(f)$ as "feedback along U", as in Figure 9. The axiomatization given here differs slightly from those in [Abr96,JSV96], although it can be shown to be equivalent. We shall leave it to the reader to draw the diagrams for the trace axioms. We note however that Vanishing

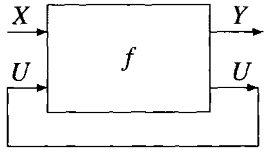

Fig. 9. The trace $Tr^U_{X,Y}(f)$.

expresses trace along a tensor $U \otimes V$ in terms of iterated traces along U and V. This is related to the so-called Bekič Lemma in Domain Theory.

The above notion is really a *parametrized trace*. The usual notion from linear algebra is when $X = Y = I$ (see Example 6.3) below.

EXAMPLE 6.2. $\mathbf{Rel}_\times$: The objects are sets, $\otimes = \times$ (Cartesian product), and maps are binary relations. Composition means composition of relations, and $x\, Tr^U_{X,Y}\, y$ iff there exists a u such that $(x, u) R (y, u)$.

EXAMPLE 6.3. $\mathbf{Vec}_{fd}$: Given $f : X \otimes U \to Y \otimes U$, define

$$Tr^U_{X,Y}(f)(x_i) = \sum_{j,k} \alpha^{kj}_{ij}\, y_k, \quad \text{where } f(x_i \otimes u_j) = \sum_{k,m} \alpha^{km}_{ij}\, y_k \otimes u_m,$$

where $(u_i), (x_j), (y_k)$ are bases for U, X, Y, resp. In the case that X, Y are one-dimensional, this reduces to the usual trace of a linear map $f : U \to U$, i.e. the usual trace determines a function $Tr_U : Hom(U, U) \to Hom(I, I)$, where $I = \mathbf{k}$.

EXAMPLE 6.4. More generally, any compact category has a canonical trace

$$Tr^U_{X,Y}(f) = X \cong X \otimes I \xrightarrow{id \otimes \eta} X \otimes U \otimes U^\perp \xrightarrow{f \otimes id} Y \otimes U \otimes U^\perp \xrightarrow{id \otimes ev'} Y \otimes I \cong Y,$$

where $ev' = ev \circ s$.

EXAMPLE 6.5. ω-$\mathbf{CPO}_\perp$: with $\otimes = \times$, $I = \{\perp\}$. In this case the dinatural least-fixed-point combinator $\mathrm{Y}_- : (-)^{(-)} \to (-)$ induces a trace, given as follows (using informal lambda calculus notation): for any $f : X \times U \to Y \times U$,

$$Tr^U_{X,Y}(f)(x) = f_1\big(x, \mathrm{Y}_U\big(\lambda u \cdot f_2(x, u)\big)\big),$$

where $f_1 = \pi_1 \circ f : X \times U \to Y$, $f_2 = \pi_2 \circ f : X \times U \to U$. Hence $Tr^U_{X,Y}(f)(x) = f_1(x, u')$, where u' is the smallest element of U such that $f_2(x, u') = u'$. A generalization of this idea to *traced Cartesian categories* is in [MHas97] and mentioned in Remark 6.16 in the next section.

Unfortunately, these examples do not really illustrate the notion of feedback as data flow: the movement of tokens through a network. More natural examples of traced monoidal

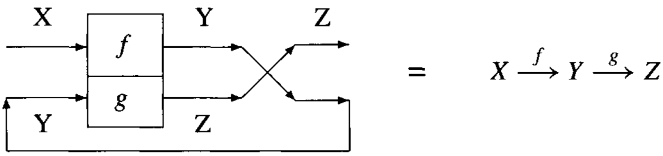

Fig. 10. Generalized Yanking.

categories in the next section, given by partially additive and similar iterative categories, more fully illustrate this aspect.

EXAMPLE 6.6. *Bicategories of Processes*: The paper of Katis, Sabadini, and Walters [KSW95] develops a general theory of processes with feedback circuits in symmetric monoidal bicategories. They prove their bicategories $\mathbf{Circ}(\mathcal{C})$ have a parametrized trace operator. A small difference with the above treatment is that their feedback is given by a family of *partially-defined* functors $fb^U_{X,Y} : \mathbf{Circ}(\mathcal{C})(X \otimes U, Y \otimes U) \to \mathbf{Circ}(\mathcal{C})(X, Y)$.

REMARK 6.7. The paper [ABP97] develops a general theory of traced ideals in tensored $*$-categories. The category $\mathcal{HILB}$, the tensored $*$-category of Hilbert spaces and bounded linear maps, illustrates the difficulty. In passing from the finite dimensional case (cf. Example 6.3 above) to the infinite dimensional one, not all endomorphisms have a trace; for example, the identity on an infinite dimensional space. However Tr_U may be defined on *traced ideals* of maps, and this extends to parametrized traces. See [ABP97] for many examples.

An amusing folklore about traced monoidal categories is that general composition is actually definable using traces of simple compositions:

PROPOSITION 6.8 (Generalized Yanking). *Let $\mathcal{C}$ be a traced symmetric monoidal category, with arrows $f : X \to Y$ and $g : Y \to Z$. Then $g \circ f = Tr^Y_{X,Z}(s_{Y,Z} \circ (f \otimes g))$.*

Although a fairly short algebraic proof is possible, the reader may wish to stare at the diagram in Figure 10 , and do a "string-pulling" argument (cf. [JSV96]). Similar calculations are in [KSW95,Mil94]

DEFINITION 6.9. Let $\mathcal{C}$ and $\mathcal{D}$ be traced symmetric monoidal categories. A strong monoidal functor $F : \mathcal{C} \to \mathcal{D}$ is *traced* if it is symmetric and satisfies

$$Tr^{FU}_{FA,FB}(\phi^{-1}_{B,U}(Ff)\phi_{A,U}) = F(Tr^U_{A,B}(f))$$

where $A \otimes U \xrightarrow{f} B \otimes U$ and $FA \otimes FU \xrightarrow{\phi_{A,U}} F(A \otimes U) \xrightarrow{Ff} F(B \otimes U) \xrightarrow{\phi^{-1}_{B,U}} FB \otimes FU$. In the case of *strict* monoidal functors, they are traced if they preserve the trace on the nose.

We define **TraMon** and $\mathbf{TraMon}_{st}$ to be the 2-categories whose 0-cells are traced monoidal categories (resp. strict traced monoidal categories), whose 1-cells are traced monoidal functors (resp. strict traced monoidal functors), and whose 2-cells are monoidal natural transformations.

6.2. *Partially additive categories*

We shall be interested in special kinds of traced monoidal categories: those whose homsets are enriched with certain partially-defined infinite sums, which permits canonical calculation of iteration and traces (see formulas (17) and (18) below). A useful example is Manes and Arbib's *partially additive categories*, which first arose in their categorical analysis of iterative and flowchart schema [MA86]. Categories with similar additive structure on the hom-sets had already been considered in the 1950's by Kuroš [Ku63] with regards to categorical Krull–Schmidt–Remak theorems.

DEFINITION 6.10. A *partially additive monoid* is a pair (M, Σ), where M is a nonempty set and Σ is a partial function which maps countable families in M to elements of M (we say that $(x_i \mid i \in I)$ is *summable* if $\Sigma(x_i \mid i \in I)$ is defined)[6] subject to the following:

1. *Partition-Associativity Axiom.* If $(x_i \mid i \in I)$ is a countable family and if $(I_j \mid j \in J)$ is a (countable) partition of I, then $(x_i \mid i \in I)$ is summable if and only if $(x_i \mid i \in I_j)$ is summable for every $j \in J$ and $(\Sigma(x_i \mid i \in I_j) \mid j \in J)$ is summable. In that case,

$$\Sigma(x_i \mid i \in I) = \Sigma\big(\Sigma(x_i \mid i \in I_j) \mid j \in J\big).$$

2. *Unary Sum Axiom.* Any family $(x_i \mid i \in I)$ in which I is a singleton is summable and $\Sigma(x_i \mid i \in I) = x_j$ if $I = \{j\}$.

3.* *Limit Axiom.* If $(x_i \mid i \in I)$ is a countable family and if $(x_i \mid i \in F)$ is summable for every finite subset F of I then $(x_i \mid i \in I)$ is summable.

We observe the following facts about partially additive monoids:

(i) Axioms 1 and 2 imply that the empty family is summable. We denote $\Sigma(x_i \mid i \in \emptyset)$ by 0, which is an additive identity for summation.

(ii) Axiom 1 implies the obvious equations of commutativity and associativity for the sum (when defined).

(iii) Although Manes and Arbib use the Limit Axiom to prove existence of Elgot-style iteration (see below), Kuroš did not have it. And for many aspects of the theory below, it is not needed.

DEFINITION 6.11. The category of *partially additive monoids*, **PAMon**, is defined as follows. Its objects are partially additive monoids (M, Σ). Its arrows $(M, \Sigma_M) \xrightarrow{f} (N, \Sigma_N)$ are maps from M to N which preserve the sum, in the sense that: $f(\Sigma_M(x_i \mid i \in I)) = \Sigma_N(f(x_i) \mid i \in I)$ for all summable families $(x_i \mid i \in I)$ in M. Composition and identities are inherited from **Sets**.

[6] We sometimes abbreviate $\Sigma(x_i \mid i \in I)$ by $\Sigma_{i \in I} x_i$. Throughout, "countable" means finite or denumerable. All index sets are countable. A *partition* $\{I_j \mid j \in J\}$ of I satisfies: $I_j \subseteq I$, $I_i \cap I_j = \emptyset$ if $i \neq j$, and $\bigcup\{I_j \mid j \in J\} = I$. But we also allow $I_j = \emptyset$ for countably many j.

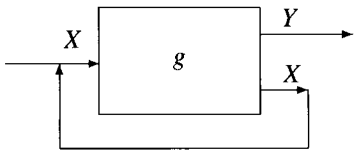

Fig. 11. Elgot dagger $g^\dagger$.

A **PAMon**-category $\mathcal{C}$ is a category enriched in **PAMon**. This means the hom-sets carry a **PAMon**-structure, compatible with composition. In particular, in each homset $Hom_{\mathcal{C}}(X, Y)$ there is a zero morphism $0_{XY} : X \to Y$, the sum of the empty family.

REMARK 6.12. In a **PAMon**-category $\mathcal{C}$

(1) The family of zero morphisms $\{0_{XY}\}_{X,Y\in\mathcal{C}}$ satisfies: $g0_{WZ} = 0_{WY} = 0_{XY}f$ for any $f : W \to X$ and $g : Z \to Y$.

(2) If $\Sigma_{i\in I} f_i = 0_{XY}$ then all summands $f_i = 0_{XY}$ in $Hom_{\mathcal{C}}(X, Y)$.

DEFINITION 6.13. Let $\mathcal{C}$ be a **PAMon**-category with countable coproducts $\bigoplus_{i\in I} X_i$. For any $j \in I$ we define *quasi projections* $PR_j : \bigoplus_{i\in I} X_i \to X_j$ as follows:

$$PR_j in_k = \begin{cases} id_{X_j} & \text{if } k = j, \\ 0_{X_k X_j} & \text{else.} \end{cases}$$

DEFINITION 6.14. A *partially additive category* (pac) $\mathcal{C}$ is a **PAMon**-category with countable coproducts which satisfies the following axioms:

(1) *Compatible Sum Axiom*: If $(f_i \mid i \in I)$ is a countable family of morphisms in $\mathcal{C}(X, Y)$ and there exists $f : X \to \coprod_I Y$ such that $PR_i f = f_i \; X \xrightarrow{f_i} Y$ (we say the f_i are *compatible*), then $\sum f_i$ exists. Here $\coprod_I Y$ is the I-fold coproduct of Y (often denoted $Y^{(I)}$ in algebra).

(2) *Untying Axiom*: If $f + g : X \to Y$ exists then so does $in_1 f + in_2 g : X \to Y + Y$.

The following facts about partially additive categories follow from Manes and Arbib [MA86]:

- *Matrix Representation of maps*: For any map $f : \bigoplus_{i\in I} X_i \to \bigoplus_{j\in J} Y_j$ there is a unique family $\{f_{ij} : X_i \to Y_j\}_{i\in I j\in J}$ with $f = \sum_{i\in I j\in J} in_j f_{ij} PR_i$ and $PR_j f in_i = f_{ij}$. *Notation*: we write $f_{X_i Y_j}$ for f_{ij}.
- *Elgot Iteration*: Given $X \xrightarrow{g} Y + X$, there exists $X \xrightarrow{g^\dagger} Y$ where $g^\dagger = \sum_{n=0}^{\infty} g_{XY} g_{XX}^n$, satisfying the *fixed-point identity*

$$[1_Y, g^\dagger]g = g^\dagger. \tag{17}$$

This corresponds to the flowchart scheme in Figure 11. The proof of Elgot iteration ([MA86], p. 83) uses the Limit Axiom.

PROPOSITION 6.15. *A partially additive category C is traced monoidal with $\otimes$ = coproduct and if $f : X \oplus U \to Y \oplus U$,*

$$Tr^U_{X,Y}(f) = [1_Y, f_2^\dagger] f_1 = f_{XY} + \sum_{n=0}^{\infty} f_{UY} f^n_{UU} f_{XU}, \tag{18}$$

where $f = [f_1, f_2]$ with $f_1 : X \to Y \oplus U$, $f_2 : U \to Y \oplus U$.

Formula (18) corresponds to the data flow interpretation of trace-as-feedback in Figure 9: see the examples below. We should also remark that Formula (18) corresponds closely to Girard's *Execution Formula* [Gi88,Abr96] and is related to a construction of Geometry of Interaction categories in the next section.

REMARK 6.16. Conversely, a traced monoidal category where $\otimes$ is coproduct has an Elgot iterator $g^\dagger = Tr^X_{X,Y}([g, g])$, where $g : X \to Y + X$. An axiomatization of the *opposite* of such categories, which correspond to categories with a parametrized Y combinator, is considered in Hasegawa [MHas97]. More generally, Hasegawa considers traced monoidal categories built over Cartesian categories and it is shown how various typed lambda calculi with *cyclic sharing* are Sound and Complete for such categorical models.

Finally, we should mention the general notion of *Iteration Theories*. These general categorical theories of feedback and iteration, their axiomatization and equational logics have been studied in detail by S.L. Bloom and Z. Ésik in their book [BE93]. A more recent 2-categorical study of iteration theories is in [BELM].

We shall now give a few important examples of pac's:

EXAMPLE 6.17. $\mathbf{Rel}_+$, the category of sets and relations. Objects are sets and maps are binary relations. Composition means relational composition. The identity is the identity relation, and the zero morphism 0_{XY} is the empty set $\emptyset \subseteq X \times Y$. Coproducts $\bigoplus_{i\in I} X_i$ are as in **Set**, i.e. disjoint union. *All* countable families are summable where $\Sigma_{i\in I}(R_i) = \bigcup_{i\in I} R_i$. Finally, let R^* be the reflexive, transitive closure of a relation R. Suppose $R : X + U \nrightarrow Y + U$. Then formula (18) becomes:

$$\begin{aligned} Tr^U_{X,Y}(R) &= R_{XY} \cup \bigcup_{n\geqslant 0} R_{UY} \circ R^n_{UU} \circ R_{XU} \\ &= R_{XY} \cup R_{UY} \circ R^*_{UU} \circ R_{XU}. \end{aligned} \tag{19}$$

EXAMPLE 6.18. **Pfn**, the category of sets and partial functions. The objects are sets, the maps are partial functions. Composition means the usual composition of partial functions. The zero map 0_{XY} is the empty partial function. A family $\{f_i \mid i \in I\}$ is said to be summable iff $\forall i, j \in I,\ i \neq j,\ Dom(f_i) \cap Dom(f_j) = \emptyset$. $\Sigma_{i\in I} f_i$ is the partial function with domain $\bigcup_i Dom(f_i)$ and where

$$(\Sigma_{i\in I} f_i)(x) = \begin{cases} f_j(x) & \text{if } x \in Dom(f_j), \\ \text{undefined} & \text{else.} \end{cases}$$

The following example comes from Giry [Giry] (inspired from Lawvere) and is mentioned in [Abr96]. The fact that this is a pac follows from work of P. Panangaden and E. Haghverdi.

EXAMPLE 6.19. **SRel**, the category of Stochastic Relations. Objects are measurable spaces (X, Σ_X) where X is a set and Σ_X is a σ-algebra of subsets of X. An arrow $f:(X, \Sigma_X) \to (Y, \Sigma_Y)$ is a transition probability, i.e. $f: X \times \Sigma_Y \to [0,1]$ such that $f(\cdot, B): X \to [0,1]$ is a measurable function for fixed $B \in \Sigma_Y$ and $f(x, \cdot): \Sigma_Y \to [0,1]$ is a subprobability measure (i.e. a σ-additive set function satisfying $f(x, \emptyset) = 0$ and $f(x, Y) \leqslant 1$). The identity morphism $id_X: (X, \Sigma_X) \to (X, \Sigma_X)$ is a map $id_X: X \times \Sigma_X \to [0,1]$, with $id_X(x, A) = \delta(x, A)$, where for A fixed, $\delta(x, A)$ is the characteristic function of A and for x fixed, $\delta(x, A)$ is the Dirac distribution.

Composition is defined as follows: given $f:(X, \Sigma_X) \to (Y, \Sigma_Y)$ and $g:(Y, \Sigma_Y) \to (Z, \Sigma_Z)$, $g \circ f: (X, \Sigma_X) \to (Z, \Sigma_Z)$ is $g \circ f(x, C) = \int_Y g(y, C)\, d\{f(x, \cdot)\}$, where the notation $d\{f(x, \cdot)\}$ means that we are fixing x and using $f(x, \cdot)$ as the measure for the integration, the function being integrated is the measurable function $g(\cdot, C)$.

Given (X, Σ_X) and (Y, Σ_Y), the *zero morphism* $0_{XY}: (X, \Sigma_X) \to (Y, \Sigma_Y)$, is given by $0_{XY}(x, B) = 0$ for all $x \in X$ and $B \in \Sigma_Y$.

The partially additive structure on the homsets of **SRel** is as follows: we say an I-indexed family of morphisms $\{f_i \mid i \in I\}$ is *summable* if for all $x \in X$ we have $\sum_{i \in I} f_i(x, Y) \leqslant 1$. Since we are dealing with bounded, positive measures it is easy to verify that the sum so defined is a subprobability measure. Note that we would have only trivial additive structure (only singleton families summable) if we had used probability distributions rather than subprobability distributions.

Finally, let $\{X_i \mid i \in I\}$ be a countable family of objects. We define the coproduct $\bigoplus_{i \in I} X_i$ as follows. We take the disjoint union of the sets X_i, equipped with the evident σ-algebra. Thus a measurable subset will look like the disjoint union of measurable subsets of each of the X_i, say $\biguplus A_i$ (of course some of the A_i may be empty, and a point will be a pair (x, i) where $i \in I$ and $x \in X_i$). The canonical injections $in_j: X_j \to \bigoplus_{i \in I} X_i$ are $in_j(x, \biguplus A_i) = \delta(x, A_j)$. Given Y and $\forall i \in I$, arrows $h_i: X_i \to Y$, we obtain the mediating morphism $h: \bigoplus_{i \in I} X_i \to Y$ by the formula $h((x, j), B) = h_j(x, B)$. The verifications are all routine.

The next example, while not a pac, is essentially similar.

EXAMPLE 6.20. **Pinj**, the category of sets and injective partial functions. This is a fundamental example that arises in Girard's Geometry of Interaction program. Although this category is traced monoidal, with an iterative trace formula given in Abramsky [Abr96], it does not have coproducts. However its pac-like aspects may be captured in a Kuroš-style presentation via a generalization of partially additive categories, in which countable coproducts are replaced by countable tensors, and in which suitable axioms guarantee (analogously to pacs): a matrix representation of maps $\bigotimes_{i \in I} X_i \to \bigotimes_{j \in J} Y_j$ and a trace formula as in (18).

6.3. *GoI categories*

Girard's Geometry of Interaction (GoI) program introduces some profound new twists into computation theory. In particular, the idea that proofs are like dynamical systems, interacting locally. The dynamics of information flow in composition, via cut-elimination, is then related to tracing out paths in certain algebraic structures (Girard originally used operator algebras but the results can be expressed without them [Gi95b]). The connection of Girard's functional analytic methods in GoI with lambda calculus and proof nets is further explored in [DR95,MaRe91].

Starting with a traced monoidal category $\mathcal{C}$, we now describe a compact category $\mathcal{G}(\mathcal{C})$ (called $Int(\mathcal{C})$ in [JSV96]) which captures in abstract form many of the features of Girard's Geometry of Interaction program, as well as the general ideas behind game semantics. We follow the evocative treatment in Abramsky [Abr96]. The idea is to create a category whose composition is given by an iterative feedback formula, using the trace.

DEFINITION 6.21. (The *Geometry of Interaction* construction). Given a traced monoidal category $\mathcal{C}$ we define a compact closed category, $\mathcal{G}(\mathcal{C})$, as follows [JSV96,Abr96]:

- Objects: Pair of objects (A^+, A^-) where A^+ and A^- are objects of $\mathcal{C}$.
- Arrows: An arrow $f : (A^+, A^-) \to (B^+, B^-)$ in $\mathcal{G}(\mathcal{C})$ is an arrow $f : A^+ \otimes B^- \to A^- \otimes B^+$ in $\mathcal{C}$.
- Identity: $1_{(A^+,A^-)} = \sigma_{A^+,A^-}$.
- Composition: given by symmetric feedback. Arrows $f : (A^+, A^-) \to (B^+, B^-)$ and $g : (B^+, B^-) \to (C^+, C^-)$ have composite $g \circ f : (A^+, A^-) \to (C^+, C^-)$ given by:

  $$g \circ f = Tr^{B^- \otimes B^+}_{A^+ \otimes C^-, A^- \otimes C^+}\big(\beta(f \otimes g)\alpha\big),$$

  where $\alpha = (1_{A^+} \otimes 1_{B^-} \otimes \sigma_{C^-,B^+})(1_{A^+} \otimes \sigma_{C^-,B^-} \otimes 1_{B^+})$ and $\beta = (1_{A^-} \otimes 1_{C^+} \otimes \sigma_{B^+,B^-})(1_{A^-} \otimes \sigma_{B^+,C^+} \otimes 1_{B^-})(1_{A^-} \otimes 1_{B^+} \otimes \sigma_{B^-,C^+})$. A picture displaying $g \circ f$ is given in Figure 12.
- Tensor: $(A^+, A^-) \otimes (B^+, B^-) = (A^+ \otimes B^+, A^- \otimes B^-)$ and for $(A^+, A^-) \to (B^+, B^-)$ and $g : (C^+, C^-) \to (D^+, D^-)$, $f \otimes g = (1_{A^-} \otimes \sigma_{B^+,C^-} \otimes 1_{D^+})(f \otimes g)(1_{A^+} \otimes \sigma_{C^+,B^-}) \otimes 1_{D^-}$.

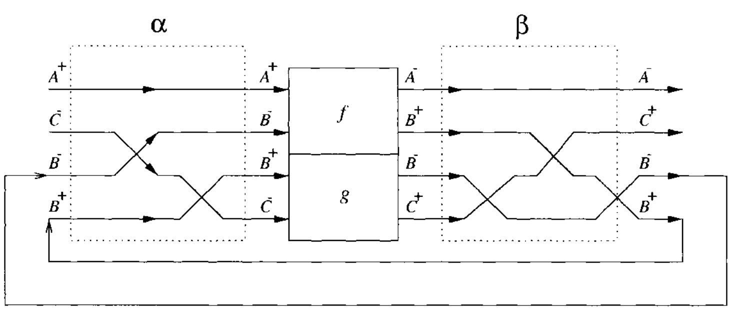

Fig. 12. Composition in $\mathcal{G}(\mathcal{C})$.

- Unit: (I, I).
- Duality: The dual of (A^+, A^-) is given by $(A^+, A^-)^\perp = (A^-, A^+)$ where the unit $\eta : (I, I) \to (A^+, A^-) \otimes (A^+, A^-)^\perp =_{def} \sigma_{A^-, A^+}$ and counit $\varepsilon : (A^+, A^-)^\perp \otimes (A^+, A^-) \to (I, I) =_{def} \sigma_{A^-, A^+}$.
- Internal Homs: As usual, $(A^+, A^-) \multimap (B^+, B^-) = (A^+, A^-)^\perp \otimes (B^+, B^-) = (A^- \otimes B^+, A^+ \otimes B^-)$.

REMARK 6.22. We have used a specific definition for α and β above; however, any other permutations $A^+ \otimes C^- \otimes B^- \otimes B^+ \xrightarrow{\cong} A^+ \otimes B^- \otimes B^+ \otimes C^-$ and $A^- \otimes B^+ \otimes B^- \otimes C^+ \xrightarrow{\cong} A^- \otimes C^+ \otimes B^- \otimes B^+$ for α and β respectively will yield the same result for $g \circ f$ due to coherence.

Translating the work of [JSV96] in our setting we obtain that $\mathcal{G}(\mathcal{C})$ is a kind of "free compact closure" of $\mathcal{C}$:

PROPOSITION 6.23. *Let $\mathcal{C}$ be a traced symmetric monoidal category*
- *$\mathcal{G}(\mathcal{C})$ defined as in Definition 6.21 is a compact closed category. Moreover, $F_\mathcal{C} : \mathcal{C} \to \mathcal{G}(\mathcal{C})$ defined by $F_\mathcal{C}(A) = (A, I)$ and $F_\mathcal{C}(f) = f$ is a full and faithful embedding.*
- *The inclusion of 2-categories* **CompCl** $\hookrightarrow$ **TraMon** *has a left biadjoint with unit having component at $\mathcal{C}$ given by $F_\mathcal{C}$.*

Following Abramsky [Abr96], we interpret the objects of $\mathcal{G}(\mathcal{C})$ in a game-theoretic manner: A^+ is the type of "moves by Player (the System)" and A^- is the type of "moves by Opponent (the Environment)". The composition of morphisms in $\mathcal{G}(\mathcal{C})$ is the key to Girard's Execution formula, especially for pac-like traces. In [Abr96] it is pointed out that $\mathcal{G}(\mathbf{Pinj})$ is essentially the original Girard formalism, while $\mathcal{G}(\omega\text{-CPO})$ is the data-flow model of GoI given in [AJ94a]. This is studied in detail in E. Haghverdi [Hagh99].

7. Literature notes

In the above we have merely touched on the large and varied literature. The journals *Mathematical Structures in Computer Science* (Cambridge University Press) and *Theoretical Computer Science* (Elsevier) are standard venues for categorical computer science. Two recent graduate texts emphasizing categorical aspects are J. Mitchell [Mit96] and R. Amadio and P.-L. Curien [AC98]. Mitchell's book has an encyclopedic coverage of the major areas and recent results in programming language theory (for example subtyping, which we have not discussed at all). The Amadio and Curien book covers many recent topics in domain theory and lambda calculi, including full abstraction results, foundations of linear logic, and sequentiality.

We regret that there are many important topics in categorical computer science which we barely mentioned. We particularly recommend the compendia [PD97,FJP,AGM]. Let us give a few pointers with sample papers:

- *Operational and Denotational Semantics:* See the surveys in the Handbook [AGM]. The classical paper on solutions of domain equations is [SP82]. For some recent directions in domain theory, see [FiPl96,ReSt97]. For recent categorical aspects of Operational Semantics, see [Pi97,TP96]. Higher-dimensional category theory has also generated considerable theoretical interest (e.g., [Ba97,HMP98]). Coalgebraic and coinductive methods are a fundamental technique and have considerable influence (e.g., see [AbJu94, Pi96a,Mul91,CSp91,JR,Mil89]).

- *Fibrations and Indexed Category Models*: This important area arising from categorical logic is fundamental in treating dependent types, System $\mathcal{F}, \mathcal{F}_\omega, \ldots$ models, and general variable-binding (quantifier-like) operations, for example "hiding" in certain process calculi. For fibred category models of dependent type-theories, see the survey by M. Hofmann [H97a] (cf. also [PowTh97,HJ95,Pi9?,See87]). Indexed category models for *Concurrent Constraint Logic Programming* are given in [PSSS,MPSS95] (see also [FRS98] for connections of this latter paradigm to LL).
- *Computational Monads*: E. Moggi introduced the categorists' notion of monads and comonads [Mac71] as a structuring tool in semantics and logics of programming languages. Moggi's highly influential approach permits a modular treatment of such important programming features as: exceptions, side-effects, non-determinism, resumptions, dynamic allocation, etc. as well as their associated logics [Mo97,Mo91]. Practical uses of monads in functional programming languages are discussed in P. Wadler ([W92]). More recently, E. Manes ([M98]) showed how to use monads to implement collection classes in Object-Oriented languages. Alternative category-theoretic perspectives on Moggi's work are in Power and Robinson's [PowRob97].
- *Categories in Concurrency Theory and Bisimulation*: This large and important area is surveyed in Winskel and Nielsen [WN97]. In particular, the fundamental notion of bisimulation via open maps, is introduced in Joyal, Nielsen, and Winskel [JNW]. Presheaf models for Milner's π-calculus ([Mil93a]) and other concurrent calculi are in [CaWi,CSW]. For categorical work on Milner's recent Action Calculi, see [GaHa, Mil93b,Mil94,P96].
- *Complexity Theory*: Characterizing feasible (e.g., polynomial-time) computation is a major area of theoretical computer science (e.g., [Cob64,Cook75]). Typed lambda calculi for feasible higher-order computation have recently been the subjects of intense work, e.g., [CoKa,CU93]. Versions of linear logic have been developed to analyze the fine structure of feasible computation ([GSS92,Gi97]). Although there are some known models ([KOS97]), general categorical treatments for these versions of LL are not yet known. Recently, M. Hofmann (e.g., [H97a]) has analyzed the work of Cook and Urquhart [CU93] as well as giving higher-order extensions of work of Bellantoni and Cook, using presheaf and sheaf categories.

Three volumes [MR97,FJP,GS89] are conferences specializing in applications of categories in computer science (note [MR97] is the 7th Biennial such meeting). Similarly, see the biennial meetings of MFPS (Mathematical Foundations of Programming Semantics) – published in either Springer Lecture Notes in Computer Science or the journal *Theoretical Computer Science*. There is currently an electronic website of categorical logic and computer science, HYPATIA (`http://hypatia.dcs.qmw.ac.uk`).

Other books covering categorical aspects of computer science and/or some of the topics covered here include [AL91,BW95,Cu93,DiCo95,Gun92,MA86,Tay98,Wal92]. In categorical logic and proof theory, we should mention our own book with J. Lambek [LS86] which became popular in theoretical computer science. The category theory book of Freyd and Scedrov [FrSc] is a source book for Representation Theorems and categories of relations.

References

[ACC93] M. Abadi, L. Cardelli and P.-L. Curien, *Formal parametric polymorphism*, Theoretical Comp. Science **121** (1993), 9–58.

[Abr93] S. Abramsky, *Computational Interpretations of linear logic*, Theoretical Comp. Science **111** (1993), 3–57.

[AbJu94] S. Abramsky and A. Jung, *Domain theory*, in: [AGM], pp. 1–168.

[Abr96] S. Abramsky, *Retracing some paths in process algebra*, CONCUR '96, U. Montanari and V. Sassone, eds, Lecture Notes in Computer Science, Vol. 1119 (1996), 1–17.

[Abr97] S. Abramsky, *Semantics of interaction: An introduction to game semantics*, in: [PD97], pp. 1–31.

[ABP97] S. Abramsky, R. Blute and P. Panangaden, *Nuclear and trace ideals in tensored $*$-categories*, Preprint (1997). (To appear in *Journal of Pure and Applied Algebra.*)

[AGM] S. Abramsky, D. Gabbay and T. Maibaum eds, *Handbook of Logic in Computer Science*, Vol. 3, Oxford (1994).

[AGN] S. Abramsky, S. Gay and R. Nagarajan, *A type-theoretic approach to deadlock-freedom of asynchronous systems*, Theoretical Aspects of Computer Software, TACS'97, M. Abadi and T. Ito, eds, Lecture Notes in Computer Science, Vol. 1281 (1997).

[AHMc98] S. Abramsky, K. Honda and G. McCusker, *A fully abstract game semantics for general references*, 13th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE Press (1998), 334–344.

[AJ94a] S. Abramsky and R. Jagadeesan, *New foundations for the geometry of interaction*, Information and Computation **111** (1) (1994), 53–119. Preliminary version, in *7th Annual IEEE Symp. on Logic in Computer Science (LICS)*, IEEE Publications (1992), 211–222.

[AJ94b] S. Abramsky and R. Jagadeesan, *Games and full completeness theorem for multiplicative linear logic*, J. Symbolic Logic **59** (2) (1994), 543–574.

[AMc98] S. Abramsky and G. McCusker, *Full abstraction for idealized Algol with passive expressions*, to appear in *Theoretical Computer Science.*

[Abru91] V.M. Abrusci, *Phase semantics and sequent calculus for pure noncommutative classical linear propositional logic*, J. Symbolic Logic **56** (1991), 1403–1456.

[Ali95] M. Alimohamed, *A characterization of lambda definability in categorical models of implicit polymorphism*, Theoretical Comp. Science **146** (1995), 5–23.

[AHS95] T. Altenkirch, M. Hofmann and T. Streicher, *Categorical reconstruction of a reduction-free normalisation proof*, Proc. CTCS '95, Lecture Notes in Computer Science, Vol. 953, 182–199.

[AHS96] T. Altenkirch, M. Hofmann and T. Streicher, *Reduction-free normalisation for a polymorphic system*, Proc. of the 11th Annual IEEE Symposium of Logic in Computer Science (1996), 98–106.

[AC98] R.M. Amadio and P.-L. Curien, *Selected Domains and Lambda Calculi*, Camb. Univ. Press (1998).

[Asp] A. Asperti, *Linear logic, comonads, and optimal reductions*, to appear in *Fund. Informatica.*

[AL91] A. Asperti and G. Longo, *Categories, Types, and Structures*, MIT Press (1991). (Currently out of print but publicly available on the Web.)

[Ba97] J.C. Baez, *An introduction to n-categories*, Category Theory and Computer Science, CTCS'97, E. Moggi and G. Rosolini, eds, Lecture Notes in Computer Science, Vol. 1290 (1997).

[BDER97] P. Baillot, V. Danos, T. Ehrhard and L. Regnier, *Believe it or not: AJM games model is a model of classical linear logic*, 12th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE Press (1997).

[B72] E.S. Bainbridge, *A unified minimal realization theory, with duality*, Dissertation, Department of Computer & Communication Sciences, University of Michigan (1972).

[B75] E.S. Bainbridge, Addressed machines and duality, *Category theory applied to computation and control*, E. Manes, ed., Lecture Notes in Computer Science, Vol. 25 (1975), 93–98.

[B76] E.S. Bainbridge, *Feedback and generalized logic*, Information and Control **31** (1976), 75–96.

[B77] E.S. Bainbridge, *The fundamental duality of system theory*, Systems: Approaches, Theories, Applications, W.E. Hartnett, ed., D. Reidel (1977), 45–61.

[BFSS90] E.S. Bainbridge, P. Freyd, A. Scedrov and P.J. Scott, *Functorial polymorphism*, Theoretical Computer Science **70** (1990), 35–64.

[Bar84] H.P. Barendregt, *The Lambda Calculus*, Studies in Logic, Vol. 103, North-Holland (1984).

[Bar92] H.P. Barendregt, *Lambda calculus with types*, Handbook of Logic in Computer Science, Vol. 2, S. Abramsky, D. Gabbay and T. Maibaum, eds, Oxford Univ. Press (1992).

[Barr79] M. Barr, *$*$-Autonomous Categories*, Lecture Notes in Mathematics, Vol. 752 (1979).

[Barr91] M. Barr, *$*$-Autonomous categories and linear logic*, Math. Structures in Computer Science **1** (1991), 159–178.

[Barr95] M. Barr, *Non-symmetric $*$-autonomous categories*, Theoretical Comp. Science **139** (1995), 115–130.

[Barr96] M. Barr, Appendix to [Blu96] (1994).

[BW95] M. Barr and C. Wells, *Category Theory for Computing Science*, 2nd ed., Prentice-Hall (1995).

[Bee85] M. Beeson, *Foundations of Constructive Mathematics*, Springer, Berlin (1985).

[BAC95] R. Bellucci, M. Abadi and P.-L. Curien, *A model for formal parametric polymorphism: A PER interpretation for system $\mathcal{R}$*, Lecture Notes in Computer Science, Vol. 902, TLCA'95 (1995), 32–46.

[Ben] D.B. Benson, *Bialgebras: Some foundations for distributed and concurrent computation*, Fundamenta Informaticae **12** (1989), 427–486.

[BBPH] B.N. Benton, G. Bierman, V. de Paiva and M. Hyland, *Term assignment for intuitionistic linear logic*, Proceedings of the International Conference on Typed Lambda Calculi and Applications, M. Bezem and J.F. Groote, eds, Lecture Notes in Computer Science, Vol. 664 (1992), 75–90.

[BS91] U. Berger and H. Schwichtenberg, *An inverse to the evaluation functional for typed λ-calculus*, Proc. of the 6th Annual IEEE Symposium of Logic in Computer Science (1991), 203–211.

[BD95] I. Beylin and P. Dybjer, *Extracting a proof of coherence for monoidal categories from a proof of normalization for monoids*, Types for Proofs and Programs, TYPES'95, S. Berardi and M. Coppo, eds, Lecture Notes in Computer Science, Vol. 1158 (1995).

[Bie95] G.M. Bierman, *What is a categorical model of intuitionistic linear logic?*, TLCA'95, M. Dezani-Ciancaglini and G. Plotkin, eds (1995), 78–93.

[Bla92] A. Blass, *A game semantics for linear logic*, Annals of Pure and Applied Logic **56** (1992), 183–220.

[BE93] S.L. Bloom and Z. Ésik, *Iteration Theories; The Equational Logic of Iterative Processes*, Springer, Berlin (1993).

[BELM] S.L. Bloom, Z. Ésik, A. Labella and E. Manes, *Iteration 2-theories: Extended abstract*, Algebraic Methodology and Software Technology, AMAST'97, M. Johnson, ed., Lecture Notes in Computer Science, Vol. 1349 (1997), 30–44.

[Blu93] R. Blute, *Linear logic, coherence and dinaturality*, Theoretical Computer Science **115** (1993), 3–41.

[Blu96] R. Blute, *Hopf algebras and linear logic*, Math. Structures in Computer Science **6** (1996), 189–217.

[BCS96] R.F. Blute, J.R.B. Cockett and R.A.G. Seely, *! and ? : Storage as tensorial strength*, Math. Structures in Computer Science **6** (1996), 313–351.

[BCST96] R.F. Blute, J.R.B. Cockett, R.A.G. Seely and T.H. Trimble, *Natural deduction and coherence for weakly distributive categories*, J. Pure and Applied Algebra **113** (1996), 229–296.

[BCS97] R.F. Blute, J.R.B. Cockett and R.A.G. Seely, *Categories for computation in context and unified logic*, J. Pure and Applied Algebra **116** (1997), 49–98.

[BCS98] R. Blute, J.R.B. Cockett and R.A.G. Seely, *Feedback for linearly distributive categories: Traces and fixpoints*, Preprint, McGill University (1998).

[BS96] R.F. Blute and P.J. Scott, *Linear Läuchli semantics*, Annals of Pure and Applied Logic **77** (1996), 101–142.

[BS96b] R.F. Blute and P.J. Scott, *A noncommutative full completeness theorem (Extended abstract)*, Electronic Notes in Theoretical Computer Science **3**, Elsevier Science B.V. (1996).

[BS98] R. Blute and P.J. Scott, *The shuffle Hopf algebra and noncommutative full completeness*, J. Symbolic Logic, to appear.

[Bor94] F. Borceux, *Handbook of Categorical Algebra*, Cambridge University Press (1994).

[Brea-T,G] V. Breazu-Tannen and J. Gallier, *Polymorphic rewriting conserves algebraic normalization*, Theoretical Comp. Science **83**.

[CPS88] A. Carboni, P. Freyd and A. Scedrov, *A categorical approach to realizability and polymorphic types*, Lecture Notes in Computer Science, Vol. 298 (Proc. 3rd MFPS), 23–42.

[CSW] G.L. Cattani, I. Stark and G. Winskel, *Presheaf models for the π-calculus*, Category Theory and Computer Science, CTCS'97, E. Moggi and G. Rosolini, eds, Lecture Notes in Computer Science, Vol. 1290 (1997).

[CaWi] G.L. Cattani and G. Winskel, *Presheaf models for concurrency*, Computer Science Logic, CSL'96, D. van Dalen and M. Bezem, eds, Lecture Notes in Computer Science, Vol. 1258 (1996).

[Cob64] A. Cobham, *The intrinsic computational difficulty of functions*, Proc. of the 1964 Internations Congress for Logic, Methodology and Philosophy of Science, Y. Bar-Hillel, ed., North-Holland, Amsterdam (1964), 24–30.

[Co93] J.R.B. Cockett, *Introduction to distributive categories*, Math. Structures in Computer Science **3** (1993), 277–308.

[CS91] J.R.B. Cockett and R.A.G. Seely, *Weakly distributive categories*, J. Pure and Applied Algebra **114** (1997), 133–173. (Preliminary version in [FJP], pp. 45–65.)

[CS96a] J.R.B. Cockett and R.A.G. Seely, *Linearly distributive functors*, Preprint, McGill University, 1996.

[CS96b] J.R.B. Cockett and R.A.G. Seely, *Proof theory for full intuitionistic linear logic, bilinear logic, and* MIX *categories*, Theory and Applications of Categories **3** (1997), 85–131.

[CSp91] J.R.B. Cockett and D.L. Spencer, *Strong categorical datatypes I*, Category Theory 1991, Montreal, R.A.G. Seely, ed., CMS Conference Proceedings, Vol. 13 (1991), 141–169.

[Cook75] S.A. Cook, *Feasibly constructive proofs and the propositional calculus*, Proc. 7th Annual ACM Symp. on Theory of Computing (1975), 83–97.

[CoKa] S.A. Cook and B.M. Kapron, *Characterizations of the basic feasible functionals of finite type*, Feasible Mathematics, S. Buss and P. Scott, eds, Birkhäuser, 71–96.

[CD97] T. Coquand and P. Dybjer, *Intuitionistic model constructions and normalization proofs*, Math. Structures in Computer Science **7** (1997), 75–94.

[Coq90] T. Coquand, *On the analogy between propositons and types*, in: [Hu90], pp. 399–417.

[CU93] S.A. Cook and A. Urquhart, *Functional interpretations of feasibly constructive arithmetic*, Ann. Pure & Applied Logic **63** (2) (1993), 103–200.

[CoGa] A. Corradini and F. Gadducci, *A 2-categorical presentation of term graph rewriting*, Category Theory and Computer Science, CTCS'97, E. Moggi and G. Rosolini, eds, Lecture Notes in Computer Science, Vol. 1290 (1997).

[Cr93] R. Crole, *Categories for Types*, Cambridge Mathematical Textbooks (1993).

[Cu93] D. Čubrić, *Results in categorical proof theory*, PhD thesis, Dept. of Mathematics, McGill University (1993).

[CDS97] D. Čubrić, P. Dybjer and P.J. Scott, *Normalization and the Yoneda embedding*, Math. Structures in Computer Science **8** (2) (1997), 153–192.

[Cur93] P.-L. Curien, *Categorical Combinators, Sequential Algorithms, and Functional Programming*, Pitman (1986) (revised edition, Birkhäuser, 1993).

[D90] V. Danos, *La logique linéaire appliquée àl'étude de divers processus de normalisation et principalement du λ-calcul*, Thèse de doctorat, Univ. Paris VII, U.F.R. de Mathématiques (1990).

[DR95] V. Danos and L. Regnier, *Proof-nets and the Hilbert space*, in: [GLR95], pp. 307–328.

[DR89] V. Danos and L. Regnier, *The structure of multiplicatives*, Arch. Math. Logic **28** (1989), 181–203.

[DiCo95] R. Di Cosmo, *Isomorphism of Types: From λ-Calculus to Information Retrieval and Language Design*, Birkhäuser (1995).

[D$^+$93] G. Dowek et al., *The Coq proof assistant user's guide, Version 5.8*, Technical Report, INRIA-Rocquencourt (Feb. 1993).

[DuRe94] D. Duval and J.-C. Reynaud, *Sketches and computation I: Basic definitions and static evaluation*, Math. Structures in Computer Science **4** (2) (1994), 185–238.

[Dy95] P. Dybjer, *Internal type theory*, Types for Proofs and Programs, TYPES'95, S. Berardi and M. Coppo, eds, Lecture Notes in Computer Science, Vol. 1158 (1995).

[EK66] S. Eilenberg and G.M. Kelly, *A generalization of the functorial calculus*, J. Algebra **3** (1966), 366–375.

[EsLa] Z. Ésik and A. Labella, *Equational properties of iteration in algebraically complete categories*, MFCS'96, Lecture Notes in Computer Science, Vol. 1113 (1996).

[FRS98] F. Fages, P. Ruet and S. Soliman, *Phase semantics and verification of concurrent constraint programs*, 13th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE Press (1998), 141–152.

[FiFrL] S. Finkelstein, P. Freyd and J. Lipton, *A new framework for declarative programming*, Th. Comp. Science (to appear).

[FiPl96] M.P. Fiore and G.D. Plotkin, *An extension of models of axiomatic domain theory to models of synthetic domain theory*, CSL'96, Lecture Notes in Computer Science, Vol. 1258 (1996).

[Fl96] A. Fleury, *Thesis*, Thèse de doctorat, Univ. Paris VII, U.F.R. de Mathématiques (1996).

[FR94] A. Fleury and C. Rétoré, *The MIX rule*, Math. Structures in Computer Science **4** (1994), 273–285.

[FJP] M.P. Fourman, P.T. Johnstone and A.M. Pitts (eds), *Applications of Categories in Computer Science*, London Math. Soc. Lecture Notes 177, Camb. Univ. Press (1992).

[Fre92] P. Freyd, *Remarks on algebraically compact categories*, in: [FJP], pp. 95–107.

[Fre93] P. Freyd, *Structural polymorphism*, Theoretical Comp. Science **115** (1993), 107–129.

[FRRa] P. Freyd, E. Robinson and G. Rosolini, *Dinaturality for free*, in: [FJP], pp. 107–118.

[FRRb] P. Freyd, E. Robinson and G. Rosolini, *Functorial parametricity*, Manuscript.

[FrSc] P. Freyd and A. Scedrov, *Categories, Allegories*, North-Holland (1990).

[Frie73] H. Friedman, *Equality between functionals*, Logic Colloquium '73, Lecture Notes in Computer Science, Vol. 453 (1975), 22–37.

[GaHa] P. Gardner and M. Hasegawa, *Types and models for higher-order action calculi*, TACS'97, Lecture Notes in Computer Science, Vol. 1281 (1997).

[Ge85] R. Geroch, *Mathematical Physics*, University of Chicago Press (1985).

[Gh96] N. Ghani, *$\beta\eta$-equality for coproducts*, TLCA '95, Lecture Notes in Computer Science, Vol. 902 (1995), 171–185.

[Gi71] J.-Y. Girard, *Une extension de l'interprétation de Gödel a l'analyse, et son application a l'elimination des coupures dans l'analyse et la théorie des types*, J.E. Fenstad, ed., Proc. 2nd Scandinavian Logic Symposium, North-Holland (1971), 63–92.

[Gi72] J.-Y. Girard, *Interprétation fonctionnelle et elimination des coupures dans l'arithmétique d'ordre supérieur*, Thèse de doctorat d'état, Univ. Paris VII (1972).

[Gi86] J.-Y. Girard, *The system* **F** *of variables types, fifteen years later*, Theoretical Comp. Science **45** (1986), 159–192.

[Gi87] J.-Y. Girard, *Linear logic*, Theoretical Comp. Science **50** (1987), 1–102

[Gi88] J.-Y. Girard, *Geometry of interaction I: Interpretation of system F*, Logic Colloquium '88, R. Ferro et al., eds, North-Holland (1989), 221–260.

[Gi89] J.-Y. Girard, *Towards a geometry of interaction*, Categories in Computer Science and Logic, J.W. Gray and A. Scedrov, eds, Contemp. Math., Vol. 92, AMS (1989), 69–108.

[Gi90] J.-Y. Girard, *Geometry of interaction 2: Deadlock-free algorithms*, COLOG-88, P. Martin-Lof and G. Mints, eds, Lecture Notes in Computer Science, Vol. 417 (1990), 76–93.

[Gi91] J.-Y. Girard, *A new constructive logic: Classical logic*, Math. Structures in Computer Science **1** (1991), 255–296.

[Gi93] J.-Y. Girard, *On the unity of logic*, Annals of Pure and Applied Logic **59** (1993), 201–217.

[Gi95a] J.-Y. Girard, *Linear logic: Its syntax and semantics*, in: [GLR95], pp. 1–42.

[Gi95b] J.-Y. Girard, *Geometry of interaction III: Accommodating the additives*, in: [GLR95], pp. 329–389.

[Gi96] J.-Y. Girard, *Coherent Banach spaces*, Preprint (1996), and lectures delived at Keio University, Linear Logic '96 (April 1996).

[Gi97] J.-Y. Girard, *Light linear logic*, Preprint (1995) (revised 1997). URL: ftp://lmd.univ-mrs.fr/pub/girard/ LL.ps.Z.

[GLR95] J.-Y. Girard, Y. Lafont and L. Regnier (eds), *Advances in Linear Logic*, London Math. Soc. Series 222, Camb. Univ. Press (1995).

[GLT] J.-Y. Girard, Y. Lafont and P. Taylor, *Proofs and Types*, Cambridge Tracts in Theoretical Computer Science, Vol. 7 (1989).

[GSS91] J.-Y. Girard, A. Scedrov and P. Scott, *Normal forms and cut-free proofs as natural transformations*, Logic From Computer Science, Mathematical Science Research Institute Publications, Vol. 21 (1991), 217–241. (Also available by anonymous ftp from:theory.doc.ic.ac uk, in: papers/Scott.)

[GSS92] J.-Y. Girard, A. Scedrov and P.J. Scott, *Bounded linear logic*, Theoretical Comp. Science **97** (1992), 1–66.

[Giry] M. Giry, *A categorical approach to probability theory*, Proc. of a Conference on Categorical Aspects of Topology, Lecture Notes in Mathematics, Vol. 915, (1981), 68–85.

[GAL92] G. Gonthier, M. Abadi and J.J. Levy, *Linear logic without boxes*, 7th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE Publications (1992), 223–234.

[GPS96] R. Gordon, A. Power and R. Street, *Coherence for tricategories*, Memoirs of the American Mathematical Society (1996).

[GS89] J. Gray and A. Scedrov (eds), *Categories in Computer Science and Logic*, Contemp. Math., Vol. 92, AMS (1989).

[Gun92] C. Gunter, *Semantics of Programming Languages*, MIT Press (1992).

[Gun94] C. Gunter, *The semantics of types in programming languages*, in: [AGM], pp. 395–475.

[Ha87] T. Hagino, *A typed lambda calculus with categorical type constructors*, Category Theory and Computer Science, Lecture Notes in Computer Science, Vol. 283 (1987), 140–157.

[Hagh99] E. Haghverdi, *Unique decomposition categories, geometry of interaction and combinatory logic*, Math. Struct. in Comput. Sci. (to appear).

[Ham99] M. Hamano, *Pontrjagin duality and a full completeness of multiplicative linear logic without MIX*, Math. Struct. in Comput. Sci. (to appear).

[Har93] T. Hardin, *How to get confluence for explicit substitutions*, Term Graph Rewriting, M. Sleep, M. Plasmeijer and M. van Eekelen, eds, Wiley (1993), 31–45.

[HM92] V. Harnik and M. Makkai, *Lambek's categorical proof theory and Läuchli's abstract realizability*, J. Symbolic Logic **57** (1992), 200–230.

[MHas97] M. Hasegawa, *Recursion from cyclic sharing: Traced monoidal categories and models of cyclic lambda calculi*, TLCA '97, Lecture Notes in Computer Science, Vol. 1210 (1997), 196–213.

[RHas94] R. Hasegawa, *Categorical data types in parametric polymorphism*, Math. Structures in Computer Science **4** (1) (1994), 71–109.

[RHas95] R. Hasegawa, *A logical aspect of parametric polymorphism*, Computer Science Logic, CSL'95, H.K. Büning, ed., Lecture Notes in Computer Science, Vol. 1092 (1995).

[Haz] M. Hazewinkel, *Introductory recommendations for the study of Hopf algebras in mathematics and physics*, CWI Quarterly, Centre for Mathematics and Computer Science, Amsterdam **4** (1) (March 1991).

[HJ95] C. Hermida and B. Jacobs, *Fibrations with indeterminates: Contextual and functional completeness for polymorphic lambda calculi*, Math. Structures in Computer Science **5** (4), 501–532.

[HMP98] C. Hermida, M. Makkai and J. Power, *Higher dimensional multigraphs*, 13th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE (1998), 199–206.

[Hi80] D. Higgs, *Axiomatic Infinite Sums–An Algebraic Approach to Integration Theory*, Contemporary Mathematics, Vol. 2 (1980).

[HR89] D.A. Higgs and K.A. Rowe, *Nuclearity in the category of complete semilattices*, J. Pure and Applied Algebra **5** (1989), 67–78.

[HPW] T. Hildebrandt, P. Panangaden and G. Winskel, *A relational model of non-deterministic dataflow*, to appear in: Concur 98, Lecture Notes in Computer Science, Vol. ??.

[HM94] J. Hodas and D. Miller, *Logic programming in a fragment of intuitionistic linear logic*, Inf. and Computation **110** (2) (1994), 327–365.

[H97a] M. Hofmann, *Syntax and semantics of dependent types*, in: [PD97], pp. 79–130.

[H97b] M. Hofmann, *An application of category-theoretic semantics to the characterisation of complexity classes using higher-order function algebras*, Bull. Symb. Logic **3** (4) (1997), 469–485.

[HJ97] H. Hu and A. Joyal, *Coherence completions of categories and their enriched softness*, Proceedings, Mathematical Foundations of Programming Semantics, Thirteenth Annual Conference, S. Brookes and M. Mislove, eds, Electronic Notes in Theoretical Computer Science **6** (1997).

[Hu90] G. Huet (ed.), *Logical Foundations of Functional Programming*, Addison-Wesley (1990).

[Hy88] M. Hyland, *A small complete category*, Ann. Pure and Appl. Logic **40** (1988), 135–165.

[Hy97] M. Hyland, *Game semantics*, in: [PD97], pp. 131–184.

[HRR] M. Hyland, E. Robinson and G. Rosolini, *The discrete objects in the effective topos*, Proc. Lond. Math. Soc. Ser. **60** (1990), 1–36.

[HP93] M. Hyland and V. de Paiva, *Full intuitionistic linear logic (Extended Abstract)*, Annals of Pure and Applied Logic **64** (3) (1993), 273–291.

[JR] B. Jacobs and J. Rutten, *A tutorial on (Co)Algebras and (Co)Induction*, EATCS Bulletin **62** (1997), 222–259.

[J90] C.B. Jay, *The structure of free closed categories*, J. Pure and Applied Algebra **66** (1990), 271–285.

[JGh95] C.B. Jay and N. Ghani, *The virtues of eta expansion*, Journal of Functional Programing **5** (2) (1995), 135–154.

[JNW] A. Joyal, M. Nielsen and G. Winskel, *Bisimulation from open maps*, Inf. and Computation **127** (2) (1996), 164–185.

[JS91b] A. Joyal and R. Street, *An introduction to Tannaka duality and quantum groups*, Category Theory, Proceedings, Como 1990, A. Carboni et al., eds, Lecture Notes in Mathematics, Vol. 1488 (1991), 411–492.

[JS91] A. Joyal and R. Street, *The geometry of tensor calculus I*, Advances in Mathematics **88** (1991), 55–112.

[JS93] A. Joyal and R. Street, *Braided tensor categories*, Advances in Mathematics **102** (1) (1993), 20–79.

[JSV96] A. Joyal, R. Street and D. Verity, *Traced monoidal categories*, Math. Proc. Camb. Phil. Soc. **119** (1996), 447–468.

[JT93] A. Jung and J. Tiuryn, *A new characterization of lambda definability*, TLCA'93, Lecture Notes in Computer Science, Vol. 664, 230–244.

[Ka95] M. Kanovich, *The direct simulation of Minksky machines in linear logic*, in: [GLR95], pp. 123–146.

[KOS97] M. Kanovich, M. Okada and A. Scedrov, *Phase semantics for light linear logic*, Proceedings, Mathematical Foundations of Programming Semantics, Thirteenth Annual Conference, S. Brookes and M. Mislove, eds, Electronic Notes in Theoretical Computer Science **6** (1997).

[KSW95] P. Katis, N. Sabadini and R.F.C. Walters, *Bicategories of processes*, Manuscript (1995). Dept. of Mathematics, Univ. Sydney.

[KSWa] P. Katis, N. Sabadini and R.F.C. Walters, *Span(Graph): A categorical algebra of transition systems*, AMAST'97, Lecture Notes in Computer Science, Vol. 1349 (1997).

[K77] G.M. Kelly, *Many-variable functorial calculus. I*, Coherence in Categories, Lecture Notes in Mathematics, Vol. 281 (1977), 66–105.

[K82] G.M. Kelly, *Basic Concepts of Enriched Category Theory*, London Math. Soc. Lecture Notes, Vol. 64, Camb. Univ. Press (1982).

[KM71] G.M. Kelly and S. Mac Lane, *Coherence in closed categories*, J. Pure and Applied Algebra **1** (1971), 97–140.

[KOPTT] Y. Kinoshita, P.W. O'Hearn, A.J. Power, M. Takeyama and R.D. Tennent, *An axiomatic approach to binary logical relations with applications to data refinement*, Lecture Notes in Computer Science, Vol. 1281, TACS '97 (1997), 191–212.

[Ku63] A.G. Kurosh, *Direct decomposition in algebraic categories*, Amer. Math. Soc. Transl. Ser. 2, Vol. 27 (1963), 231–255.

[Lac95] S.G. Lack, *The algebra of distributive and extensive categories*, PhD thesis, University of Cambridge (1995).

[Laf88] Y. Lafont, *Logiques, catégories et machines*, Thèse de doctorat, Université Paris VII, U.F.R. de Mathématiques (1988).

[Laf95] Y. Lafont, *From proof nets to interaction nets*, in: [GLR95], pp. 225–247.

[L68] J. Lambek, *Deductive systems and categories I*, J. Math. Systems Theory **2**, 278–318.

[L69] J. Lambek, *Deductive systems and categories II*, Lecture Notes in Mathematics, Vol. 87, Springer, Berlin (1969).

[L74] J. Lambek, *Functional completeness of cartesian catgories*, Ann. Math. Logic **6**, 259–292.

[L88] J. Lambek, *On the unity of algebra and logic*, Categorical Algebra and its Applications, F. Borceux, ed., Lecture Notes in Mathematics, Vol. 1348, Springer (1988).

[L89] J. Lambek, *Multicategories revisited*, Contemp. Math. **92**, 217–239.

[L90] J. Lambek, *Logic without structural rules*, Manuscript, Mcgill University (1990).

[L93] J. Lambek, *From categorial grammar to bilinear logic*, Substructural Logics, Studies in Logic and Computation, Vol. 2, K. Došen and P. Schroeder-Heister, eds, Oxford Science Publications (1993).

[L95] J. Lambek, *Bilinear logic in algebra and linguistics*, in: [GLR95].

[LS86] J. Lambek and P.J. Scott, *Introduction to Higher Order Categorical Logic*, Cambridge Studies in Advanced Mathematics, Vol. 7, Cambridge University Press (1986).

[Lau] H. Läuchli, *An abstract notion of realizability for which intuitionistic predicate calculus is complete*, Intuitionism and Proof Theory, North-Holland (1970), 227–234.

[Law66] F.W. Lawvere, *The category of categories as a foundation for mathematics*, Proc. Conf. on Categorical Algebra, S. Eilenberg et al., eds, Springer (1965).

[Law69] F.W. Lawvere, *Adjointness in foundations*, Dialectica **23**, 281–296.

[Lef63] S. Lefschetz, *Algebraic Topology*, Amer. Math. Soc. Colloq. Publ., Vol. 27 (1963).

[Li95] P. Lincoln, *Deciding provability of linear logic formulas*, in: [GLR95], pp. 109–122.

[Luo94] Z. Luo, *Computation and Reasoning: A Type Theory for Computer Science*, Oxford Univ. Press (1994).

[LP92] Z. Luo and R. Pollack, *LEGO proof development system: User's manual*, Technical Report ECS-LFCS-92-211, Univ. Edinburgh (1992).

[MaRey] Q. Ma and J.C. Reynolds, *Types, abstraction, and parametric polymorphism, Part 2*, Lecture Notes in Computer Science, Vol. 598 (1991), 1–40.

[Mac71] S. Mac Lane, *Categories for the Working Mathematician*, Springer (1971).

[Mac82] S. Mac Lane, *Why commutative diagrams coincide with equivalent proofs*, Contemp. Math. **13** (1982), 387–401.

[MM92] S. Mac Lane and I. Moerdijk, *Sheaves in Geometry and Logic*, Springer (1992).

[MaRe91] P. Malacaria and L. Regnier, *Some results on the interpretation of λ-calculus in operator algebras*, 6th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE Press (1991), 63–72.

[MA86] E. Manes and M. Arbib, *Algebraic Approaches to Program Semantics*, Springer (1986).

[M98] E. Manes, *Implementing collection classes with monads*, Math. Structures in Computer Science **8** (3), 231–276.

[ML73] P. Martin-Löf, *An intuitionistic theory of types: Predicative part*, Logic Colloquium '73, H.E. Rose and J.C. Shepherdson, eds, North-Holland (1975), 73–118.

[ML82] P. Martin-Löf, *Constructive mathematics and computer programming*, Proc. Sixth International Congress for Logic, Methodology and Philosophy of Science, North-Holland (1982).

[Mc97] G. McCusker, *Games and definability for FPC*, Bull. Symb. Logic **3** (3) (1997), 347–362.

[MPSS95] N. Mendler, P. Panangaden, P. Scott and R. Seely, *The logical structure of concurrent constraint languages*, Nordic Journal of Computing **2** (1995), 181–220.

[Mét94] F. Métayer, *Homology of proof nets*, Arch. Math Logic **33** (1994) (see also [GLR95]).

[Mill] D. Miller, *Linear logic as logic programming: An abstract*, Logical Aspects of Computational Linguistics, LACL'96, C. Retoré, ed., Lecture Notes in Artificial Intelligence, Vol. 1328 (1996).

[MNPS] D. Miller, G. Nadathur, F. Pfenning and A. Scedrov, *Uniform proofs as a foundation for logic programming*, Ann. Pure and Applied Logic **51** (1991), 125–157.

[Mil89] R. Milner, *Communication and Concurrency*, Prentice-Hall (1989).

[Mil77] R. Milner, *Fully abstract models of typed lambda calculi*, Theoretical Comp. Science **4** (1977), 1–22.

[Mil93a] R. Milner, *The polyadic π-calculus: A tutorial*, Logic and Algebra of Specification, F.L. Bauer, W. Brauer and H. Schwichtenberg, eds, Springer (1993).

[Mil93b] R. Milner, *Action calculi, or concrete action structures*, Mathematical Foundations of Computer Science, Lecture Notes in Computer Science, Vol. 711 (1993), 105–121.

[Mil94] R. Milner, *Action calculi V: Reflexive molecular forms*, Manuscript, University of Edinburgh (June, 1994).

[Min81] G.E. Mints, *Closed Categories and the theory of proofs*, J. Soviet Math. **15** (1981), 45–62.

[Mit96] J.C. Mitchell, *Foundations for Programming Languages*, MIT Press (1996).

[Mit90] J.C. Mitchell, *Type systems for programming languages*, Handbook of Theoretical Computer Science, Vol. B (Formal Models and Semantics), J. Van Leeuwen, ed., North-Holland (1990), 365–458.

[MitSce] J.C. Mitchell and A. Scedrov, *Notes on sconing and relators*, Lecture Notes in Computer Science, Vol. 702, CSL'92 (1992), 352–378.

[MS89] J.C. Mitchell and P.J. Scott, *Typed lambda models and Cartesian closed categories*, in: [GS89], pp. 301–316.

[MiVi] J.C. Mitchell and R. Visvanathan, *Effective models of polymorphism, subtyping and recursion (Extended abstract)*, Automata, Languages and Programming, ICALP'96, F.M. auf der Heide and B. Monien, eds, Lecture Notes in Computer Science, Vol. 1099 (1996).

[Mo91] E. Moggi, *Notions of computation and monads*, Inf. & Computation **93** (1) (1991), 55–92.

[Mo97] E. Moggi, *Metalanguages and applications*, in: [PD97], pp. 185–239.

[MR97] E. Moggi and G. Rosolini (eds), *7th International Conference on Category Theory and Computer Science*, Lecture Notes in Computer Science, Vol. 1290 (1997).

[Mul91] P. Mulry, *Strong monads, algebras and fixed points*, in: [FJP], pp. 202–216.

[OHT92] P. O'Hearn and R.D. Tennent, *Semantics of local variables*, Applications of Categories in Computer Science, London Math. Soc. Lecture Series, Vol. 177, Camb. Univ. Press (1992), 217–236.

[OHT93] P. O'Hearn and R.D. Tennent, *Relational parametricity and local variables*, 20th Symp. on Principles of Programming Languages, Assoc. Comp. Machinery (1993), 171–184.

[OHRi] P. O'Hearn and J. Riecke, *Kripke logical relations and PCF*, to appear in *Information and Computation.*

[Ol85] F.J. Oles, *Type algebras, functor categories and block structure*, Algebraic Methods in Semantics, M. Nivat and J. Reynolds, eds, Camb. Univ. Press (1985).

[PSSS] P. Panangaden, V. Saraswat, P.J. Scott and R.A.G. Seely, *A categorical view of concurrent constraint programming*, Proceedings of REX Workshop, Lecture Notes in Computer Science, Vol. 666 (1993), 457–476.

[PR89] R. Paré and L. Román, *Monoidal categories with natural numbers object*, Studia Logica **48** (1989), 361–376.

[P96] D. Pavlović, *Categorical logic of names and abstraction in action calculi*, Math. Structures in Computer Science **7** (6) (1997), 619–637.

[PaAb] D. Pavlovič and S. Abramsky, *Specifying interaction categories*, CTCS'97, E. Moggi and G. Rosolini, eds, Lecture Notes in Computer Science, Vol. 1290 (1997).

[Pen93] M. Pentus, *Lambek grammars are context free*, 8th Annual IEEE Symp. on Logic in Computer Science (LICS), Montreal (1993), 429–433.

[PDM89] B. Pierce, S. Dietzen and S. Michaylov, *Programming in higher-order typed lambda-calculi*, Reprint CMU-CS-89-111, Dept. of Computer Science, Carnegie-Mellon University (1989).

[Pi87] A. Pitts, *Polymorphism is set-theoretic, constructively*, Category Theory and Computer Science, Lecture Notes in Computer Science, Vol. 283, D.H. Pitt, ed. (1987), 12–39.

[Pi96a] A. Pitts, *Relational properties of domains*, Inf. and Computation **127** (2) (1996), 66–90.

[Pi96b] A. Pitts, *Reasoning about local variables with operationally-based logical relations*, 11th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE Press (1996), 152–163.

[Pi97] A. Pitts, *Operationally-based theories of program equivalence*, in: [PD97], pp. 241–298.

[Pi9?] A.M. Pitts, *Categorical logic*, Handbook of Logic in Computer Science, Vol. 6, S. Abramsky, D.M. Gabbay and T.S.E. Maibaum, eds, Oxford University Press (1996) (to appear).

[PD97] A. Pitts and P. Dybjer (eds), *Semantics and Logics of Computation*, Publications of the Newton Institute, Camb. Univ. Press (1997).

[Plo77] G.D. Plotkin, *LCF considered as programming language*, Theoretical Comp. Science **5** (1977), 223–255.

[Plo80] G.D. Plotkin, *Lambda definability in the full type hierarchy*, To H.B. Curry: Essays on Combinatory Logic, Lambda Calculus, and Formalism, Academic Press (1980), 363–373.

[PlAb93] G. Plotkin and M. Abadi, *A logic for parametric polymorphism*, TLCA'93, Lecture Notes in Computer Science, Vol. 664 (1993), 361–375.

[Pow89] J. Power, *A general coherence result*, J. Pure and Applied Algebra **57** (1989), 165–173.

[PK96] J. Power and Y. Kinoshita, *A new foundation for logic programming*, Extensions of Logic Programming '96, Springer (1996).

PowRob97] A.J. Power and E.P. Robinson, *Premonoidal categories and notions of computation*, Math. Structures in Computer Science **7** (5) (1997), 453–468.

[PowTh97] J. Power and H. Thielecke, *Environments, continuation semantics and indexed categories*, TACS'97, Lecture Notes in Computer Science, Vol. 1281 (1997), 191–212.

[Pra95] V. Pratt, *Chu spaces and their interpretation as concurrent objects*, Computer Science Today, J. van Leeuwen, ed., Lecture Notes in Computer Science, Vol. 1000 (1995), 392–405.

[Pra97] V. Pratt, *Towards full completeness for the linear logic of Chu spaces*, Proc. MFPS (MFPS'97), Electronic Notes of Theoretical Computer Science (1997).

[Pr65] D. Prawitz, *Natural Deduction*, Almquist & Wilksell, Stockhom (1965).

[Rgn92] L. Regnier, *Lambda-calcul et réseaux*, Thèse de doctorat, Univ. Paris VII, U.F.R. Mathématiques (1992).

[ReSt97] B. Reus and Th. Streicher, *General synthetic domain theory – A logical approach (Extended abstract)*, CTCS'97, Lecture Notes in Computer Science, Vol. 1290 (1997).

[Rey74] J. Reynolds, *Towards a theory of type structure*, Programming Symposium, Lecture Notes in Computer Science, Vol. 19 (1974), 408–425.

[Rey81] J.C. Reynolds, *The essence of Algol*, Algorithmic Languages, J.W. de Bakker and J.C. van Vliet, eds, North-Holland (1981), 345–372.

[Rey83] J.C. Reynolds, *Types, abstraction, and parametric polymorphism*, Information Processing '83, R.E.A. Mason, ed., North-Holland (1983), 513–523.

[RP] J. Reynolds and G. Plotkin, *On functors expressible in the polymorphic typed lambda calculus*, Inf. and Computation, reprinted in: [Hu90], pp. 127–152.

[Rob89] E. Robinson, *How complete is PER?*, 4th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE Publications (1989), 106–111.

[RR90] E. Robinson and G. Rosolini, *Polymorphism, set theory, and call-by-value*, 5th Annual IEEE Symp. on Logic in Computer Science (LICS), IEEE Publications (1990), 12–18.

[Rom89] L. Roman, *Cartesian categories with natural numbers object*, J. Pure and Applied Algebra **58** (1989), 267–278.

[Ros90] G. Rosolini, *About modest sets*, Int. J. Found. Comp. Sci. **1** (1990), 341–353.

[Sc93] A. Scedrov, *A brief guide to linear logic*, Current Trends in Theoretical Computer Science, G. Rozenberg and A. Salomaa, eds, World Scientific Publishing Co. (1993), 377–394.

[Sc95] A. Scedrov, *Linear logic and computation: A survey*, Proof and Computation, H. Schwichtenberg, ed., NATO Advanced Science Institutes, Series F, Vol. 139, Springer, Berlin (1995), 379–395.

[Sc72] D. Scott, *Continuous lattices*, Lecture Notes in Mathematics, Vol. 274, 97–136.

[See87] R.A.G. Seely, *Categorical semantics for higher order polymorphic lambda calculus*, J. Symb. Logic **52** (1987), 969–989.

[See89] R.A.G. Seely, *Linear logic, $*$-autonomous categories and cofree coalgebras*, J. Gray and A. Scedrov, eds, Categories in Computer Science and Logic, Contemporary Mathematics, Vol. 92, Amer. Math. Soc. (1989).

[Sie92] K. Sieber, *Reasoning about sequential functions via logical relations*, in: [FJP], pp. 258–269.

[Si93] A.K. Simpson, *A characterization of least-fixed-point operator by dinaturality*, Theoretical Comp. Science **118** (1993), 301–314.

[SP82] M.B. Smyth and G.D. Plotkin, *The category-theoretic solution of recursive domain equations*, SIAM J. Computing **11** (1982), 761–783.

[So87] S. Soloviev, *On natural transformations of distinguished functors and their superpositions in certain closed categories*, J. Pure and Applied Algebra **47** (1987), 181–204.

[So95] S. Soloviev, *Proof of a S. MacLane conjecture (Extended abstract)*, CTCS'95, Lecture Notes in Computer Science, Vol. 953, 59–80.

[St96] R.F. Stärk, *Call-by-value, call-by-name and the logic of values*, Computer Science Logic, CSL'96, D. van Dalen and M. Bezem, eds, Lecture Notes in Computer Science, Vol. 1258 (1996).

[St85] R. Statman, *Logical relations and the typed lambda calculus*, Inf. and Control **65** (1985), 85–97.

[St96] R. Statman, *On Cartesian monoids*, CSL'96, Lecture Notes in Computer Science, Vol. 1258 (1996).

[ST96] R. Street, *Categorical structures*, Handbook of Algebra, Vol. 1, M. Hazewinkel, ed., Elsevier (1996).

[Tay98] P. Taylor, *Practical Foundations*, Camb. Univ. Press (1998).

[Ten94] R.D. Tennent, *Denotational semantics*, in: [AGM], pp. 169–322.

[TP96] D. Turi and G. Plotkin, *Towards a mathematical operational semantics*, 12th Annual IEEE Symp. on Logic in Computer Science (LICS) (1996), 280–291.

[Tr92] A. Troelstra, *Lectures on Linear Logic*, CSLI Lecture Notes, Vol. 29 (1992).

[TrvD88] A.S. Troelstra and D. van Dalen, *Constructivism in Mathematics*, Vols. I and II, North-Holland (1988).

[W92] P. Wadler, *The essence of functional programming*, Symp. on Principles of Programming Languages (POPL'92), ACM Press, 1–14.

[W94] P. Wadler, *A syntax for linear logic*, Lecture Notes in Computer Science, Vol. 802, Springer, Berlin (1994), 513–528.

[WW86] K. Wagner and G. Wechsung, *Computational Complexity*, D. Reidel (1986).

[Wal92] R.F.C. Walters, *Categories and Computer Science*, Camb. Univ. Press (1992).

[WN97] G. Winskel and M. Nielsen, *Categories in concurrency*, in: [PD97], pp. 299–354.

[Wr89] G. Wraith, *A note on categorical datatypes*, Category Theory in Computer Science, Lecture Notes in Computer Science, Vol. 389 (1989), 118–127.

[Y90] D. Yetter, *Quantales and (noncommutative) linear logic*, Journal of Symbolic Logic **55** (1990), 41–64.

This Page Intentionally Left Blank

Algebra, Categories and Databases

B. Plotkin

Institute of Mathematics, Hebrew University, Givat Ram, Jerusalem, Israel

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

Introduction

1. This paper is in some sense, the survey of connections between algebra and databases. We will discuss mainly an algebraic model of DB's[1] and applications of this model. This paper is far from being a survey on DBs themselves. Speaking about algebra we mean universal algebra and model theory, and also logic, in particular, algebraic logic, and theory of categories. Groups which appear here as groups of automorphisms of DBs play a significant role and are used for solving specific problems. For example, Galois theory of DBs is used for solving the problem of equivalence of two DBss.

We are interested in two sides of the relations considered. One side is the evaluation of the possibilities of algebra in the theory of DBs, and another deals with understanding of the fact, that algebra itself can profit much from these links. This survey is devoted only to the first side, however the second one is also visible to some extent. At last, it is clear, that modern algebra gets a lot of benefits from its various applications.

2. A database is an information system which allows to store and process information as well as to query its contents. It is possible to query either the data directly, stored in the DB, or some information which can be somehow derived from the basic data. It is assumed that the derived information is expressed in terms of basic data, through some algebraic means.

It is customary to distinguish three main traditional approaches in DBs. These are hierarchical, network and relational approaches. We work with relational DBs, since in fact, up to now, they display the interaction between universal algebra, logic, category theory and databases in the best way.

The relational approach was discovered by E. Codd in the early 70s [Co1,Co2,Co3,Co4, Co5]. In this approach, information is stored in the form of relations, replies on queries are constructed in a special relational algebra, called algebra of relations. There is also queries language which is compatible with the operations of relational algebra. This efficient way of presenting queries and replies is one of the main advantages of the relational approach to DB organization.

However, the language of queries in this theory is not sufficiently natural. It is adapted to the calculation of a reply to query, but it is not very convenient for the description of the query. We find that the traditional language of first-order logic is more natural for this purpose.

The second problem of relational approach deals with the algebra of replies. The relational algebra is defined not axiomatically. This means that there is no good connection between queries and replies. Recently, J. Cirulis has found an axiomatic approach [Ci1] to Codd's algebras.

Perhaps, the situation would change, if we proceed from the ideas of algebraic logic. In early 80's applications of algebraic logic in databases were outlined. In the paper [IL] transitions between Codd's relational algebras and cylindric Tarski algebras were studied, in [BPl1,BPl2] polyadic Halmos algebras were used. These transitions allow us to work on the level of modeling with machinery of algebraic logic, and then transfer to the language and algebras of Codd.

[1] "DB" stands for "database".

3. In the theory of relational DBs there are two main directions, namely deductive approach and model-theoretic approach. In this survey we pay attention, mainly, to the model-theoretic approach (some connections with the deductive one are sketched). In particular, a DB state is considered as a model in the sense of model theory. It has the form (D, Φ, f), where D is a data algebra, Φ is a set of symbols of relations and f is a realization of this set Φ in D. Algebra, logic and model theory are applied in various specific problems of DBs theory, for instance, in problems related to the modeling of semantics, in the theory of functional dependences, in construction of normal forms. There are new algebraic approaches to the theory of computations, which are based on categorical ideas and notions of monad and comonad. Of course, algebra is used also as an algebra of relations.

This survey is devoted, as it was already mentioned, to the development of a DB model. In this model a DB appears as an algebraic structure. Thus, we can speak about the category of DBs depending on some data type. In the framework of this category it is possible to treat many problems. Among them the problem of comparing of different DBs, the problem of equivalence of two DBs and so on.

4. In the theory and practice of DBs three big parts can be singled out. These are semantics modeling problems, computational problems and creation of various DBMSs.[2] Semantics modeling in DBs is described in the book of Tsalenko [Ts3]. In this book there is a complete survey of this topic, accompanied by a rich bibliography, which allows to follow the development of the direction. Besides, in the book there are algebraic results. In particular, a useful algebra of files is constructed. In this book can be found important algebraic results obtained by E. Beniaminov [Ben1,Ben2,Ben3,Ben4], V. Borschew and M. Khomjakov [BK].

Theory of computations has become a very significant branch of mathematics. Algebra and categories play a special role in this theory, see papers of D. Scott [Sc,GS], G. Plotkin [GPl1,GPl2,GPl3] and many others. An important series of works in this field with applications to DBs was carried out by J. Gurevich [Gur1,Gur2]. In these works all DBs are finite and investigations are combined with the theory of finite models. There are collections of papers [Cat,Appl] dedicated to applications of categorical theory in computer science, in particular, in DBs. See also [AN,BN,GB1,GB2,Go,GoTWW,GoTW,GS,Ja,Ka,MG,Ol].

There exists an almost boundless bibliography on various DBMSs. See, for example, the book [Sy] and references included.

The set of textbooks on DBs is well known. There are the books of J. Ullman [Ulm], C. Date [Da], J. Martin [Ma], D. Maier [Ma] and others. We would like to mention in particular the book of P. Gray "Logic, Algebra and Databases" [Gr]. See also [AL,LS].

5. Let us sketch the main ideas, which lie at the base of the algebraic model of DB. We assume that a query is expressed as formula of first-order logic. In a query there are also functional symbols, which are connected with operations in data algebras. These operations satisfy some identities, which are used for the verifying of correctness of operations fulfillment. The identities determine some data type, [Ag,Za], whose algebraic equivalent

---

2 "DBMS" stands for database management system.

is a variety Θ, [BPl3]. Thus, we use not necessarily only pure logic of first-order but also some Θ-logic, which is associated with data type Θ. To this logic corresponds a variety of algebras L_Θ. This L_Θ may be a variety of cylindric algebras [HMT], or polyadic algebras [Hal,BPl3].

For each algebra $D \in \Theta$ there is an algebra V_D from the variety L_Θ, which serves as an algebra of replies to queries. It plays the role of Codd's relational algebra. Each query is written by means of Θ-logic, however one should distinguish a query and its notation. It is clear that one and the same query can be expressed in different ways. This means that a query must be considered as a class of equivalent notations. The set of such classes form an universal algebra of queries U, which also belongs to the variety L_Θ. Elements from U are also formulas, but considered up to a certain equivalence.

U and V_D are the objects of algebraic logic. To each state (D, Φ, f), $D \in \Theta$ of a DB corresponds a canonical homomorphism $\hat{f}: U \to V_D$. The reply to a query u in the state f is $\hat{f}(u)$. This idea lies at the base of DB model construction. In a first approximation, a DB is considered as a triple (F, Q, R) where F is a set of feasible states, and Q and R are algebras of queries and replies respectively, which belong to one and the same variety L_Θ. There is an operation $*: F \times Q \to R$, such that $f * q$, $f \in F$, $q \in Q$ is a reply to a query q in the state f. Let $\hat{f}: Q \to R$ be a map, defined by the formula $\hat{f}(q) = f * q$. We assume that this map is a homomorphism of algebras in L_Θ. The last means that the structure of a reply is well-coordinated with the structure of a query. The triple (F, Q, R) has to be compatible with the data algebra D. In particular, R is a subalgebra in V_D, and, besides that, the algebra Q is a result of compression of the query algebra U, depending on the choice of the set F. The set F can be given by a set of axioms $T \subset U$. It also can be defined in other ways.

This is the model-theoretic approach to DBs. Its combination with deductive ideas allows to realize a reply by deductive means. This, in its turn, gives connection with syntax and computations based on it. DB model construction can be grounded also on the categorical approach of F. Lawvere to first-order logic [Law1,Law2,Law3]. Observe here, that DB models uses complicated structures of algebra, algebraic logic and category theory. We will see, that results obtained with the help of this model may serve as a tool for quite practical applications in real DBs.

This survey consists of three sections. The first section describes necessary algebraic machinery, in the second one the DB model is constructed. The last section is devoted to applications and other questions. In particular, applications of monads and comonads in DBs are considered. Monads are used for enrichment of DB structure, while comonads work in the model of computations.

We essentially use the book [BPl3]. The results taken from it are given without proofs. New results are provided, as a rule, by proofs.

Not only relational DBs use the theory of categories. During the last decade a plenitude of works on this topic appeared (see [ABCD] and bibliography included).

In preparation of this survey I obtained significant help from J. Cirulis, Z. Diskin, E. & T. Plotkin. I wish to thank also L. Bokut and B. Trakhtenbrot for their useful remarks.

1. Algebra for databases

1.1. *Many-sorted algebra*

1.1.1. *The main notions.* In DBs *many-sorted sets* and *many-sorted algebras* are considered. "Many-sorted" means not necessarily one-sorted. Systems of data sets or data algebras are usually many-sorted. Algebraic operations on data appear when we deals with characteristics of objects. These characteristics are computed by formulas, using other, possibly observable, characteristics. This is the typical situation.

Many-sorted algebras are met frequently in algebra. For instance, affine space is a two-sorted algebra, a group together with its action on some linear space is also a two-sorted algebra. An automaton gives an example of three-sorted algebra, and so on.

The general point of view on algebra as many-sorted algebra was suggested by A. Mal'cev, P. Higgins, G. Birkhoff and others (see [Mal1,Hig,BL,BW]). G. Birkhoff has called such algebras *heterogeneous*. Let us pass to general definitions.

First of all, fix a set of sorts Γ. Correspondingly, we consider the many-sorted set $D = (D_i,\ i \in \Gamma)$. Each D_i is a domain of the sort i. Denote by Ω a set of symbols of operations. Each $\omega \in \Omega$ has a definite type $\tau = \tau(w) = (i_1, \dots, i_n;\ j)$, where $i, j \in \Gamma$. Such a symbol ω is realized in D as an *operation*, i.e. a map

$$\omega : D_{i_1} \times \cdots \times D_{i_n} \to D_j.$$

In each algebra D all symbols $\omega \in \Omega$ are assumed to be realized as operations. So, D is an Ω-algebra.

Given Γ and Ω, *morphisms* of algebras $D \to D'$ have the form

$$\mu = (\mu_i,\ i \in \Gamma) : D = (D_i,\ i \in \Gamma) \to D' = (D'_i,\ i \in \Gamma),$$

where each $\mu_i : D_i \to D'_i$ is a map of sets. *Homomorphisms* of algebras are morphisms, which preserve operations. For $\mu = (\mu_i,\ i \in \Gamma)$ and $\omega \in \Omega$ this means that if $\tau(\omega) = (i_1, \dots, i_n;\ j)$ and $a_1 \in D_{i_1}, \dots, a_n \in D_{i_n}$, then

$$(a_1 \dots a_n \omega)^{\mu_j} = a_1^{\mu_{i_1}} \dots a_n^{\mu_{i_n}} \omega.$$

The multiplication of such many-sorted maps is defined componentwise, and the product of homomorphisms is a homomorphism. Sometimes instead of a^{μ_i} we write a^{μ}. If all μ_i are surjections (injections), then μ is also surjection (injection). If all μ_i are bijections, then μ is a bijection and $\mu^{-1} = (\mu_i^{-1},\ i \in \Gamma)$. A bijective homomorphism μ is an isomorphism.

The *kernel* of a homomorphism $\mu : D \to D'$ has the form $\rho = (\rho_i,\ i \in \Gamma)$, where each ρ_i is the kernel equivalence of the map $\mu_i : D_i \to D'_i$. *A many-sorted equivalence* ρ is a *congruence* if it preserves all operations $\omega \in \Omega$. This means that if $\tau(\omega) = (i_1, \dots, i_n;\ j)$, $a_1, a'_1 \in D_{i_1}, \dots, a_n, a'_n \in D_{i_n}$ then

$$a_1 \rho_{i_1} a'_1, \dots, a_n \rho_{i_n} a'_n \Rightarrow (a_1 \dots a_n \omega) \rho_j (a'_1 \dots a'_n \omega).$$

If ρ is a congruence in D, then one can consider the factor algebra $D/\rho = (D_i/\rho_i,\ i \in \Gamma)$. The notions of subalgebra and Cartesian product of algebras are defined in a natural way. If, for example, $D^\alpha = (D_i^\alpha,\ i \in \Gamma)$ are Ω-algebras, $\alpha \in I$, then

$$\prod D^\alpha = \left(\prod_\alpha D_i^\alpha,\ i \in \Gamma\right)$$

and for each $\tau(\omega) = (i_1, \ldots, i_n; j)$, if $a_1 \in \prod_\alpha D_{i_1}^\alpha, \ldots, a_n \in \prod_\alpha D_{i_n}^\alpha$, then

$$(a_1 \ldots a_n\omega)(\alpha) = a_1(\alpha)\ldots a_n(\alpha)\omega.$$

1.1.2. *Varieties.* Varieties of algebras are defined by identities. Let us consider the many-sorted case in more detail. Let $X = (X_i,\ i \in \Gamma)$ be a many-sorted set with the countable sets X_i. Starting from the set of symbols of operations Ω we can construct terms over X. Denote the system of such terms by $W = (W_i,\ i \in \Gamma)$. The inductive definition of W is as follows. All the sets X_i are contained in W_i, whose elements are terms of the sort i. If $\omega \in \Omega$, $\tau(\omega) = (i_1, \ldots, i_n; j)$ and $w_1, \ldots, w_n$ are terms of the sorts $i_1, \ldots, i_n$, then $w_1 \ldots w_n\omega$ is the term of the sort j. If in Ω there are symbols of nullary operations ($n = 0$) of the sort i, then they also belong to W_i. The algebra W is, naturally, an Ω-algebra, and it is called the *absolutely free Ω-algebra.*

Identities are formulas of the kind $w \equiv w'$, where w and w' are terms from W of the same sort, say i. Such a formula *is fulfilled in an Ω-algebra* $D = (D_i,\ i \in \Gamma)$ (holds in D), if for any homomorphism $\mu : W \to D$ there holds $w^{\mu_i} = w'^{\mu_i}$.

A set of identities defines a *variety of Ω-algebras, the class of all algebras satisfying all identities of a given set.*

In every variety Θ the set $X = (X_i,\ i \in \Gamma)$ defines a *free algebra*, which is a factor algebra of the absolutely free algebra W.

Birkhoff's theorem takes place also in the many-sorted case.

A class Θ is a variety, if and only if it is closed under Cartesian products, subalgebras and homomorphic images.

Some variety Θ can be taken as an initial variety and one can classify various subvarieties and other axiomatizable classes in Θ. If $X = (X_i,\ i \in \Gamma)$ is a many-sorted set, then a free algebra in Θ, associated with this X, is also denoted by $W = (W_i,\ i \in \Gamma)$. Subvarieties in Θ are defined by the identities of the form $w \equiv w'$, where w and w' are terms of the same sort in this new W. Fully characteristic congruences in W correspond to closed sets of identities.

For an arbitrary algebra $D = (D_i,\ i \in \Gamma)$ from the variety Θ, we can consider the set of homomorphisms $\mathrm{Hom}(W, D)$. We regard homomorphisms $\mu : W \to D$ as rows. Let us explain this idea. First, take a map

$$\mu = (\mu_i,\ i \in \Gamma) : X = (X_i,\ i \in \Gamma) \to D = (D_i,\ i \in \Gamma).$$

Each map of such kind can be viewed as a row

$$\mu = (\mu_1, \ldots, \mu_k) \in D_1^{X_1} \times \cdots \times D_k^{X_k},$$

where $k = |\Gamma|$. In turn, each $\mu_i \in D_i^{X_i}$ is also a row. To a row μ corresponds homomorphism $\mu : W \to D$, which is an element from the set $\mathrm{Hom}(W, D)$. Elements from $\mathrm{Hom}(W, D)$ can be also regarded as rows. Each row μ first defines values of variables in algebra D, and then it calculates values of terms for the given values of variables.

1.1.3. *Group of automorphisms and semigroup of endomorphisms.* For each algebra $D = (D_i,\ i \in \Gamma)$ there is a group $\mathrm{Aut}\, D$ and a semigroup $\mathrm{End}\, D$. In order to construct $\mathrm{Aut}\, D$ we denote by S_i the group of all substitutions (permutations) of the set D_i and let S be the Cartesian product of all S_i, $i \in \Gamma$. $\mathrm{Aut}\, D$ is a subgroup in S, consisting of all $g = (g_i,\ i \in \Gamma) \in S$, which preserve operations of the set Ω, i.e. $(a_1 \ldots a_n \omega)^{g_j} = a_1^{g_{i_1}} \ldots a_n^{g_{i_n}} \omega$.

The semigroup $\mathrm{End}\, D$ is built in a similar way. We pay particular attention to the semigroup $\mathrm{End}\, W$. Its elements are determined by maps

$$s = (s_i,\ i \in \Gamma) : X = (X_i,\ i \in \Gamma) \to W = (W_i,\ i \in \Gamma)$$

which assign terms (Θ-terms) to variables.

1.2. *Algebraic logic*

1.2.1. *First-order Θ-logic.* We again consider the many-sorted case with a set of sorts Γ. Let X be a set of variables, with stratification map $n : X \to \Gamma$. This map is surjective and divides X in sets X_i, $i \in \Gamma$. Each X_i consists of the variables of the sort i, where $n(x) = i$ is the sort of variable x. So, we have the many-sorted set $X = (X_i,\ i \in \Gamma)$.

Let us fix a set of symbols of operations Ω; to each $\omega \in \Omega$ corresponds a Γ-type τ, and we select the variety Θ of Ω-algebras. With this Θ we associate a logic, which is called *Θ-logic*. Let $W = (W_i,\ i \in \Gamma)$ be the free over X algebra in Θ. We fix also a set of symbols of relations Φ; each $\varphi \in \Phi$ has a type $\tau = \tau(\varphi) = (i_1, \ldots, i_n)$. The symbol φ is realized in the algebra $D = (D_i,\ i \in \Gamma)$ as a subset in the Cartesian product $D_{i_1} \times \cdots \times D_{i_n}$.

Now, we can construct the set of formulas of Θ-logic. First, we define elementary formulas. They have the form

$$\varphi(w_1, \ldots, w_n)$$

where $\tau(\varphi) = (i_1, \ldots, i_n)$, $w_1 \in W_{i_1}, \ldots, w_n \in W_{i_n}$. Denote the set of all elementary formulas by ΦW. Let $L = \{\vee, \wedge, \neg, \exists x, x \in X\}$ be a signature of logical symbols and we construct the absolutely free algebra over ΦW of this signature. The corresponding algebra of formulas is denoted by $L\Phi W$.

Among the formulas of the set $L\Phi W$ there is a part, whose elements are logical axioms. These elements are the usual axioms for calculus with functional symbols. Terms in this calculus are Θ-terms. The rules of inference are standard:

(1) *Modus ponens*: u and $u \to v$ imply v,

(2) *Generalization rule*: u implies $\forall x u$.

Here $u, v \in L\Phi W$, $u \to v$ stands for $\neg u \vee v$, and $\forall x u$ means $\neg \exists x \neg u$.

Formulas together with axioms and rules of inference constitute the Θ-logic of the first-order. In particular, we can speak about the logic of group theory, logic of ring theory, etc.

A set of formulas is said to be *closed* if it contains all the axioms and is invariant with respect to the rules of inference.

With each $\varphi \in \Phi$ of type $\tau = (i_1, \ldots, i_n)$ we associate some set of different variables $x_{i_1}^{\varphi}, \ldots, x_{i_n}^{\varphi}$. Then we have a formula $\varphi(x_{i_1}^{\varphi}, \ldots, x_{i_n}^{\varphi})$, which is called a basic one. All variables included into basic formulas are called *attributes*. The set of attributes is denoted by X_0, the set of basic formulas is denoted by ΦX_0. This is a small part of the set of elementary formulas ΦW.

In order to apply logic in DBs, we need the logic to be a logic with equalities.

An *equality* is a formula of the kind $w \equiv w'$, where w and w' have the same sort in W. Such a formula can be regarded either as an equality or as an identity. In a *logic with equalities* the axioms of equality are added to the set of axioms of the logic. Equalities are added to elementary formulas. In logic with equalities the absolutely free algebra of formulas also can be constructed. We denote it by the same abbreviation $L\Phi W$.

1.2.2. *Axiomatic definition of quantifiers.* Quantifiers are logical symbols, but they also can be defined as operations of Boolean algebra:

Given a Boolean algebra H, an existential quantifier $\exists$ is a map $\exists : H \to H$ subject to the following conditions:

(1) $\exists 0 = 0$,

(2) $a \leqslant \exists a$,

(3) $\exists(a \wedge \exists b) = \exists a \wedge \exists b$,

where $a, b \in H$, 0 is the zero element in H.

Every quantifier is a closure operator in H. Different existential quantifiers can be non-permutable. Besides, an existential quantifier is additive: $\exists(a \vee b) = \exists a \vee \exists b$. The set of all $a \in H$, such that $\exists a = a$ is a subalgebra of the Boolean algebra H.

A universal quantifier $\forall : H \to H$ is defined dually:

(1) $\forall 1 = 1$,

(2) $\forall a \leqslant a$,

(3) $\forall(a \vee \forall b) = \forall a \vee \forall b$.

Analogously we have: $\forall(a \wedge b) = \forall a \wedge \forall b$.

There is known correspondence between the existential and universal quantifiers, which allows to pass from one to another.

1.2.3. *Halmos algebras.* Algebraic logic constructs and studies algebraic structures of logic. For example, with the propositional calculus there are associated Boolean algebras, with the intuitionistic propositional calculus Heyting algebras are connected.

There are three approaches to algebraization of the first-order logic, namely, cylindric Tarski algebras [HMT], polyadic Halmos algebras [Hal] and categorical Lawvere's approach [Law1,Law2]. These approaches are based on deep analysis of calculus.[3]

There are also algebraic equivalents of nonclassical first-order logic [Ge]. Similar constructions are developed for other logics [Dis1].

[3] As a rule, will use the word "calculus" for the first-order Θ-logic.

We consider corresponding algebraizations for the case of Θ-logic. This generalization is necessary for DBs with *data type* Θ, and for algebra itself.

Let us proceed from a fixed *scheme, consisting of a set* $X = (X_i,\ i \in \Gamma)$*, a variety* Θ*, and of the algebra* W *free in* Θ *over* X. We also take into account the semigroup End W, whose elements are considered as additional operations.

DEFINITION 1.1. Given a scheme (X, Θ), an algebra H is a Halmos algebra in this scheme, if:

(1.1) H is a Boolean algebra.

(1.2) The semigroup End W acts on H as a semigroup of Boolean endomorphisms.

(1.3) The action of quantifiers of the form $\exists(Y)$, $Y \subset X$ is defined.

These actions are connected by the following conditions:

(2.1) $\exists(\emptyset)$ acts trivially.

(2.2) $\exists(Y_1 \cup Y_2) = \exists(Y_1)\exists(Y_2)$.

(2.3) $s_1\exists(Y) = s_2\exists(Y)$, if $s_1, s_2 \in \mathrm{End}\, W$ and $s_1(x) = s_2(x)$ for $x \in X \setminus Y$.

(2.4) $\exists(Y)s = s\exists(s^{-1}Y)$ for $s \in \mathrm{End}\, W$, if the following conditions are fulfilled:

(1) $s(x_1) = s(x_2) \in Y$ implies $x_1 = x_2$,

(2) If $x \notin s^{-1}Y$, then $\Delta s(x) \cap Y = \emptyset$.

Here, $s^{-1}Y = \{x \mid s(x) \in Y\}$ and $\Delta s(x)$ is a *support* of the element $w = s(x) \in W$, i.e. the set of all $x \in X$, *which is involved in the expression of the element* w.

All Halmos algebras in the given scheme form a variety, denoted by HA_Θ. We will deal with such Halmos algebras and sometimes call them *HAs specialized in* Θ. In the next three subsections we will introduce examples of Halmos algebras.

1.2.3.1. Given $D \in \Theta$, consider $\mathrm{Hom}(W, D)$. Denote by M_D the set of all subsets in $\mathrm{Hom}(W, D)$, i.e. $M_D = \mathrm{Sub}(\mathrm{Hom}(W, D))$. The set M_D is a Boolean algebra. If $A \in M_D$, $\mu \in \mathrm{Hom}(W, D)$, $s \in \mathrm{End}\, W$, then define μs by the rule $\mu s(x) = \mu(s(x))$. An action of the semigroup End W in M_D is defined by

$$\mu \in sA \Leftrightarrow \mu s \in A.$$

For $Y \subset X$ we define $\mu \in \exists(Y)A$, if there is $\nu : W \to D$, $\nu \in A$, such that $\mu(x) = \nu(x)$ for every $x \in X \setminus Y$. This defines the action of quantifiers in M_D. It can be checked that all the axioms of Halmos algebra are fulfilled in M_D and it is the first example of HA_Θ.

For h from a Halmos algebra H, denote its *support* by Δh.

$$\Delta h = \{x \in X,\ \exists x h \neq h\}.$$

If Δh is finite, then the element h is called an element of *finite support*. In the previous subsection we used supports of elements of the algebra W, while here we regard supports of elements of Halmos algebra. In every Halmos algebra H all its elements with finite support constitute a subalgebra. This subalgebra is called a *locally finite part of* H. *A Halmos algebra* H *is locally finite, if all its elements have finite support.*

1.2.3.2. Denote by V_D the locally finite part of the Halmos algebra M_D. Here $A \in V_D$ if for some finite subset $Y \subset X$ and elements $\mu, \nu \in \mathrm{Hom}(W, D)$, the equality $\mu(x) = \nu(x)$ for every $x \in Y$ implies that

$$\mu \in A \Leftrightarrow \nu \in A.$$

In other words, belonging of a row to the set A is checked on a finite part of X.

1.2.3.3. Now let us consider the main example: the HA_Θ-algebra of the first order calculus. Suppose that all X_i from the set $X = (X_i,\ i \in \Gamma)$ are infinite. Besides that, the symbols of relations Φ are added to the scheme and we have the algebra of formulas $L\Phi W$.

Variables occur in the formulas. These occurrences can be free or bound. Let us define the action of elements $s \in \mathrm{End}\, W$ on the set $L\Phi W$. If u is a formula and $x_1, \dots, x_n$ occur free in it, then we substitute them by $sx_1, \dots, sx_n$, and we do it for all free occurrences of variables. So we get su. For example, it follows from this definition that if $u = \varphi(w_1, \dots, w_n)$ is an elementary formula, then $su = \varphi(sw_1, \dots, sw_n)$. However, the definition doesn't give a representation of the semigroup End W on the set $L\Phi W$. Simple examples show that the condition $(s_1 s_2)u = s_1(s_2 u)$ need not be fulfilled. Let us define equivalence ρ on the set $L\Phi W$ by the rule: $u\rho v$ if u and v differ only by the names of bound variables. Take the quotient set $L\Phi W/\rho = \overline{L\Phi W}$, and call this transition *quotienting by renaming of bound variables.* Elements of $\overline{L\Phi W}$ we also call formulas, but these formulas are regarded up to renaming of bound variables. It is easy to see that the equivalence ρ is compatible with the signature, but it is not compatible with the action of elements from End W. Therefore all operations of the set $L = \{\vee, \wedge, \neg, \exists x,\ x \in X\}$ are defined on $\overline{L\Phi X}$, but the action of elements from End W has to be defined separately. It is done as follows.

Let u be a formula, $x_1, \dots, x_n$ are all its free variables, and take $sx_1, \dots, sx_n$, $s \in \mathrm{End}\, W$. We say that u is *open* for s if there are no bound variables which occur in the sets $\Delta sx_1, \dots, \Delta sx_n$. For each u we denote by $\bar{u}$ the corresponding class of equivalent elements. For $s \in \mathrm{End}\, W$ we always can find some formula u' in the class $\bar{u}$ which is open for s. Then we set $s\bar{u} = \overline{su'}$. It is easy to see that if we have another formula u'' in $\bar{u}$ which is opened for s, then $su'\rho su''$ and $\overline{su'} = \overline{su''}$. Hence the definition of $s\bar{u}$ is correct. This rule gives the representation of the semigroup End W as a semigroup of transformations of the set of formulas $\overline{L\Phi W}$.

From now on we work with the set of formulas $\overline{L\Phi W}$. Axioms and rules of inference are related to this set too. They are standard, as before.

Now we pass to the *Lindenbaum–Tarski algebra.* We have an equivalence τ which is defined as follows: $\bar{u}\tau\bar{v}$ if the formula $(\bar{u} \to \bar{v}) \wedge (\bar{v} \to \bar{u})$ is derivable. It can be verified that τ is congruence on $\overline{L\Phi W}$, and this τ is also compatible with the action of the semigroup End W.

Denote by U the result of the quotienting of $\overline{L\Phi W}$ by τ. Define an equivalence η on $L\Phi W$ by the rule: $u\eta v \Leftrightarrow \bar{u}\tau\bar{v}$. Since $\rho \subset \eta$, the set $U = \overline{L\Phi W}/\tau$ can be identified with $L\Phi W/\eta$. It is important to emphasize that the equivalence η can also be defined by means of the Lindenbaum–Tarski scheme and U is the *Lindenbaum–Tarski algebra.* It can be proved that:

(1) *U is a Boolean algebra for the operations $\vee, \wedge, \neg$;*
(2) *The semigroup End W acts on U as a semigroup of endomorphisms of this algebra;*
(3) *All $\exists x$ for different x commute. This allows to define in U quantifiers $\exists(Y)$ over all $Y \subset X$.*

All the above gives the following result:

THEOREM 1.2. *The algebra U with the given operations is an algebra in HA_Θ.*

This is the syntactical approach to the definition of the Halmos algebra of first-order Θ-logic. This approach is due to Z. Diskin [Dis1]. There is also semantical approach, which is described in [BPl3]. Both approaches give one and the same result. Finally, we can use the verbal congruence of the variety HA_Θ, and obtain once more the same algebra U.

1.2.4. *Homomorphisms of Halmos algebras.* We start from a very important feature of the algebra U, and note first of all, that this algebra is locally finite. Take the basic set ΦX_0 in the algebra of formulas $L\Phi W$, and let U_0 be a corresponding basic set in U. The set U_0 generates the algebra U.

THEOREM 1.3. *Let H be an arbitrary Halmos algebra, and let $\zeta : \Phi X_0 \to H$ be a map, such that for every $u = \varphi(x_1, \dots, x_n) \in \Phi X_0$ we have $\Delta\zeta(u) \subset \{x_1, \dots, x_n\}$. Such ζ gives a map $\zeta : U_0 \to H$ and the last map uniquely extends up to a homomorphism $\zeta : U \to H$.*

Given a model (D, Φ, f), $D \in \Theta$, take the special case, when $H = V_D$. Define $\zeta = \hat{f} : \Phi X_0 \to V_D$ by the rule: if $u = \varphi(x_1, \dots, x_n) \in \Phi X_0$, then:

$$\hat{f}(u) = \{\mu,\ \mu \in \mathrm{Hom}(W, D),\ (\mu(x_1), \dots, \mu(x_n)) \in f(\varphi)\}.$$

Then $\Delta\hat{f}(u) = \{x_1, \dots, x_n\}$ and we have a homomorphism

$$\hat{f} : U \to V_D.$$

From this definition follows, that if $u = \varphi(w_1, \dots, w_n)$ is an elementary formula, then

$$\hat{f}(u) = \{\mu,\ \mu \in \mathrm{Hom}(W, D),\ (\mu(w_1), \dots, \mu(w_n)) \in f(\varphi)\}.$$

So, for every model (D, Φ, f), $D \in \Theta$, we have a canonical homomorphism $\hat{f} : U \to V_D$, and it can be proved that every homomorphism $U \to V_D$ is, in fact, some $\hat{f}$. Now we can say that $\hat{f}(u)$ for every $u \in U$ is the value of u in the model D. If $f(u) = 1$, it means that u holds in D.

Now some words about kernels of homomorphisms in the variety HA_Θ. If $\sigma : H \to H'$ is a homomorphism in HA_Θ, then we have two kernels, viz. the coimage of zero and the coimage of the unit. The coimage of zero is an ideal, the coimage of the unit is a filter. A subset T of H is a *filter* if

(1) *$a \wedge b \in T$ if a and b belongs to T.*
(2) *From $a \in T$ and $b \in H$ follows $a \vee b \in T$.*

(3) $\forall(Y)a \in T$ *if* $a \in T$, $Y \subset X$.

The definition of ideal is dual. For every filter T we can consider a quotient algebra H/T, which is simultaneously H/F, where F is the ideal defined by the rule: $h \in T$ if and only if $\bar{h} \in F$.

If T is a subset of H, then the filter in H, generated by T, consists of elements of the form

$$\forall(X)a_1 \wedge \cdots \wedge \forall(X)a_n \vee b, \quad a_i \in T,\ b \in H.$$

Every filter is closed under existential quantifiers, while every ideal is closed under universal quantifiers. Besides, it can be proved that every ideal and every filter is closed under the action of the semigroup End W.

THEOREM 1.4 (see, e.g., [BPl3]). *A set* $T \subset H$ *is a filter in* H, *if and only if it satisfies the conditions*:

(1) $1 \in T$

(2) *If* $a \in T$, $a \to b \in T$, *then* $b \in T$.

(3) *If* $a \in T$, *then* $\forall(Y)a \in T$, $Y \subset X$.

Here, $a \to b = \neg a \vee b$.

We will consider two rules of inference in HAs:

(1) *From* a *and* $a \to b$ *follows* b.

(2) *From* a *follows* $\forall(Y)a$, $Y \subset X$.

If H is locally finite, then the second rule can be replaced by

(2′) *From* a *follows* $\forall x a$, $x \in X$.

For every set T, one can consider the set of elements which are derived from T.

THEOREM 1.5. *If* T *is a set in* H, *which contains a unit, then the filter, generated by* T, *is a set of all* $h \in H$, *which are derived from* T.

The notion of derivability in HAs agrees with the notion of derivability of formulas in logic. The notion of filter in HAs corresponds to the notion of *closed* set of formulas.

Given a model (D, Φ, f), we have $\hat{f}: U \to V_D$. The corresponding filter Ker $\hat{f}$ we can consider as the *elementary theory* (Θ*-theory*) *of the given model.*

THEOREM 1.6. *Let* T *be a subset in* U, *let* $T' = K$ *be the axiomatizable class of models defined by the set* T (*not empty*), *and let* $T'' = K'$ *be all axioms of* K *in* U (*a closure of* T). *Then* T'' *is the filter in* U *generated by* T.

1.2.5. *Equalities in Halmos algebras. Cylindric algebras.* An equality in Halmos algebra H is not a formula, but an element in H of the form $w \equiv w'$, where w, $w' \in W$ and have the same sort.

DEFINITION 1.7. $H \in HA_\Theta$ is an algebra with equalities, when all $w \equiv w'$ are defined as elements in H and the following axioms hold:

(1) $s(w \equiv w') = (sw \equiv sw')$, $s \in \operatorname{End} W$,
(2) $(w \equiv w) = 1$, $w \in W$,
(3) $(w_1 \equiv w_1') \wedge \cdots \wedge (w_n \equiv w_u') < w_1 \dots w_n \omega \equiv w_1' \dots w_w' \omega$ if $\omega \in \Omega$ is of the appropriate type,
(4) $s_w^x a \wedge (w \equiv w') < s_{w'}^x a$, $a \in H$, s_w^x transforms x into w of the same sort and leaves $y \neq x$ fixed.

In the algebra V_D the equality is defined by the rule: $w \equiv w'$ is the set of all $\mu : W \to D$, for which $w^\mu = w'^\mu$ is satisfied in D. In the algebra U with equalities, the symbol $\equiv$ is added to the set Φ, and the standard axioms of equality are added to the initial axioms of Θ-logic. In this case the set Φ can be empty. Such an algebra U arises from Θ-logic with equalities.

Each equality $w \equiv w'$ in H is considered as a new nullary operation, therefore a subalgebra of an algebra with equalities should contain all the elements $w \equiv w'$. The same remark is valid for homomorphisms of algebras with equalities. If H is an algebra with equalities, then so is H/T, where T is a filter.

DEFINITION 1.8. A Boolean algebra H is a cylindric algebra in the scheme (X, Θ) if:
(1) To every $x \in X$ an existential quantifier $\exists x$ of the Boolean algebra H is assigned, and all these quantifiers commute.
(2) The nullary operations $d_{w,w'}$, where $w, w' \in W$, which are called diagonals, lie in H.
(3) The following conditions hold:
   (a) $d_{w,w} = 1$.
   (b) If $x \notin \Delta w$, $x \notin \Delta w'$, then $\exists x(d_{x,w} \wedge d_{x,w'}) = d_{w,w'}$.
   (c) If $x \notin \Delta w$ and h is arbitrary element in H, then $\exists x(h \wedge d_{x,w}) \wedge \exists x(\neg h \wedge d_{x,w}) = 0$.

There is also a condition (d) in which transformations of variables are hidden, (see [Ci2] for the details).

The Halmos algebras with equalities U, M_D and V_D can also be viewed as cylindric algebras.

The following theorem holds

THEOREM 1.9. *Given $X = (X_i,\ i \in \Gamma)$, all X_i being infinite, the category of locally finite HA_Θ with equalities and the category of locally finite cylindric algebras with the same X and Θ are equivalent categories.*

This theorem for pure first-order logic was proved in [Ga]. The proof for the Θ-logic requires insignificant modifications.

Along with HAs, cylindrical algebras can be used in databases. However, an important role is played by the semigroup End W in the signature, and therefore we place the emphasis on Halmos algebras. See [Ci2,Fe].

1.3. *The categorical approach to algebraic logic*

1.3.1. *Algebraic theories.* In this subsection we consider another approach to the algebraization of first-order logic, specialized in some variety Θ. We will speak about relational algebras, which are tied with the idea of doctrines in categorical logic of Lawvere [Law1, Law2,Law3], see also [Ben1] and [BPl1].

We start from the notion of *algebraic theory*. It is a special kind of category T. The objects of this category are of the type $\tau = (i_1, \dots, i_n)$, $i_s \in \Gamma$, where Γ, as before, is a set of sorts. The arity n can be also zero. It is assumed that the set of morphisms contains *projectors* $\delta_\tau^s \colon \tau \to i_s$, which allow to represent every object $\tau = (i_1, \dots, i_n)$ as a product $\tau = i_1 \times \cdots \times i_n$ in the category T.

An algebra D in the theory T is a covariant functor $D : T \to$ Set, *which is compatible with products in T.*

For each $i \in \Gamma$, taken as an object, we have $D(i) = D_i$, and so we can take $D = (D_i,\ i \in \Gamma)$. If further $\omega : \tau \to j$ is a morphism in T, then we have:

$$D(\omega) : D(\tau) = D_{i_1} \times \cdots \times D_{i_n} \to D_j.$$

This is an operation in D of the type $(i_1, \dots, i_n, j)$. All algebras in the theory T form a variety $\Theta = \operatorname{Alg} T$, which depends on T. On the other hand, for every variety Θ there can be constructed a theory T, leading to the given Θ.

1.3.2. *Relational algebras.* We denote by Bool the category of Boolean algebras, and let Sub : Set $\to$ Bool be the functor, that assigns to every set M the Boolean algebra of all subsets of M. This is a contravariant functor. Consider the composition of functors $D : T \to$ Set and Sub : Set $\to$ Bool. Then we have the contravariant functor

$$(\mathrm{Sub} \circ D) : T \to \mathrm{Bool}.$$

For every $\tau = (i_1, \dots, i_n)$ the algebra $(\mathrm{Sub} \circ D)(\tau)$ is the Boolean algebra of all subsets in $D_i \times \cdots \times D_{i_n}$. The functor $(\mathrm{Sub} \circ D)$ is an example of a *relational algebra*, connected to the algebra D, and we denote this algebra by $V_D = (\mathrm{Sub} \circ D)$. It plays the same role as V_D in Halmos algebras.

Let us turn now to the general definition. *A relational algebra R in the theory T is defined as a contravariant functor*

$$R : T \to \mathrm{Bool},$$

subject to some additional conditions.

It is required that for every $\alpha : \tau \to \tau'$ in T along with the homomorphism of Boolean algebras $\alpha_* = R(\alpha) : R(\tau') \to R(\tau)$, there exists a conjugate map, which we denote by $\alpha^* = R'(\alpha) : R(\tau) \to R(\tau')$. The map α^* is not a homomorphism of Boolean algebras, but it preserves zero element and addition. Taking $R'(\tau) = R(\tau)$, we consider R' as a covariant functor. If now we define

$$\exists(\alpha) = \alpha_* \alpha^* : R(\tau) \to R(\tau),$$

then $\exists(\alpha)$ is an *existential quantifier* of the Boolean algebra $R(\tau)$. We define now some new operation. Let τ_1 and τ_2 be two objects in T and let $\tau_1 \times \tau_2$ be the product in T. For every $a \in R(\tau_1)$ and $b \in R(\tau_2)$ we define an element $a \times b \in R(\tau_1 \times \tau_2)$. We do it in the following way: let $\pi_1 : \tau_1 \times \tau_2 \to \tau_1$ and $\pi_2 : \tau_1 \times \tau_2 \to \tau_2$ be the projectors, corresponding $\tau_1 \times \tau_2$. Then we have:

$$(\pi_1)_* : R(\tau_1) \to R(\tau_1 \times \tau_2),$$

$$(\pi_2)_* : R(\tau_2) \to R(\tau_1 \times \tau_2).$$

Define

$$a \times b = (\pi_1)_*(a) \wedge (\pi_2)_*(b).$$

It can be proved, that, given $\alpha : \tau_1 \to \tau_1'$ and $\beta : \tau_2 \to \tau_2'$, then, for every $a \in R(\tau_1')$ and $b \in R(\tau_2')$

$$(\alpha \times \beta)_*(a \times b) = \alpha_*(a) \times \beta_*(b).$$

This is a property of the functor R. Such property for the functor R' need not always hold, and we *include it into the definition of relational algebra* as an *additional axiom*. This axiom is as follows: for all $a \in R(\tau_1)$ and $b \in R(\tau_2)$ holds

$$(\alpha \times \beta)^*(a \times b) = \alpha^*(a) \times \beta^*(b). \qquad (*)$$

All the above gives the definition of a relational algebra. It can be verified that $V_D = (\mathrm{Sub} \circ D)$ is an example of a relational algebra.

Summarizing, we can say that a relational algebra is a contravariant functor $R : T \to$ Bool *such that there exists an appropriate covariant functor* $R' : T \to$ Bool, *satisfying the axiom* $(*)$.

We can also speak about relational algebras of first-order calculus, i.e. algebras of the type U. The construction of such U can be accomplished with the help of the Theorem 1.10 of this section. We sketch here the approach to a proof of this result. As we know, for every variety Θ there exists an algebraic theory T, such that $\Theta = \mathrm{Alg}(T)$. For the construction of T we must take $X = (X_i,\ i \in \Gamma)$, where all X_i are countable. Let W be the free algebra in Θ over the given X. On the basis of this W we construct the theory T. The algebra W can be considered as a functor $W : T \to$ Set. A special approach allows to take another functor $T \to \Theta$ (contravariant one) for the same W, which we also denote by W. In this case $W(\tau)$ is the subalgebra in W, generated by a set of variables connected with the object τ. The additional condition requires that for every $i \in \Gamma$ in the set of operations Ω there exist nullary operations of the type i. This new W allows to consider an antiisomorphism of the semigroups $\mathrm{End}\,\tau$ and $\mathrm{End}\,W(\tau)$. On the other hand, if R is a relational algebra, then for every $\alpha : \tau \to \tau$ we have $\alpha_* : R(\tau) \to R(\tau)$. This gives a representation of the semigroup $\mathrm{End}\,\tau$ as a semigroup of endomorphisms of Boolean algebra $R(\tau)$ with contravariant action. Hence the semigroup $\mathrm{End}\,W(\tau)$ acts in $R(\tau)$ in a usual way as a semigroup of endomorphisms.

The proof of the following result is based on the remarks above.

THEOREM 1.10 [Vo2,Ro]. *Given $X = (X_i,\ i \in \Gamma)$ and Θ, all X_i being infinite, the category of locally finite Halmos algebras with equalities and the category of relational algebras are equivalent.*

In particular, from the Halmos algebra U we can get a relational algebra of the type U. We see that *all approaches to algebraization of first-order logic are equivalent.*

1.3.3. *Changing the variety Θ.* We want to know what happens with HA_Θ when the variety Θ is changed. In the case when Θ' is a subvariety of Θ, every algebra in $\mathrm{HA}_{\Theta'}$ can be considered as an algebra in HA_Θ. This gives an injection of the variety $\mathrm{HA}_{\Theta'}$ into HA_Θ as a subvariety. The Θ-identities of Θ' can be transformed into identities of $\mathrm{HA}_{\Theta'}$ in HA_Θ. On the other hand, if U is the algebra of calculus in HA_Θ and U' is the algebra of calculus in $\mathrm{HA}_{\Theta'}$, then we have a canonical epimorphism $U \to U'$, and the kernel of this epimorphism can be well described if we consider algebras with equalities [BPl3].

Let now T and T' be two algebraic theories. A functor $\psi : T \to T'$ is called a *morphism of theories* if $\psi(\tau) = \tau'$ *for every object* τ *and there is compatibility with projectors.* If $D : T' \to \mathrm{Set}$ is an algebra in $\mathrm{Alg}\,T'$, then $D\psi : T \to \mathrm{Set}$ is an algebra in $\mathrm{Alg}\,T$. So we have a functor

$$\mathrm{Alg}\,T' \to \mathrm{Alg}\,T.$$

Also, if V_D is $(\mathrm{Sub} \circ D) : T' \to \mathrm{Bool}$, then $V_{D\psi} = V_D\psi$ is

$$(\mathrm{Sub} \circ D)\psi : T \to \mathrm{Bool}.$$

Let now $R : T' \to \mathrm{Bool}$ be an arbitrary relational algebra. It can be proved that $R\psi : T \to \mathrm{Bool}$ is also relational algebra. So, a morphism $\psi : T \to T'$ gives a functor

$$\psi^* : \mathit{Rel}(T') \to \mathit{Rel}\,T,$$

where $\mathit{Rel}(T)$ is the category of relational algebras in T.

Let now Θ be a variety and Θ' be its subvariety. We can construct theories T and T' with a morphism $\psi : T \to T'$. This morphism gives an injection $\psi^* : \mathit{Rel}(T') \to \mathit{Rel}(T)$, which is compatible with the injection $\mathrm{HA}_{\Theta'} \to \mathrm{HA}_\Theta$, considered before.

1.3.4. *Generalizations.* There are two directions: the generalization of the notion of algebra on the one hand, and of logic on the other hand. First, we explore the notion of algebra. One of the approaches to its generalization, which uses the algebraic theory T, is based on the idea to take some arbitrary *topos* Ξ instead of the category Set in the definition of algebra. A topos is a category, which is in some sense close to the category of sets [Jo,Gol, MacL,BPl3]. We can, for example, consider the topos Ξ of *fuzzy sets*, where fuzzy set is defined as a fuzzy quotient set of a usual set.

If Θ is a variety of algebras, then for Θ we take the theory T and consider functors $D:T \to \Xi$. So we have *fuzzy Θ-algebras*. They may be fuzzy groups, fuzzy rings and so on.

Now we generalize the notion of relational algebra. For every object of the topos the system of its subobjects form a Heyting algebra which is a generalization of Boolean algebra, connected with intuitionistic logic [Gol]. This gives a contravariant functor $\mathrm{Sub}:\Xi \to \mathrm{Heyt}$, where Heyt is the category of Heyting algebras. Having at the same time an "algebra" $D:T \to \Xi$ we can construct the composition $(\mathrm{Sub} \circ D):T \to \mathrm{Heyt}$. It is an example of relational algebra, associated with intuitionistic logic. *A relational algebra in intuitionistic logic is a functor* $R:T \to \mathrm{Heyt}$, *satisfying some special conditions.*

The presence of the theory T here means that we consider relational algebras, specialized in the variety $\mathrm{Alg}\,T$.

Some yeas ago, see [BPl3], a problem was stated concerning generalization of Theorem 1.9 for arbitrary, not only classical, logics. This problem is solved by Z. Diskin [Dis1] in sufficiently general situation, including also modalities. Instead of categories Bool and Heyt, he takes an appropriate category L and defines relational algebras of the type $R:T \to L$. This definition uses doctrines [KR] in categorical logic. Halmos algebras, associated with L and $\mathrm{Alg}\,T = \Theta$, are also defined.

In these terms the solution of the problem is given. Besides, there is constructed some logical calculus, connected with L. For this calculus a Lindenbaum–Tarski algebra, which is a generalization of a Halmos algebra with given L, is built. See also [ANS,BP,Ne,Ni, WBT,Man].

2. Algebraic model of databases

2.1. *Passive databases*

2.1.1. *Preliminaries.* We distinguish passive and active DBs. To be more precise we speak about corresponding models of DBs. Passive DBs have poor algebraic structure, while the structure of active ones is sufficiently rich. Passive DBs are simpler and not so sophisticated; active DBs are much more complicated, but they are also much more interesting. There are natural transitions from passive DBs to active ones.

We proceed from a definite *database scheme*. It includes *permanent* and *varying* parts. *Permanent part* consists of a variety Θ, which serves as *data type*, and a system of variables X. In fact, this is the scheme of a HA. *The varying part* consists of the set of symbols of relations Φ.

We recall briefly the main notions, described in the first section. All algebras from Θ are Ω-algebras where Ω is a set of symbols of operations. All algebras are generally speaking many-sorted as well as the set of variables X. The set of sorts Γ is simultaneously a set of names of domains. Each algebra $D \in \Theta$ has the form $D = (D_i,\ i \in \Gamma)$, where D_i are domains of D. There is a map $n:X \to \Gamma$ and for each $i \in \Gamma$ the set X_i is the set of all variables of the sort i. Then $X = (X_i,\ i \in \Gamma)$. Let $W = (W_i,\ i \in \Gamma)$ be the free Θ-algebra, corresponding to X. This is the algebra of Θ-terms. For each operation $\omega \in \Omega$ its type has the form $\tau = \tau(\omega) = (i_1, \dots, i_n, j)$. The type of a relation $\varphi \in \Phi$ has the form $\tau = \tau(\varphi) = (i_1, \dots, i_n)$, all $i, j \in \Gamma$.

If $D = (D_i,\ i \in \Gamma)$ is an algebra in Θ, then the set Φ can be realized in D. Let f be a function embodying this realization. If $\varphi \in \Phi$ and $\tau(\varphi) = (i_1, \dots, i_n)$, then $f(\varphi)$ is a subset in the Cartesian product $D_{i_1} \times \cdots \times D_{i_n}$. This subset can be represented as a table. To a tuple $(i_1, \dots, i_n)$ we assign a set of different variables – *attributes* $(x_{i_1}, \dots, x_{i_n})$. Here i_α and i_β may coincide, but the corresponding x_{i_α} and x_{i_β} are different. The table $f(\varphi)$ consists of rows (functions) r, such that $r(x_{i_\alpha}) \in D_{i_\alpha}$. If $\omega \in \Omega$ and $\tau(\omega) = (i_1, \dots, i_n, j)$, then the operation ω is realized as a map $\omega : D_{i_1} \times \cdots \times D_{i_n} \to D_j$.

All this was already mentioned in the previous section. Let us note that in the case of DBs a variety Θ is defined by identities, used for verification of correctness of operations execution in algebras from the data type Θ.

Recall that the set of symbols of operations Φ defines a Halmos algebra U in the given scheme. The algebra $L\Phi W$ is built over the set of formulas of first-order Θ-logic. Elements of U are formulas as well, but they are considered up to a special equivalence (see 1.2.3.3). An algebra $D \in \Theta$ in the same scheme also determines a HA, denoted by V_D. It is defined on the set of subsets of the set $\mathrm{Hom}(W, D)$. In turn $\mathrm{Hom}(W, D)$ can be regarded as $\prod_{i \in \Gamma} D_i^{X_i}$ or as $\prod_{x \in X} D_x$, $D_x = D_{n(x)}$.

Thus, elements from V_D can be treated as relations. In the previous section we defined the notion of support of a subset of $\mathrm{Hom}(W, D)$. All relations under consideration have finite supports.

We are dealing with relational databases. Queries in a relational DB, connected with Θ-logic, are written by formulas from the language $L\Phi W$. One and the same query may correspond to different formulas. Thus, a query should be viewed as a class of equivalent formulas. The equivalence arising here, is the same that appeared in the definition of the Halmos algebra U. Now it is clear, that *the algebra U may be considered as an algebra of queries to a DB. It is the universal algebra of queries; in a sense it does not depend on data algebra D. Given $D \in \Theta$, the algebra V_D is the universal algebra of replies to queries.*

Let F_D be the set of all realizations f of the set Φ in D. F_D is easily equipped with the structure of a Boolean algebra. We have a triple (F_D, U, V_D).

A canonical homomorphism $\hat{f} : U \to V_D$ corresponds to every $f \in F_D$.

It is easy to verify that the transition $\hat{}: F_D \to \mathrm{Hom}(U, V_D)$ is a bijection of sets.

Let us define an operation $* : F_D \times U \to V_D$, setting $f * u = \hat{f}(u)$. It leads to a $*$-automaton (see 2.2.1)

$$\mathrm{Atm}\, D = (F_D, U, V_D),$$

which is simultaneously treated as an *universal active database for a given data algebra D.*

Elements from F_D are called *states* of the database $\mathrm{Atm}\, D$, and $f * u$ is a *reply to a query u in a state f*. The transition from a query to a reply is realized by a homomorphism of HAs.

We remark here, that the structure of $\mathrm{Atm}\, D$ depends on the algebra D and on the fixed scheme. In DBs this structure is used as the universal active DB. All the other active DBs are defined on the based of this universal one. Besides that, this structure has applications in some problems of algebra and model theory [BPl5].

If u is considered as a formula in the language $L\Phi W$, then we also write $f * u$, assuming $f * u = f * \bar{u}$ ($\bar{u}$ is the corresponding element in the algebra U). An element $f * u$ can be defined independently from the transition to the Halmos algebra U.

A given a homomorphism of algebras $\delta : D' \to D$ defines a map

$$\tilde{\delta} : \mathrm{Hom}(W, D') \to \mathrm{Hom}(W, D)$$

by the rule: if $\mu \in \mathrm{Hom}(W, D')$, then $\tilde{\delta}(u) = \mu\delta$. The map $\tilde{\delta}$ defines, in turn, a homomorphism of Boolean algebras

$$\delta_* : V_D \to V_{D'}$$

by the rule: $\mu \in \delta_*(A)$ if and only if $\mu\delta \in A$, $A \in V_D$.

It is proved that if $\delta : D' \to D$ is an epimorphism, then $\delta_* : V_D \to V_{D'}$ is a monomorphism of HAs.

Let $\delta : D' \to D$ be an epimorphism. For every $f \in F_D$ we consider $\hat{f} : U \to V_D$ and the composition of this homomorphism with $\delta_* : V_D \to V_{D'}$. We have $\hat{f}\delta_* : U \to V_{D'}$. An element $f^{\delta^*} \in F_{D'}$ defined by the condition $\widehat{f^{\delta^*}} = \hat{f}\delta_*$ corresponds to this homomorphism.

The map $\delta^* : F_D \to F_{D'}$ is a monomorphism of Boolean algebras.

The pair of maps (δ^*, δ_*) determines a transition $\mathrm{Atm}\, D \to \mathrm{Atm}\, D'$, and the following condition is satisfied:

$$(f * u)^{\delta_*} = f^{\delta^*} * u.$$

We have here an embedding of universal DBs, which does not change the algebra of queries U.

Let us consider now a situation when the algebra U is also changed. Given homomorphisms $\zeta : U' \to U$, $\hat{f} : U \to V_D$ and $\delta_* : V_D \to V_{D'}$, we can compose

$$\zeta \hat{f} \delta_* : U' \to V_{D'}.$$

Once more we denote by f^{δ^*} the element of $F_{D'}$ which corresponds to this homomorphism. We preserve notation in this more general case. A triple $(\delta^*, \zeta, \delta_*)$ is associated with the transition $\mathrm{Atm}\, D \to \mathrm{Atm}\, D'$ according to the condition

$$f^{\delta^*} * u = \left(f * u^{\zeta}\right)^{\delta_*}, \quad u \in U', \ f \in F_D.$$

2.1.2. *Passive databases.* A database is a system which can accept, store and process information, and which allows as well as the querying of its contents. A DB contains initial information, which is used when constructing a reply to a query to the DB. This information can be formalized as a *passive* DB. An *active* DB reflects the processing of a query in order to get a reply to this query. It is possible to query either directly the data stored in the DB, or some information which can be somehow derived from the basic one represented by the source data.

Now we concentrate our attention on passive DBs. A passive DB assumes the presence of a data algebra. Operations of this algebra are used in data processing. It is also supposed, that in a passive DB there is a set of symbols (names) of relations and a set of possible states of this DB. States realize symbols of relations as relations in the data algebra. These states can be selected according to some conditions.

DEFINITION 2.1. A passive database is a triple (D, Φ, F) where $D \in \Theta$ is a data algebra, Φ is a set of symbols of relations, F is a set of feasible states of a database.

Recall that elements of F are realizations f of the set Φ in D. For each $f \in F$ there is a model (D, Φ, f). A passive DB is the set of such models, D and Φ being fixed. It is clear that one can consider the case of empty set of operations Ω. In this case D is not an algebra, but a many-sorted set. Recall that, if (D, Φ, f) and (D', Φ, f') are two models with one and the same Φ, then a *homomorphism of models* $(D', \Phi, f') \to (D, \Phi, f)$ is a homomorphism $\delta : D' \to D$ of algebras from Θ, which is compatible with all relations of the set Φ under the map $f' \to f$.

This definition of a homomorphism of models is easily translated into the language of HAs. As we know, the algebra of formulas $L\Phi W$ contains the set of elementary formulas ΦW. Among the elementary formulas there is a set of basic formulas ΦX_0, where X_0 is the set of attributes. Basic formulas have the form

$$\varphi\left(x_1^{\varphi}, \ldots, x_n^{\varphi}\right), \quad \varphi \in \Phi, \ x_i^{\varphi} \in X_0,$$

where all attributes $x_1^{\varphi}, \ldots, x_n^{\varphi}$ are distinct and compatible with the type of the formula φ. Regarding basic formulas as elements of the algebra U we can state that these elements generate U.

PROPOSITION 2.2. *A homomorphism $\delta : D' \to D$ together with a transition $f' \to f$ defines a homomorphism of models if and only if for every basic element $u \in U'$ there is an inclusion*

$$f' * u \subset (f * u)^{\delta_*}.$$

Let us consider now homomorphisms of models with varying Φ.

We have (D, Φ, f) and (D', Φ', f') and, correspondingly, Halmos algebras U and U'. Take a homomorphism of algebras $\delta : D' \to D$ and a homomorphism of Halmos algebras $\zeta : U' \to U$.

DEFINITION 2.3. A pair (δ, ζ) defines a homomorphism of models $(D', \Phi', f') \to (D, \Phi, f)$ if for every basic element $u \in U'$ there is an inclusion

$$f' * u \subset \left(f * u^{\zeta}\right)^{\delta_*}.$$

It makes more sense to speak of epimorphisms δ in this definition. However, we need this restriction not always.

DEFINITION 2.4. Two models (D, Φ, f) and (D', Φ', f') are called similar under the map $f' \to f$, if there is an isomorphism of algebras $\delta : D' \to D$ and homomorphisms $\zeta : U' \to U$, $\zeta' : U \to U'$, such that for any basic elements $u \in U'$, $u' \in U$ hold

$$(f' * u) \subset \left(f * u^{\zeta}\right)^{\delta_*}; \qquad f * u' \subset \left(f' * u'^{\zeta'}\right)^{\delta_*^{-1}}.$$

The notion of similarity of models generalizes the notion of isomorphism of models. We will use it in the sequel.

In passive databases, D, Φ and F can be changed. It leads to a category of DBs.

2.1.3. *Category of passive databases.* Let us define a *category whose objects are passive databases*. If (D, Φ, F) and (D', Φ', F') are two passive DBs, then a *morphism* of the first one into the second is a map $\alpha : F \to F'$, an epimorphism $\delta : D' \to D$ and a homomorphism $\zeta : U' \to U$, such that for every $f \in F$ and $u \in U'$

$$f^{\alpha} * u = \left(f * u^{\zeta}\right)^{\delta_*}.$$

It is evident that a category really arises here, and besides that the pair (δ, ζ) under the map $f^{\alpha} \to f$ always defines a homomorphism of models $(D', \Phi', f^{\alpha}) \to (D, \Phi, f)$.

We translate now the notion of morphism of passive databases to a language which is free from HAs.

We proceed from the algebras $L\Phi'W$ and $L\Phi W$ of formulas, and for the first one we take a system of basic formulas $\Phi' X_0'$. Recall that a map $\nu : \Phi' X_0' \to L\Phi W$, is called correct, if for every $u = \varphi(x_1, \ldots, x_n) \in \Phi' X_0'$ all free variables from the formulas $\nu(u)$ lie in the set $\{x_1, \ldots, x_n\}$.

We know from the previous section that every correct map induces a homomorphism $\zeta : U' \to U$, and every ζ is induced by some correct ν. Given (D, Φ, F), (D', Φ', F'), $\alpha : F \to F'$, $\delta : D' \to D$ and a correct map $\nu : \Phi' X_0' \to L\Phi W$, we assume that for any $u \in \Phi' X_0'$ and $f \in F$ the following formula holds

$$f^{\alpha} * u = \left(f * u^{\nu}\right)^{\delta_*}.$$

It is easy to check that the transition $\nu \to \zeta$ leads to a morphism of DBs, and every morphism of DBs can be obtained in such a way. However, this possibility does not allow yet to get rid from HAs while defining a category of passive DBs. The reason is that we can not well define composition of correct maps.

2.1.4. *Similarity of passive databases.* The notion of similarity of passive DBs plays an essential role in the problem of DB equivalence.

DEFINITION 2.5. Two passive databases (D, Φ, F) and (D', Φ', F') are called similar if there exist bijection $\alpha : F \to F'$, isomorphism $\delta : D' \to D$ and homomorphisms $\zeta : U' \to U$ and $\zeta' : U' \to U$, such that the triple (α, ζ, δ) defines a morphism $(D, \Phi, F) \to (D', \Phi', F')$ and the triple $(\alpha^{-1}, \zeta', \delta^{-1})$ defines a morphism $(D', \Phi', F') \to (D, \Phi, F)$.

Taking into consideration the above remarks, we can say that if such similarity takes place, then for every $f \in F$ the models (D, Φ, f) and $(D', \Phi', f' = f^\alpha)$ are similar.

PROPOSITION 2.6. *Databases (D, Φ, F) and (D', Φ', F') are similar if and only if there exist a bijection $\alpha : F \to F'$, an isomorphism $\delta : D' \to D$ and homomorphisms $\zeta : U \to U'$ and $\zeta' : U' \to U$ such that:*

$$f * u = \left(f^\alpha * u^\zeta\right)^{\delta_*^{-1}}, \qquad \left(f' * u'\right) = \left(f'^{\alpha^{-1}} * u'^{\zeta'}\right)^{\delta_*},$$

where $u \in U$, $u' \in U'$, $f \in F$, $f' \in F'$.

The proof of this proposition directly follows from the definitions.

Our nearest goal is to free the definition of similarity of passive DBs from the use of HAs in it.

THEOREM 2.7. *Two databases (D, Φ, F) and (D', Φ', F') are similar if and only if there exists a bijection $\alpha : F \to F'$, an isomorphism $\delta : D' \to D$ and maps*

$$\nu : \Phi X_0 \to L\Phi' W \quad \text{and} \quad \nu' : \Phi' X_0' \to L\Phi W$$

such that

$$(f * u) = \left(f^\alpha * \nu(u)\right)^{\delta_*^{-1}}, \qquad f' * u' = \left(f'^{\alpha^{-1}} * \nu'(u')\right)^{\delta_*},$$
$$f \in F,\ f' \in F',\ u \in \Phi X_0,\ u' \in \Phi' X_0'.$$

PROOF. First of all let us emphasize that we do not need the maps ν and ν' to be correct. In order to prove the theorem, first prove the following remark.

Let $\alpha : H \to H'$ be a homomorphism of locally finite HAs. As we know, the inclusion $\Delta\alpha(h) \subset \Delta h$ holds always. Let us show that there exists an $h_1 \in H$ with $\alpha(h) = \alpha(h_1)$ and $\Delta(h_1) = \Delta\alpha(h_1)$. Let us take $Y = \Delta h \setminus \Delta\alpha(h)$ and $h_1 = \exists(Y)h$. Since Y does not intersect with the support of the element $\alpha(h)$, we have $\alpha(h_1) = \exists(Y)\alpha(h) = \alpha(h)$. Finally, $\Delta(h_1) = \Delta h \setminus Y = \Delta\alpha(h) = \Delta\alpha(h_1)$.

The next remark is the most important one. Given an epimorphism $\alpha : H \to H'$ and a homomorphism $\beta : U \to H'$, we have $\gamma : U \to H$ with commutative diagram

$$\begin{tikzcd} U \arrow[rr, "\gamma"] \arrow[dr, "\beta"'] & & H \arrow[dl, "\alpha"] \\ & H' & \end{tikzcd}$$

Let us take a basic set U_0 in U. For every $u \in U_0$ we take $\beta(u)$. h is an element in H for which $\alpha(h) = \beta(u)$. We select h_1 for h so that $\alpha(h_1) = \alpha(h) = \beta(u)$ and $\Delta(h_1) = \Delta\alpha(h_1) = \Delta\beta(u) \subset \Delta(u)$. $\gamma(u)$ is defined as $\gamma(u) = h_1$, and this is done for every $u \in U_0$. The map $\gamma : U_0 \to H$ agrees with supports and it can be extended to homomorphism

$\gamma : U \to H$. For every $u \in U_0$ we have $\alpha\gamma(u) = \alpha(h_1) = \alpha(h) = \beta(u)$. Since U_0 generates U, we really get a commutative diagram.

Returning to the proof of the theorem, let us take the algebra $L\Phi W$ of formulas and the Halmos algebra U. For a given database (D, Φ, F) a filter T is picked out in U by the rule: $u \in T$ if $f * u = 1$ holds for every $f \in F$. Consider the transitions

$$L\Phi W \to U \to Q = U/T.$$

Denoting by $\bar{u}$ a corresponding element in U for each formula $u \in L\Phi W$, we have further $\bar{\bar{u}}\, q \in Q$. For every $q \in Q$ we can select u such that $\bar{\bar{u}} = q$, and for $f \in F$ we have

$$f * u = f * \bar{u} = f * \bar{\bar{u}} = f * q \in V_D.$$

If $f * q_1 = f * q_2$ holds for every $f \in F$, then $q_1 = q_2$.

Now let us take $u \in L\Phi W$ and denote by $\sigma(u)$ the union of the sets $\Delta(f * u)$ for all $f \in F$. It is clear that $\sigma(u)$ lies in the set of free variables of the formula u. Besides that, since every $f \in F$ defines a homomorphism $Q \to V_D$, transferring q into $f * q = f * u$, then $\Delta(f * u) \subset \Delta(q)$. This holds for each $f \in F$, and $\sigma(u) \subset \Delta(q)$. Let us show that this is not merely an inclusion, but equality $\sigma(u) = \Delta(q)$. If $x \in \Delta(q)$, then $\exists x q$ and q are different elements in Q. This means that $f * \exists x q = \exists x(f * q) \neq f * q$ for some $f \in F$. Hence, x lies in the support of $\Delta(f * q) = \Delta(f * u)$, and $x \in \sigma(u)$.

For the second database (D', Φ', F') the transitions look like

$$L\Phi' W \to U' \to Q'$$

and here also for every $u' \in L\Phi' W$ we have $\sigma(u') = \Delta(q')$.

Assuming that the conditions of the theorem hold for (D, Φ, F) and (D', Φ', F'), let us check that for every $u \in \Phi X_0$ the set $\sigma(\nu(u))$ lies in the set of free variables of the formula u. It is enough to verify that $\sigma(u) = \sigma(\nu(u))$. We have

$$\sigma(u) = \bigcup_f \Delta(f * u).$$

Since δ_* is isomorphism, then $\Delta(f * u) = \Delta(f * u)^{\delta_*} = \Delta(f^\alpha * \nu(u))$. According to the definition,

$$\sigma\big(\nu(u)\big) = \bigcup_{f^\alpha} \Delta\big(f^\alpha * \nu(u)\big).$$

Therefore, $\sigma(u) = \sigma(\nu(u))$.

Analogously, $\sigma(u') = \sigma(\nu'(u'))$ for $u' \in \Phi X_0'$. So, $\sigma(\nu'(u'))$ lies in the set of free variables of a basic formula u'. Defining a map $\bar{\nu} : \Phi X_0 \to Q'$ by the rule $\bar{\nu}(u) = \overline{\overline{\nu(u)}} = q'$, we get $\Delta(q') = \sigma(\nu(u))$, and $\sigma(\nu(u))$ is among the attributes of the formula u. Now we know that the map $\bar{\nu}$ uniquely extends to a homomorphism $\mu : U \to Q'$, and besides

$\mu(\bar{u}) = \bar{\nu}(u)$. Similarly, we have a map $\bar{\nu}' : \Phi' X_0' \to Q$ which extends to a homomorphism $\mu' : U' \to Q$. So, we have got two diagrams

$$
\begin{tikzcd}
U \arrow[dr, "\mu"'] & & U' \arrow[dl, "\eta'"] \\
& Q' &
\end{tikzcd}
\qquad
\begin{tikzcd}
U' \arrow[dr, "\mu'"'] & & U \arrow[dl, "\eta"] \\
& Q &
\end{tikzcd}
$$

with natural homomorphisms η and η'. Both these diagrams can be converted to commutative diagrams:

$$
\begin{tikzcd}
U \arrow[rr, "\zeta"] \arrow[dr, "\mu"'] & & U' \arrow[dl, "\eta'"] \\
& Q' &
\end{tikzcd}
\qquad
\begin{tikzcd}
U' \arrow[rr, "\zeta'"] \arrow[dr, "\mu'"'] & & U \arrow[dl, "\eta"] \\
& Q &
\end{tikzcd}
$$

It is left to verify, that α, δ, ζ and ζ' satisfy conditions of Proposition 2.6. We will first check it for the homomorphisms μ and μ'. Since the set of all elements from U with the required property is a subalgebra in U and this set contains a basic set, then this set is the whole algebra U. The same holds for the elements from U'.

In order to verify that

$$f * u = \left(f^{\alpha} * u^{\zeta}\right)^{\delta_*^{-1}}$$

holds for every $u \in U$, we use that $\zeta\eta' = \mu$ and $\zeta'\eta = \mu'$. We have

$$f * u = \left(f^{\alpha} * u^{\mu}\right)^{\delta_*^{-1}} = \left(f^{\alpha} * \left(u^{\zeta}\right)^{\eta'}\right)^{\delta_*^{-1}} = \left(f^{\alpha} * u^{\zeta}\right)^{\delta_*^{-1}}.$$

Similarly we can verify that

$$f' * u' = \left(f'^{\alpha^{-1}} * u'^{\zeta'}\right)^{\delta_*},$$

and the theorem is proved in one direction. Let us now prove the converse. Given ζ and ζ' with the necessary properties, we must construct ν and ν'. Take $u \in \Phi X_0$ and let $\bar{u}$ be a corresponding element in U. Proceed to $\bar{u}^{\zeta} \in U'$, and let u' be an element in $L\Phi W$, for which $\bar{u}' = \bar{u}^{\zeta}$. Let $u^{\nu} = u'$. Then

$$f * u = f * \bar{u} = \left(f^{\alpha} * \bar{u}^{\zeta}\right)^{\delta_*^{-1}} = \left(f^{\alpha} * \bar{u}'\right)^{\delta_*^{-1}} = \left(f^{\alpha} * u'\right)^{\delta_*^{-1}} = \left(f^{\alpha} * u^{\nu}\right)^{\delta_*^{-1}}.$$

The map $\nu' : \Phi' X_0' \to L\Phi W$ is constructed in exactly the same way. The theorem is proved.

Based on this theorem, we have the new definition of similarity of passive DBs which does not use HAs.

Let us repeat briefly what had been done. In order to define a category of passive DBs we had to pass to HAs. As for individual morphisms, they can be constructed in the language of formulas, but in this case we cannot compose these morphisms. We define naturally in categorical terms the notion of similarity of two passive bases. At last, we proved that the same notion can be defined in more simple way. □

2.1.5. *Constructions. A product* (D, Φ, F) *of two passive databases* (D_1, Φ_1, F_1) and (D_2, Φ_2, F_2) is defined as follows. An algebra D is a Cartesian product $D_1 \times D_2 = (D_{1i} \times D_{2i},\ i \in \Gamma)$. The set Φ is a set of pairs $\varphi = (\varphi_1, \varphi_2)$, $\varphi_1 \in \Phi_1$, $\varphi_2 \in \Phi_2$, with $\tau(\varphi_1) = \tau(\varphi_2)$. Finally, $F = F_1 \times F_2$. Elements $f \in F$ are written as $f_1 \times f_2$. For every $\varphi = (\varphi_1, \varphi_2)$ of type $\tau = (i_1, \ldots, i_n)$ we set $(f_1 \times f_2)(\varphi) = f_1(\varphi_1) \times f_2(\varphi_2)$. This means that $(f_1 \times f_2)(\varphi)$ is a subset of the Cartesian product $D_{i_1} \times \cdots \times D_{i_n}$ which is represented as $(D_{1i_1} \times \cdots \times D_{1i_n}) \times (D_{2i_1} \times \cdots \times D_{2i_n})$.

There are other constructions for passive DBs, for example, cascade connections and wreath products of DBs (see [BPl3]). Similar constructions are also defined for active DBs, and they are compatible if we pass from passive DBs to active ones. (See [BPl3] for details.)

2.2. *Active databases. The general definition of a database*

2.2.1. *$*$-automata.* Recall, that the scheme of HAs is fixed.

In a first approximation an *active database is a triple* (F, Q, R), *where* F *is a set of states,* Q *is a query algebra and* R *is an algebra of replies to the queries. Both* Q *and* R *are Halmos algebras in the given scheme. It is assumed that there is an operation* $* : F \times Q \to R$ *such that if* $f \in F$ *and* $q \in Q$, *then* $f * q$ *is a reply to the query* q *in the state* f. *For every* $f \in F$ *we consider the map* $\hat{f} : Q \to R$ *which is defined as* $\hat{f}(q) = f * q$, *and assume that* $\hat{f}$ *is always homomorphism of algebras.*

The triple (F, Q, R) *with operation* $*$ *is called also* $*$*-automaton*. It is an automaton of input-output type [PGG]. In the definition of $*$-automaton need not necessarily be based on a variety of HAs but on an arbitrary variety $\mathcal{L}$. Let us give some examples of $*$-automata, connected with such generalization.

2.2.1.1. Given algebras Q and R in $\mathcal{L}$ and a subset F in $\mathrm{Hom}(Q, R)$, consider the triple (F, Q, R) and the set $f * q = f(q)$. Then (F, Q, R) is $*$-automaton in $\mathcal{L}$.

2.2.1.2. Let F be a set and $R \in \mathcal{L}$. Consider the triple (F, R^F, R). R^F is a Cartesian power of algebra R, which also lies in $\mathcal{L}$. For $q \in R^F$ and $f \in F$ we have $f * q = q(f)$. So, we get a $*$-automaton in $\mathcal{L}$.

For an arbitrary (F, Q, R) we have a homomorphism $\nu : Q \to R^F$, defined by $q^\nu(f) = f * q = f * q^\nu$, $q \in Q$, $f \in F$.

Now let us define two kinds of homomorphisms for $*$-automata.

Let (F, Q, R) and $F', Q', R')$ be two $*$-automata. *A homomorphism of the first kind* has the form

$$\rho = (\alpha, \beta, \gamma) : (F, Q, R) \to (F', Q', R'),$$

with a map of sets $\alpha : F \to F'$, homomorphisms of algebras $\beta : Q \to Q'$ and $\gamma : R \to R'$, and the condition:

$$(f * q)^{\gamma} = f^{\alpha} * q^{\beta}, \quad f \in F,\ q \in Q.$$

In the definition of a *homomorphism of the second kind* we proceed from $\beta : Q' \to Q$ and demand

$$\left(f * q^{\beta}\right)^{\gamma} = f^{\alpha} * q.$$

PROPOSITION 2.8. *Given a bijection $\alpha : F \to F'$ and an isomorphism $\gamma : R \to R'$, a homomorphism of $*$-automata*

$$\rho = (\alpha, \beta, \gamma) : (F, Q, R) \to (F', Q', R')$$

is a homomorphism of the first (the second) kind if and only if

$$\rho' = (\alpha^{-1}, \beta, \gamma^{-1}) : (F', Q', R') \to (F, Q, R)$$

is a homomorphism of the second (the first) kind.

From now on $\mathcal{L}$ is always a variety of HAs in the given scheme.

2.2.2. *Definition of an active database.* The universal DB (see 2.1.1)

$$\operatorname{Atm} D = (F_D, U, V_D)$$

is an example of $*$-automaton.

DEFINITION 2.9. An arbitrary active DB is a $*$-automaton (F, Q, R) with a given representation (homomorphism of the second kind)

$$\rho = (\alpha, \beta, \gamma) : (F, Q, R) \to \operatorname{Atm} D.$$

This is an abstract DB. A concrete DB is a subautomaton in $\operatorname{Atm} D$. It is formed in a following way. We take an arbitrary subset F in F_D and subalgebra R in V_D, containing all elements $f * u$, $f \in F$, $u \in U$. Thus, we get an automaton (F, U, R). Take further in U some filter T with the condition $f * u = 1$ for every $f \in F$ and $u \in T$. Passing to $Q = U/T$ with the natural homomorphism $\beta : U \to Q$, we obtain a subautomaton (F, Q, R) in $\operatorname{Atm} D$, which is defined by $f * q = f * u$, if $u^{\beta} = q$. The corresponding representation ρ is defined by the identity embeddings F into F_D, R into V_D and by the homomorphism $\beta : U \to Q$. Usually we will deal with concrete DBs.

PROPOSITION 2.10.

(1) *Let*

$$\nu = (\nu_1, \nu_2, \nu_3) : (F, Q, R) \to (F', Q', R')$$

be a homomorphism of the first kind of concrete DBs. Then the maps ν_1 *and* ν_3 *are injective.*

(2) *If* ρ *is a homomorphism of the second kind and* $\nu_2 : Q' \to Q$ *is a surjection, then* ν_1 *and* ν_3 *are injections.*

PROOF. As for ν_3, it is always an injection because algebra R is a simple one. So, we should concentrate on ν_1.

Let ν be a homomorphism of the first kind, $f_1^{\nu_1} = f_2^{\nu_1}$, and the first DB is associated with the algebra U, for which we have a surjection $\beta : U \to Q$. For an arbitrary $u \in U$

$$(f_1 * u)^{\nu_3} = (f_1 * u\beta)^{\nu_3} = f_1^{\nu_1} * u^{\beta\nu_2} = f_2^{\nu_1} * u^{\beta\nu_2} = (f_2 * u)^{\nu_3}.$$

Since ν_3 is injection, then $f_1 * u = f_2 * u$. This is true for every $u \in U$, hence $f_1 = f_2$. The proof of the second point can be found in [BPl3].

It is evident that homomorphisms of databases connect in some sense evaluation of a reply to a query in these databases. Let, for example,

$$\nu = (\nu_1, \nu_2, \nu_2) : (F, Q, R) \to (F', Q', R')$$

be a homomorphism of the first kind. Then

$$(f * q)^{\nu_3} = f^{\nu_1} * q^{\nu_2}.$$

Due to the fact that ν_3 is injection, we can say that a reply to a query in the first DB can be obtained by means of the second one. The opposite is not true in general, because it can be that $f' \neq f^{\nu_1}$ for some $f' \in F'$, or $q' \in Q'$ could be not equal to some q^{ν_2}, $q \in Q$.

A homomorphism of databases ν is an *isomorphism* if all the three maps ν_1, ν_2 and ν_3 are bijections. If two DBs are isomorphic, then a reply to a query in each of them can be obtained by the means of another one.

A database (F, Q, R) *is called faithful, if for every pair of elements* $q_1, q_2 \in Q$ *the equalities* $f * q_1 = f * q_2$ *for all* $f \in F$ *imply* $q_1 = q_2$. To every database (F, Q, R) corresponds a faithful database (F, Q_0, R), where Q_0 is a result of squeezing (quotienting) the algebra Q. □

DEFINITION 2.11. Let (F, Q, R) and (F', Q', R') be two active databases. A generalized isomorphism of these databases is a tuple $(\alpha, \eta, \eta', \gamma)$, where $\alpha : F \to F'$ is a bijection, $\gamma : R \to R'$ is an isomorphism and $\eta : Q \to Q'$, $\eta' : Q' \to Q$ are homomorphisms, such that

$$(f * q)^\gamma = f^\alpha * q^\eta, \qquad f' * q' = \left(f'^{\alpha^{-1}} * q'^{\eta'}\right)^\gamma,$$

$f \in F$, $f' \in F'$, $q \in Q$, $q' \in Q'$.

PROPOSITION 2.12. *If there exists a generalized isomorphism* $(\alpha, \eta, \eta', \gamma)$ *for databases* (F, Q, R) *and* (F', Q', R'), *then the corresponding faithful databases are isomorphic.*

PROOF. Let (F, Q_0, R) and (F', Q'_0, R') be the corresponding faithful DBs, $\beta : Q \to Q_0$ and $\beta' : Q' \to Q'_0$ be the natural homomorphisms, and T and T' be the kernels of these homomorphisms. We want to verify that $T \subset \operatorname{Ker} \eta\beta'$. Take $q \in T$ and $f \in F$. Then $f^\alpha * q^{\eta\beta'} = f^\alpha * q^\eta = (f * q)^\gamma = 1^\gamma = 1$. Since it holds for every $f \in F$, then $q^{\eta\beta'} = 1$, $q \in \operatorname{Ker} \eta\beta'$. We can similarly check that $T' \subset \operatorname{Ker} \eta'\beta$. So, $\eta : Q \to Q'$ induces $\bar\eta : Q_0 \to Q'_0$, and $\eta' : Q' \to Q$ induces $\bar\eta' : Q'_0 \to Q_0$. If $q \in Q$, then $\bar\eta(q^\beta) = q^{\eta\beta'}$, and for $q' \in Q'$ we have $\bar\eta'(q'^{\beta'}) = q'^{\eta'\beta}$.

If, further, $f \in F$, $q_0 \in Q_0$, $q_0 = q^\beta$, then

$$(f * q_0)^\gamma = \left(f * q^\beta\right)^\gamma = (f * q)^\gamma = f^\alpha * q^\eta = f^\alpha * q^{\eta\beta'} = f^\alpha * q_0^{\bar\eta}$$

and, similarly,

$$f' * q'_0 = \left(f'^{\alpha^{-1}} * q_0^{\bar\eta'}\right)^\gamma.$$

It means that the tuple $(\alpha, \bar\eta, \bar\eta', \gamma)$ defines a generalized isomorphism of faithful databases (F, Q_0, R) and (F', Q'_0, R'). It is left to verify, that $\bar\eta$ and $\bar\eta'$ are inverse to each other.

Take $f \in F$, $q_0 \in Q_0$, $q_0 = q^\beta$. Then

$$\begin{aligned} f * q_0^{\bar\eta\bar\eta'} &= f * q^{\beta\bar\eta\bar\eta'} = f * q^{\eta\beta'\bar\eta'} = f * q^{\eta\eta'\beta} = f * q^{\eta\eta'} \\ &= \left(f^\alpha\right)^{\alpha^{-1}} * \left(q^\eta\right)^{\eta'} = \left(f^\alpha * q^\eta\right)^{\gamma^{-1}} = (f * q)^{\gamma\gamma^{-1}} = f * q = f * q_0. \end{aligned}$$

Since it holds for every $f \in F$, then $q_0^{\bar\eta\bar\eta'} = q_0$. Similarly, if $q'_0 \in Q'_0$, then $q_0'^{\bar\eta'\bar\eta} = q'_0$. Therefore, $\bar\eta$ and $\bar\eta'$ are inverse to each other. So, a triple $(\alpha, \bar\eta, \gamma)$ defines isomorphism of the databases (F, Q_0, R) and (F', Q'_0, R').

A database (F, Q, R) *is called reduced, if it is faithful and the algebra* R *is generated by the elements of the kind* $f * q$, $f \in F$, $q \in Q$. *With each* (F, Q, R) *a certain reduced database can be associated.* □

DEFINITION 2.13. Databases (F, Q, R) and (F', Q', R') are called similar, if the corresponding reduced ones are isomorphic.

PROPOSITION 2.14. *Let* $\operatorname{Atm} D = (F_D, U, V_D)$ *be the universal database, and* F *be a subset in* F_D. *Then all possible* (F, Q, R) *in* $\operatorname{Atm} D$, *with the given* F *are similar pairwise.*

PROOF. Given F, let R_F be a subalgebra in V_D, generated by all elements of the kind $f * u$, $f \in F$, $u \in U$. We have (F, U, R_F) and denote by (F, Q_F, R_F) the corresponding faithful DB. Then the latter DB is reduced, and each (F, Q, R) can be reduced to the base of such kind.

How can the set F be described? One of the approaches is as follows: let T be some set of formulas in U. Define $T' = F$ by the rule: $f \in F$, if $f * v = 1$ for every $v \in T$. Let

$T'' = F'$ be a set of all $u \in U$, such that $f * u = 1$ for every $f \in F$. If $Q = U/T''$, and R is the subalgebra in V_D, generated by all $f * u$, $f \in F$, $u \in U$, then (F, Q, R) is the reduced DB, defined by the set of axioms T.

Consider now the elementary theories of models from the point of view of isomorphisms of DBs. □

PROPOSITION 2.15. *Let $\rho = (\alpha, \xi, \gamma) : (F, U, R) \to (F', U', R')$ be an isomorphism of databases, where (F, U, R) and (F', U', R') are associated with data algebras D and D' respectively. Then, for every $f \in F$ we have*

$$(\operatorname{Ker} \hat{f})^{\xi} = \operatorname{Ker} \hat{f}^{\alpha}.$$

This means that the elementary theories of the models (D, Φ, f) and (D', Φ', f^{α}) are compatible.

In the next proposition all algebras of queries and algebras of replies are HAs with equalities. Denote by U_0 a HA with equalities and with empty Φ. For every U there is a canonical embedding $\pi : U_0 \to U$.

PROPOSITION 2.16. *If there is an isomorphism $\nu = (\nu_1, \nu_2, \nu_3) : (F, Q, R) \to (F', Q', R')$ of concrete databases with data algebras D and D', then the elementary theories of D and D' coincide.*

PROOF. Let us take the corresponding Halmos algebras U and V_D and U' and $V_{D'}$, with homomorphisms $\beta : U \to Q$ and $\beta' : U' \to Q'$. Consider homomorphisms

$$f_D : U_0 \to V_D \quad \text{and} \quad f_{D'} : U_0 \to V_{D'}.$$

We have to verify that their kernels coincide. Let $u \in \operatorname{Ker} f_D$. Then $f_D * u = f * \pi(u) = 1$ for an arbitrary $f \in F$. We have

$$\big(f * \pi(u)\big)^{\nu_3} = \big(f * \pi(u)^{\beta}\big)^{\nu_3} = f^{\nu_1} * \pi(u)^{\beta\nu_2} = 1.$$

Let now $\pi' : U_0 \to U'$ be a canonical embedding of the algebra U_0 into U'. Then $\pi(u)^{\beta\nu_2} = \pi'(u)^{\beta'}$, which leads to

$$f^{\nu_1} * \pi'(u)^{\beta'} = f^{\nu_1} * \pi'(u) = f_{D'} * u = 1.$$

So, $u \in \operatorname{Ker} f_{D'}$. Similarly, $u \in \operatorname{Ker} f_{D'}$ implies $u \in \operatorname{Ker} f_D$. Hence, $\operatorname{Ker} f_D = \operatorname{Ker} f_{D'}$. In particular we can say that if algebras D and D' are finite, then they are isomorphic. □

2.2.3. *Category of active databases.* Let us start with the recalling that if $\operatorname{Atm} D = (F_D, U, V_D)$ and $\operatorname{Atm} D' = (F_{D'}, U', V_{D'})$ are two universal databases, $\delta : D' \to D$ is an epimorphism of data algebras, and $\zeta : U' \to U$ is a homomorphism of HAs, then the pair (δ, ζ) defines an embedding

$$\eta = (\delta^*, \zeta, \delta_*) : \operatorname{Atm} D' \to \operatorname{Atm} D,$$

which is an injection of the second kind. Assume, that for $*$-automata (F, Q, R) and (F', Q', R') the representations

$$\rho = (\alpha, \beta, \gamma) : (F, Q, R) \to \operatorname{Atm} D$$

and

$$\rho' = (\alpha', \beta', \gamma') : (F', Q', R') \to \operatorname{Atm} D'$$

are given.

Finally, suppose that there is a homomorphism of the second kind

$$\nu = (\nu_1, \nu_2, \nu_3) : (F, Q, R) \to (F', Q', R').$$

We consider now the problem: which conditions have to be fulfilled in order that one can speak about homomorphisms of the corresponding abstract DBs. First of all, it is natural to take the diagram

$$\begin{tikzcd}
(F, Q, R) \arrow[r, "\rho"] \arrow[d, "\nu"'] & \operatorname{Atm} D \arrow[d, "\eta"] \\
(F', Q', R') \arrow[r, "\rho'"] & \operatorname{Atm} D'
\end{tikzcd}$$

Commutativity of this diagram means that there are the following three commutative diagrams

$$\begin{tikzcd}
F \arrow[r, "\alpha"] \arrow[d, "\nu_1"'] & F_D \arrow[d, "\delta^*"] \\
F' \arrow[r, "\alpha'"] & F_{D'}
\end{tikzcd}
\qquad
\begin{tikzcd}
Q & U \arrow[l, "\beta"'] \\
Q' \arrow[u, "\nu_2"] & U' \arrow[l, "\beta'"] \arrow[u, "\zeta"']
\end{tikzcd}
\qquad
\begin{tikzcd}
R \arrow[r, "\gamma"] \arrow[d, "\nu_3"'] & V_D \arrow[d, "\delta_*"] \\
R' \arrow[r, "\gamma'"] & V_{D'}
\end{tikzcd}
\qquad (*)$$

These diagrams can be taken as a definition of a homomorphism of DBs. However, such a definition is too rigid since we need not be bound so tightly to the initial η, which is based on the pair (δ, ζ).

Also it is not necessary that the homomorphism ν_3 is induced by the homomorphism δ_*, and ν_1 is not necessarily rigidly tied up with δ^*. Thus, we will only demand the presence of the second diagram which connects ν_2 and ζ, while in the first diagram we proceed from the following weakened form of commutativity: for every $f \in F$

$$f^{\nu_1 \alpha'} \leqslant f^{\alpha \delta^*}.$$

This condition is equivalent to the following one: for every basic formula $u \in U'$ there holds

$$f^{\nu_1 \alpha'} * u \subseteq f^{\alpha \delta^*} * u = \left(f^{\alpha} * u^{\zeta}\right)^{\delta_*}.$$

As we know from the previous section, the last means that the homomorphism $\delta : D' \to D$ together with the map $f^{\nu_1\alpha'} \to f^{\alpha}$ define homomorphism of models

$$\left(D', \Phi', f^{\nu_1\alpha'}\right) \to \left(D, \Phi, f^{\alpha}\right).$$

We does not assume any conditions of commutativity for the third diagram.

Thus, a *homomorphism of databases* has the form $\nu = (\nu_1, \nu_2, \nu_3) : (F, Q, R) \to (F', Q', R')$, subject to conditions, described by the second commutative diagram, and the first weakened commutative diagram in $(*)$.

Homomorphisms of DBs, as just defined, lead to a category of active DBs. Separately a category of concrete DBs can be considered.

In this section we discussed active DBs. The next subsection is devoted to the aggregation of active and passive databases into a common structure.

2.2.4. *General definition of a database.* First, we describe a transition from passive DBs to active ones. Let (D, Φ, F) be a passive DB. Given an algebra $D \in \Theta$, take $\operatorname{Atm} D = (F_D, U, V_D)$ in a fixed scheme. Take a subset $F \subset F_D$ and pass to the active database (F, U, V_D).

Let (D', Φ', F') be another passive DB. Recall that a morphism of passive databases $(D, \Phi, F) \to (D', \Phi', F')$ is defined by an epimorphism $\delta : D' \to D$, by a map $\alpha : F \to F'$, by a homomorphism $\zeta : U' \to U$, and the condition

$$\left(f * u'^{\zeta}\right)^{\delta_*} = f^{\alpha} * u', \quad u' \in U'.$$

The same condition means that the triple $(\alpha, \zeta, \delta_*)$ determines a homomorphism (of the second kind) $(F, U, V_D) \to (F', U', V_{D'})$ for active DBs. Thus there arises a *functor from a category of passive DBs to a category of active ones.* However, this functor does not allow to recover uniquely a passive DB from an active one. Basic information, which we must not forget, cannot be recovered. This leads us to a new view on the notion of DB.

DEFINITION. *A database is an* object of the form

$$(D, \Phi, F) \to (F, U, V_D) \to (F, Q, R)$$

where the first arrow is associated with the transition from a passive to an active DB, and the second one is defined by the representation $\rho : (F, Q, R) \to (F, U, V_D)$ which is the identity on F.

If, moreover, ρ is the identity on R, then it is a concrete DB.

Now a passive DB is the passive part of a general construction of a DB, while an active DB is its active part. The passive part is source information while the active one reflects the functioning of the system "Query–Reply". *Morphisms in the category of DBs* now look like

$$\begin{array}{ccccc} (D,\Phi,F) & \longrightarrow & (F,U,V_D) & \longrightarrow & (F,Q,R) \\ \downarrow & & \downarrow & & \downarrow \\ (D',\Phi',F') & \longrightarrow & (F',U',V_{D'}) & \longrightarrow & (F',Q',R') \end{array}$$

with the first vertical arrow being defined by an epimorphism $\delta : D' \to D$, by a homomorphism $\zeta : U' \to U$ and by a map $\alpha : F \to F'$. It is a morphism in a category of passive DBs. The second vertical arrow is a triple $(\alpha, \zeta, \delta_*)$, and the third one is $\nu = (\nu_1, \nu_2, \nu_3)$. The second and the third arrows have to be compatible with the corresponding representations ρ and ρ', described in the previous section. The second and the third vertical arrows determine a morphism in the category of active DBs.

The developed model of a DB agrees with the intuition of a DB could be: first a DB is defined on the level of basic, source information, then it generates a lot of information – the "world" of all associated information, and finally this very big world is reduced to some reasonable size.

2.2.5. *Similarity of passive and active databases.* We will somehow narrow the definition of similarity of active DBs. Now we call two databases (F, Q, R) and (F', Q', R') *similar*, if for the corresponding reduced active data bases (F, Q_0, R_0) and (F', Q_0, R_0') there is an isomorphism $\rho = (\alpha, \beta, \gamma)$, in which $\gamma : R_0 \to R_0'$ is induced by an isomorphism $\delta : D' \to D$ of data algebras. As we will see (2.3.1), this additional condition holds automatically in the natural situation of finite D and D' and presence of equalities in Q and R.

THEOREM 2.17. *Two passive databases (D, Φ, F) and (D', Φ', F') are similar if and only if the corresponding active databases (F, U, V_D) and $(F', U', V_{D'})$ are similar.*

PROOF. We remind that passive databases (D, Φ, F) and (D', Φ', F') are similar if there exist a bijection $\alpha : F \to F'$, an isomorphism $\delta : D' \to D$ and maps

$$\nu : \Phi X_0 \to L\Phi' W \quad \text{and} \quad \nu' : \Phi' X_0' \to L\Phi W,$$

for which hold

$$f * u = \left(f^\alpha * \nu(u)\right)^{\delta_*^{-1}}; \qquad f' * u' = \left(f'^{\alpha^{-1}} * \nu'(u')\right)^{\delta_*},$$

where $f \in F$, $f' \in F'$, $u \in \Phi X_0$, $u' \in \Phi' X_0$.

Let all these conditions be fulfilled. It follows from the definition of similarity that, given ν and ν', we can build homomorphisms $\zeta : U \to U'$ and $\zeta' : U' \to U$ such that for every $u \in U$ and $u' \in U'$ hold

$$f * u = \left(f^\alpha * u^\zeta\right)^{\delta_*^{-1}}, \qquad f' * u' = \left(f'^{\alpha^{-1}} * u'^{\zeta'}\right)^{\delta_*}.$$

This means that the tuple $(\alpha, \zeta, \zeta', \delta_*)$ defines a generalized isomorphism of active databases (F, U, V_D) and $(F', U', V_{D'})$. This generalized isomorphism induces an ordinary isomorphism

$$(\alpha, \zeta_0, \delta_*) : (F, Q, V_D) \to (F', Q', V_{D'})$$

for corresponding faithful DBs. This isomorphism is compatible with the homomorphism $(\alpha, \zeta, \delta_*) : (F, U, V_D) \to (F', U', V_{D'})$. The inverse one $(\alpha^{-1}, \zeta_0^{-1}, \delta_*^{-1}) : (F', Q', V_{D'}) \to (F, Q, V_D)$ is compatible with the homomorphism $(\alpha^{-1}, \zeta', \delta_*^{-1}) : (F', U', V_{D'}) \to (F, U, V_D)$. It is also clear that $(\alpha, \zeta_0, \delta_*)$ is simultaneously an isomorphism of reduced DBs, and DBs (F, U, V_D) and $(F', U', V_{D'})$ are similar in the pointed above strengthened sense. $\square$

We omit the proof of the converse statement (see [TPl2]).

2.3. *Groups and databases*

2.3.1. *Galois theory of the algebra V_D.* Let G be the group of all automorphisms of a data algebra D, $G = \text{Aut } D$. Every $g \in G$ is an isomorphism $g : D \to D$, which induces isomorphism $g_* : V_D \to V_D$. It gives a representation of the group G as a group of automorphisms of algebra V_D:

$$G = \text{Aut } D \to \text{Aut } V_D.$$

THEOREM 2.18. *The representation $G \to$ Aut V_D is an isomorphism of groups* Aut D *and* Aut V_D.

If $A \in V_D$, then we write gA instead of g_*A for every $g \in G$.

Now let us consider *Galois correspondence*. Let R be a subset in the algebra V_D. Denote by $R' = H$ the set of all $g \in G$ for which $gA = A$ holds for every $A \in R$. H is a subgroup in G. If, further, H is a subset in G, then $H' = R$ denotes a set of all $A \in V_D$ for which $gA = A$ holds for every $g \in H$. R is a subalgebra in V_D.

For a subset $R \subset V_D$ we have $R' = H$, $H' = R''$, $R \subset R''$. The subalgebra R'' is the Galois closure of the set R. This Galois closure is always a subalgebra in V_D. The same holds for a subset H in G: we take its Galois closure H'', $H \subset H''$ and H'' is always a subgroup in G.

We will consider the algebra V_D as an algebra with equalities. The theorem also holds in this case, and each algebra of the form $R = H'$ is an algebra with equalities.

Besides that, we will restrict ourselves with the case when all data algebras D are finite and all X_i in the set $X = (X_i,\ i \in \Gamma)$ are countable. In Theorem 2.19 we assume that these conditions hold.

THEOREM 2.19. *For every $H \subset G$, its closure H'' is a subgroup in G, generated by the set H. If R is a subset in V_D, then R'' is a subalgebra (with equalities), generated by the set R.*

If follows from this theorem, that there is a *one-to-one correspondence between the subgroups in G and the subalgebras with equalities in* V_D. In particular, there is only finite number of different subalgebras in V_D.

Let us now consider two algebras D_1 and D_2 (in particular, they may coincide), and the corresponding algebras V_{D_1} and V_{D_2}. We have $G_1 = \mathrm{Aut}\ D_1$ and $G_2 = \mathrm{Aut}\ D_2$.

THEOREM 2.20. *Let R_1 be a subalgebra in V_{D_1} and R_2 a subalgebra in V_{D_2}. Then*

(1) *R_1 and R_2 are isomorphic if and only if the corresponding subgroups $H_1 = R_1' \subset G_1$ and $H_2 = R_2' \subset G_2$ are conjugated by some isomorphism $\delta : D_2 \to D_1$.*
(2) *If R_1 and R_2 are isomorphic and $\gamma : R_1 \to R_2$ is the isomorphism, then this isomorphism is induced by some isomorphism $\delta : D_2 \to D_1$.*
(3) *If $R_2 = R_1^{\gamma} = R_1^{\delta_*}$, then $R_1' = \delta^{-1} R_2' \delta$, that is $H_1 = \delta^{-1} H_2 \delta$.*

Here, if $g \in G_2$, $g : D_2 \to D_2$ is an automorphism and $\delta : D_2 \to D_1$ is an isomorphism, then $\delta^{-1} g \delta$ is an automorphism of algebra D_1, i.e. an element in G_1. $\delta^{-1} H_2 \delta$ is, as usual, a set of all $\delta^{-1} g \delta$, $g \in H_2$.

It is also obvious that for each isomorphism $\delta : D_2 \to D_1$ and each $R_1 \subset V_{D_1}$, if $R_1^{\delta_*} = R_2$, then δ_* gives the isomorphism between R_1 and R_2.

These two useful theorems go back to the works of M.I. Krasner [Kr] which describe pure relational algebras. F. Bancillhon [Ban] and E. Beniaminov [Ben3] showed how to use them in databases. Beniaminov has constructed also an interesting categorical proof which allows to make generalizations. These results were extended to pure HAs by various authors. For HAs in some variety Θ see [BPl3] and [BPl4]. See also [Ts3,Dg,Bo,Maf, MSS].

2.3.2. *Additional remarks.* There are three Galois theories with one and the same group $G = \mathrm{Aut}\ D$. One of them deals with subalgebras of V_D, the second one is aimed at subbases of the universal DB, and the last one relates to subalgebras in the algebra D. The first theory was discussed in 2.3.1. In the book [BPl3] Galois theory for a universal database (F_D, U, V_D) is developed. It allows to classify concrete DBs with the given data algebra D. The group of automorphisms of a universal DB is treated in [MP]. Finally, the Galois correspondence between subgroups of $G = \mathrm{Aut}\ D$ and subalgebras in D, is in some sense connected with the material of the above subsection (see [BPl3] for details). In this book also some applications to DBs are given. In the next section we will present one more example of this kind.

3. Applications and other aspects

3.1. *The problem of equivalence of databases*

3.1.1. *Statement of the problem.* In this subsection we focus our attention on informational equivalence of DBs. Intuitively, the problem is to point out conditions, under which a reply to a query in one DB can be obtained by means of the second one, and vice versa.

Such a problem arises, for example, in the following situation. Let two forms keep a look-out for one and the same object in the same conditions, but the observations are organized differently. Data are stored in, possibly, different systems of relations. In which cases we can assert that the collected data are equivalent and, based on them, we can get well corresponding inferences?

The problem is also in coordination of the corresponding means and, first of all, in recognition of the possibility of such coordination.

As we know, a DB includes some source (basic) information and we can speak of derivative information which is derived from the basic information.

The equivalence problem can be solved on the level of basic information, as well as on a higher level, taking into consideration all possible derived information. Intuitively it is clear that the solution of the problem on the level of basic information implies its solution on any desired high level. But we need precise constructions in order to obtain the necessary maps.

The problem of DBs equivalence has attracted attention of specialists for many years (e.g., [BMSU,Ulm,BK,Ts3,Br1,TPPK,BN], etc.), and always the solution depends on the chosen approach to the definition of the notion of DB (i.e. on the chosen model of a DB). We will proceed from the model considered above, and limit ourselves to concrete DBs (see details in [TPl2]).

3.1.2. *Definition of the notion of database equivalence.*

DEFINITION 3.1. Two concrete databases

$$(D, \Phi, F) \to (F, U, V_D) \to (F, Q, R),$$

$$(D', \Phi', F) \to (F, U', V_{D'}) \to (F', Q', R')$$

are called equivalent, if one of the following conditions takes place:

(1) The passive parts of these DBs are similar,

(2) *The active parts of these DBs are similar.*

Equivalence of these two conditions was considered in 2.2.5. It confirms the feeling that DBs' equivalence is recognized on the *level of basic information.*

We want to check that this definition of equivalence reflects the intuition of the idea of equivalence, as stated above. Let (F, Q_0, R_0) and (F', Q'_0, R'_0) be reduced $*$-automata for (F, Q, R) and (F', Q', R') respectively. If DBs are equivalent, then by the definition of similarity of active DBs, there is an isomorphism

$$\rho = (\alpha, \beta, \gamma) : (F, Q_0, R_0) \to \left(F', Q'_0, R'_0\right),$$

but here we do not require the isomorphism γ to be induced by an isomorphism $\delta : D' \to D$.

Let us consider also the natural homomorphisms $\beta_0 : Q \to Q_0$ and $\beta_0' : Q' \to Q_0'$. Take $f \in F$, $q \in Q$, $f' \in F'$, $q' \in Q'$. Then

$$f * q = f * q^{\beta_0} = \left(f^{\alpha} * q^{\beta_0 \beta}\right)^{\gamma^{-1}},$$

$$f' * g' = f' * q'^{\beta_0'} = \left(f'^{\alpha^{-1}} * q'^{\beta_0' \beta^{-1}}\right)^{\gamma}.$$

Hence, a reply to a query in the first DB can be obtained via the second one, and vice versa.

In fact, the definition of equivalence is much more rigid than is required to realize the initial premise. It links too tightly the structures of the corresponding DBs. Therefore, we will weaken this definition.

3.1.3. *Local isomorphism.* Given a database (F, Q, R), consider a homomorphism $\hat{f} : Q \to R$, $f \in F$. Let R_f be the image of this homomorphism and T_f be its kernel. Passing to $Q_f = Q/T_f$, we obtain a *local database* (f, Q_f, R_f) with one state f, which, obviously, is reduced.

DEFINITION 3.2. Two active databases (F, Q, R) and (F', Q', R') are called *locally isomorphic*, if there is a bijection $\alpha : F \to F'$, such that for every $f \in F$ the algebras R_f and R'_{f^α} are isomorphic.

We will write R_{f^α} instead of R'_{f^α}, T_{f^α} instead of T'_{f^α} and Q_{f^α} instead of Q'_{f^α}, because the presence of f^α indicates that the algebras belong to the second DB. Let $\gamma_f : R_f \to R_{f^\alpha}$ be some isomorphism. Then we have a commutative diagram of isomorphisms

$$\begin{tikzcd} Q_f \arrow[r] \arrow[d, "\beta_f"'] & R_f \arrow[d, "\gamma_f"] \\ Q_{f^\alpha} \arrow[r] & R_{f^\alpha} \end{tikzcd}$$

whose horizontal arrows are associated with $\hat{f}$ and $\hat{f}^\alpha$. Now for every $q \in Q_f$ we have

$$(f * q)^{\gamma_f} = f^\alpha * q^{\beta_f},$$

which means that under the conditions above the local databases (f, Q_f, R_f) and $(f^\alpha, Q_{f^\alpha}, R_{f^\alpha})$ are also isomorphic.

PROPOSITION 3.3. *If two databases (F, Q, R) and (F', Q', R') are similar, then they are locally isomorphic.*

PROOF. Let us pass to corresponding reduced DBs with isomorphism

$$(\alpha, \beta, \gamma) : (F, Q_0, R_0) \to (F', Q_0', R_0').$$

The homomorphism $\hat{f}: Q \to R$ induces a homomorphism $\hat{f}: Q_0 \to R_0$ and the images of these homomorphisms coincide with R_f. It is easy to understand that the isomorphism γ induces an isomorphism $\gamma_f: R_f \to R_{f^\alpha}$. We can also check that there is a commutative diagram with natural homomorphisms β_0 and β_0' and $\beta_f: Q_f \to Q_{f^\alpha}$ as follows:

$$\begin{tikzcd} Q_0 \arrow[r, "\beta"] \arrow[d, "\beta_0"'] & Q_0' \arrow[d, "\beta_0'"] \\ Q_f \arrow[r, "\beta_f"] & Q_{f^\alpha} \end{tikzcd}$$

On the other hand, reduced locally isomorphic DBs can be not isomorphic.

Now, let us look at the notion of local isomorphism from the other point of view. We will proceed from the idea of generalized isomorphism, considered in Definition 2.11. A generalized isomorphism implies similarity and, hence, local isomorphism.

Given (F, Q, R) and (F', Q', R'), consider a tuple $(\alpha, \beta, \beta', \gamma)$ with a bijection $\alpha: F \to F'$ and functions β, β', γ. The function β for every $f \in F$ gives a homomorphism $\beta_f: Q \to Q'/T_f'$, where $T_f' \subset \operatorname{Ker}\widehat{f^\alpha}$ is a filter in Q'. It follows from the definition of β_f, that an element $f^\alpha * q^{\beta_f}$ is always determined. The function β' acts similarly in the opposite direction, defining $\beta'_{f'}: Q' \to Q/T_{f'}$. And finally, the function γ for every $f \in F$ gives an isomorphism $\gamma_f: R_f \to R_{f^\alpha}$.

Suppose the following conditions hold:

$$(f * q)^{\gamma_f} = f^\alpha * q^{\beta_f}; \qquad f' * q' = \left(f'^{\alpha^{-1}} * q'^{\beta'_{f'}}\right)^{\gamma_f}$$

with $f \in F, q \in Q, q' \in Q', f' \in F'$.

Under this condition a tuple $(\alpha, \beta, \beta', \gamma)$ is called a *crossed isomorphism of DBs*. It is evident that an isomorphism and a generalized isomorphism are particular cases of a crossed isomorphism. In these cases the corresponding functions β, β', γ do not depend on f, they are constants, and all filters $T_{f'}$ and T_f are trivial.

It is also clear that if there is a crossed isomorphism between DBs, then the main premises of equivalence in the sense of evaluation of a reply to a query hold. □

PROPOSITION 3.4. *Two databases are locally isomorphic if and only if one can establish a crossed isomorphism between them.*

PROOF. It directly follows from the definition that if between DBs there is a crossed isomorphism, then they are locally isomorphic. Let us prove the opposite. Let the databases (F, Q, R) and (F', Q', R') be locally isomorphic. We have a bijection $\alpha: F \to F'$ and an isomorphism $\gamma_f: R_f \to R_{f^\alpha}$ for every $f \in F$. Fix this γ_f for each f, and take $T_f = \operatorname{Ker}\hat{f}$ and $T_{f'} = \operatorname{Ker}\widehat{f'}$ for $f' \in F'$.

In the commutative diagrams

$$\begin{tikzcd} Q/T_f \arrow[r, "\bar{f}"] \arrow[d, "\psi_f"'] & R_f \arrow[d, "\gamma_f"] \\ Q'/T_{f'} \arrow[r, "\bar{f}'"] & R_{f'} \end{tikzcd} \qquad \begin{tikzcd} Q/T_f \arrow[r, "\bar{f}"] & R_f \\ Q'/T_{f'} \arrow[r, "\bar{f}'"] \arrow[u, "\psi_f^{-1}"] & R_{f'} \arrow[u, "\gamma_f^{-1}"'] \end{tikzcd}$$

$f' = f^{\alpha}$, the isomorphism ψ_f was earlier denoted by $\beta_{f'}$, and $\bar{f}$ is the isomorphism, induced by $\hat{f}$. Let us now take the natural homomorphisms

$$\eta_f : Q \to Q/T_f \quad \text{and} \quad \eta'_{f'} : Q' \to Q'/T_{f'}.$$

Let $\beta_f = \eta_f \psi_f : Q \to Q'/T'_f$, where $T'_f = T_{f'}$, and $\beta'_{f'} = \eta'_{f'} \psi_f^{-1} : Q' \to Q/T_{f'}$, where $T_{f'} = T_f$. We have a tuple $(\alpha, \beta, \beta', \gamma)$ and we must check the relations between its components. Let $f \in F$, $f' = f^{\alpha} \in F'$, $q \in Q$ and $q' \in Q'$. Then

$$(f * q)^{\gamma_f} = \left(f * q^{\eta_f}\right)^{\gamma_f} = q^{\eta_f \bar{f} \gamma_f} = q^{\eta_f \psi_f \bar{f}'} = f^{\alpha} * q^{\beta_f}.$$

Similarly we can check that $f' * q' = (f * q'^{\beta'_{f'}})^{\gamma_f}$. The proposition is proved. □

3.1.4. *The general notion of equivalence.* As it was done before, we start from passive databases (D, Φ, F) and (D', Φ', F'). The modified definition of similarity, using the idea of localization, is as follows.

DEFINITION. Two passive databases (D, Φ, F) and (D', Φ', F') are similar, if there is a bijection $\alpha : F \to F'$, isomorphisms $\delta_f : D' \to D$ for every f, and maps

$$\nu_f : \Phi X_0 \to L\Phi' W \quad \text{and} \quad \nu'_{f'} : \Phi' X'_0 \to L\Phi W$$

such that the equalities

$$(f * u)^{\delta_f *} = f^{\alpha} * \nu_f(u), \qquad f' * u' = \left(f * \nu'_{f'}(u')\right)^{\delta_f *}$$

hold for

$$f \in F, \qquad f' = f^{\alpha} \in F', \qquad u \in \Phi X_0, \qquad u' \in \Phi' X'_0.$$

Let us now redefine the notion of similarity of active DBs.

DEFINITION. Active databases (F, Q, R) and (F', Q', R') with data algebras D and D' are similar, if they are locally isomorphic and every $\gamma_f : R_f \to R_{f^{\alpha}}$ under the corresponding $\alpha : F \to F'$ is induced by some $\delta_f : D' \to D$.

THEOREM 3.5. *Two passive databases (D, Φ, F) and (D', Φ', F') are similar if and only if the corresponding active databases (F, U, V_D) and $(F', U', V_{D'})$ are similar.*

PROOF. Let the passive DBs be similar and let the tuple $(\alpha, \nu, \nu', \delta)$ give the similarity. Here ν, ν' and δ depend on $f \in F$ or $f' \in F'$. Let us pass to the active databases (F, U, V_D) and $(F', U', V_{D'})$ and prove that they are similar. We will show that for every $f \in F$ the isomorphism $\delta_f : D' \to D$ induces an isomorphism between the corresponding R_f and R_{f^α} with $R = V_D$ and $R' = V_{D'}$. Denote the image of the set ΦX_0 in U by U_0 and the image of $\Phi' X_0'$ in U' by U_0'. This U_0 generates an algebra U, while U_0 generates U'. The algebra R_f is generated by all $f * u$, $u \in \Phi X_0$, and R_{f^α} is generated by all $f^\alpha * u'$, $u' \in \Phi X_0'$. We have

$$(f * u)^{\delta_{f^*}} = f^\alpha * \nu_f(u) \in R_{f^\alpha}, \quad u \in \Phi X_0.$$

Hence, $(R_f)^{\delta_{f^*}} \subset R_{f^\alpha}$. Analogously, $(R_{f^\alpha})^{\delta_{f^*}^{-1}} \subset R_f$. Therefore, $R_f^{\delta_{f^*}} = R_{f^\alpha}$, and the databases (F, U, V_D) and $(F', U', V_{D'})$ are similar.

Now let us prove the converse. Let (F, U, V_D) and $(F', U', V_{D'})$ be similar. Then there is a crossed isomorphism

$$(\alpha, \beta, \beta', \gamma) : (F, U, V_D) \to \left(F', U', V_{D'}\right),$$

where every γ_f is induced by some $\delta_f : D' \to D$. Let us construct a similarity for (D, Φ, F) and (D', Φ', F'). Since the maps α and δ are already given, we must construct ν and ν'.

Given $f \in F$, we have $\beta_f : u \to U'/T_f'$ and also $\beta'_{f'} : U' \to U/T_{f'}$. In order to form ν_f and $\nu'_{f'}$ we use transitions $L\Phi W \to U$ and $L\Phi' W \to U'$. If $u \in L\Phi W$, then $\bar{u} \in U$, and the same for $u' \in L\Phi' W$. Let now $u \in \Phi X_0$. Take $\bar{u}^{\beta_f} \in U'/T_f'$, and let $\bar{u}^{\beta_f} = (\bar{u}')^{\eta_f}$, where $u' \in L\Phi' W$ and η_f is a natural homomorphism. Define $\nu(u) = u'$.

We have

$$\begin{aligned}(f * u)^{\delta_{f*}} &= (f * \bar{u})^{\delta_{f*}} = f^\alpha * \bar{u}^{\beta_f} = f^\alpha * \bar{u}'^{\eta_f} \\ &= f^\alpha * \bar{u}' = f^\alpha * u' = f^\alpha * \nu_f(u).\end{aligned}$$

Using $\beta'_{f'}$ one can construct $\nu'_{f'}$ in a similar way and check the necessary second condition. □

The theorem proved allows to introduce the following new definition of DBs' equivalence, which looks exactly like the previous one, but is based on the modified notion of similarity.

DEFINITION 3.6. The concrete databases

$$(D, \Phi, F) \to (F, U, V_D) \to (F, Q, R)$$

and

$$(D', \Phi', F') \to \big(F', U', V_{D'}\big) \to (F', Q', R')$$

are called equivalent, if one of the following conditions holds:
(1) The passive parts are similar.
(2) The active parts are similar.

Note here that similarity between active DBs, standing in the middle, is equivalent to the similarity of the right parts.

3.1.5. *Galois theory and the problem of equivalence.* We use the definition of DBs equivalence given above, and consider an algorithm, based on Galois theory, which recognizes such equivalence.

We assume, that all data algebras are finite, that in all $X = (X_i,\ i \in \Gamma)$ the sets X_i are countable, and that all Halmos algebras are algebras with equalities. Under these conditions, if $R_1 \subset V_{D_1}$ and $R_2 \subset V_{D_2}$, then each isomorphism $\gamma : R_1 \to R_2$ is induced by some isomorphism $\delta : D_2 \to D_1$.

The problem is how to recognize the local isomorphism of databases

$$(F_1, Q_1, R_1) \quad \text{and} \quad (F_2, Q_2, R_2).$$

If the databases (F_1, Q_1, R_1) and (F_2, Q_2, R_2) are equivalent, i.e. they are locally isomorphic, then $|F_1| = |F_2|$ and the algebras D_1 and D_2 have to be isomorphic. In particular, if in the algebras $D_1 = (D_i^1,\ i \in \Gamma)$ and $D_2 = (D_i^2,\ i \in \Gamma)$ for some i $|D_i^1| \neq |D_i^2|$, then the DBs are not equivalent.

An intermediate step in solving the equivalence problem is the following. Let $f \in F_1$ and $f' \in F_2$ and consider the algebras $R_f \subset V_{D_1}$ and $R_{f'} \subset V_{D_2}$. We want to find out, whether these algebras are isomorphic or not. For this purpose the Galois theory of the algebra V_D is used.

First of all, given D_1 and D_2, take the groups Aut D_1 and Aut D_2. These groups are isomorphic, if D_1 and D_2 are isomorphic. Let S_{D_1} and S_{D_2} be the many-sorted symmetric groups, which are the direct products of the symmetric groups of the domains. Then $G_1 =$ Aut D_1 and $G_2 =$ Aut D_2 are subgroups in S_{D_1} and S_{D_2}, consisting of the elements which preserve operations of the set Ω. It means, that we can obtain these groups. Consider the algebra R_f, which is generated by the elements of the form $f * u$, where $u \in \Phi X_0$. Assume that the set Φ is finite, then ΦX_0 is also finite. Denote by M the finite set of all elements of the form $f * u$, $u \in \Phi X_0$. Then $M' = R'_f = H_f \subset G_1$. Here the symbol $'$ comes from the Galois correspondence (see 2.3.1).

We will describe explicitly the subgroup H_f just defined. An element $g \in G_1$ belongs to H_f if and only if $g(f * u) = f * u$ for every $u \in \Phi X_0$. Let $u = \varphi(x_1, \ldots, x_n)$ and $\tau(\varphi) = (i_1, \ldots, i_n)$. Then, the relation $f(\varphi)$ is a subset in the Cartesian product

$$D_{i_1} \times \cdots \times D_{i_n}.$$

The group G_1 acts on $D_{i_1} \times \cdots \times D_{i_n}$ by the rule

$$a^g = \left(a_1^{g_{i_1}}, \ldots, a_n^{g_{i_n}}\right),$$

where $a = (a_1, \ldots, a_n) \in D_{i_1} \times \cdots \times D_{i_n}$ and $g = (g_i,\ i \in \Gamma) \in G_1$. It is easy to check that the condition $g(f * u) = f * u$ is equivalent to that for each $a \in f(\varphi)$ the elements a^g lie in $f(\varphi)$. *Thus, for each $f * u$ the set of all corresponding g can be constructed, using tables, containing the original information.* Since the group H_f consists of elements $g \in G_1$, such that a^g lies in $f(\varphi)$ for all $a \in f(\varphi)$ and all $\varphi \in \Phi$, then we can compute H_f. The group $H_{f'} \subset G_2 = \operatorname{Aut} D_2$ is constructed in a similar way. Now, we use the fact that the algebras R_f and $R_{f'}$ are isomorphic if and only if the groups H_f and $H_{f'}$ are conjugated by an isomorphism $\delta_f : D_2 \to D_1$. In particular, H_f and $H_{f'}$ have to be isomorphic. So, if H_f and $H_{f'}$ are not isomorphic for some reason, this means that R_f and $R_{f'}$ are not isomorphic.

Since D_1 and D_2 are finite, the existence of conjugation between H_f and $H_{f'}$ can be established in a finite number of steps. Thus, we can speak about an algorithm which checks whether R_f and $R_{f'}$ are isomorphic.

Let us take the Cartesian product $F \times F'$ and consider the relation τ between elements from F and F', defined by the rule: $f \tau f'$ if and only if R_f and $R_{f'}$ are isomorphic. The relation τ is computable and a certain subset in $F \times F'$ corresponds to it.

Let now $\alpha : F \to F'$ be a bijection. We say that α is compatible with the relation τ if for each $f \in F$ there holds $f \tau f^\alpha$. For a given τ there is a simple algorithm recognizing whether there exists a bijection α compatible with τ. If such bijection exists, then the DBs are locally isomorphic. In the opposite case the DBs are not equivalent.

Let us consider an example: $F = \{f_1, \ldots, f_5\}$ and $F' = \{f_1', \ldots, f_5'\}$. Let τ be given by the graph

$$\begin{tikzcd}
& f_1 \arrow[dl] \arrow[d] \arrow[dr] & & f_2 \arrow[d] & f_3 \arrow[dr] & f_4 \arrow[d] & f_5 \arrow[dl] \\
f_1' & f_2' & f_3' & f_4' & & f_5' &
\end{tikzcd}$$

It is clear, that there is no bijection in this case.

The above algorithm can be realized more effectively in specific cases. Besides, if we fix an isomorphism between D_1 and D_2, the groups G_1 and G_2 can be identified. This also gives some simplification.

Note in conclusion that other questions related to the problem of equivalence of DBs are considered in [Br1,Br2,TPl1,TPl2,BPl3,BMSU]. The situation when the algebras D_1 and D_2 do not have nontrivial automorphisms can be easily considered separately.

3.2. *Deductive approach in databases*

3.2.1. *Preliminaries.* We proceed from a universal DB with the data algebra $D \in \Theta$. It looks like $\operatorname{Atm} D = (F_D, U, V_D)$.

A reply to a query $u \in U$ in a state $f \in F_D$ has the form $f * u$. It is an element in the algebra V_D, and simultaneously it is a subset in the set $\mathrm{Hom}(W, D)$. To evaluate a reply to a query means to point out a rule, according to which an inclusion $\mu \in f * u$ can be recognized for every $\mu : W \to D$.

Our goal is to reduce this problem of recognition to some syntactical problem, which is associated with application of deductive means. A query u is a syntactical object, while f and μ are not. Both μ and f are connected with the data algebra D, which is also a non-syntactical object. We want to describe f, μ and D syntactically.

To a state f corresponds a homomorphism $\hat{f} : U \to V_D$. Its kernel $T = \mathrm{Ker}\,\hat{f}$ is a filter in U, and this is an object of syntactical nature, which somehow describes the state.

DEFINITION 3.7. A description of the state f is an arbitrary subset T_0 of the set $T = \mathrm{Ker}\,\hat{f}$, generating the filter T.

As we already noted in the first section, there is a notion of derivability in HAs, which is associated with the derivability in first-order logic. If Ψ_0 is a set of formulas in $L\Phi W$ and T_0 is the corresponding set of elements in U, then the formula ψ is derivable from Ψ_0 if and only if the corresponding element $u = \bar{\psi} \in U$ is derivable from T_0. We know two rules of construction of a filter T over a given subset T_0. One of them is algebraic, and the second is deductive. According to the second rule, the filter T consists of all $u \in U$ which are derivable from the set T_0.

So, if T_0 is a description of a state f, then the filter $\mathrm{Ker}\,\hat{f}$ consists of elements $u \in U$, derivable form the description T_0.

A description of a state belongs to the elementary theory of this state and generates it, but it determines the state not necessarily uniquely.

3.2.2. *Categoricity and D-categoricity of a set of formulas.* It is known that a set of formulas $T \subset U$ is called a *categorical*, if any two models of the kind (D, Φ, f), $D \in \Theta$, satisfying the set T, are isomorphic. The set of symbols of relations Φ is assumed to be fixed. It is included in the scheme along with the variety Θ, and defines the algebra U.

Let D be an algebra in Θ. The set of formulas $T \subset U$ will be called *D-categorical*, if for any its model (D_1, Φ, f), $D_1 \in \Theta$, the algebra D_1 is isomorphic to D.

Given an algebra $D \in \Theta$, take the database $\mathrm{Atm}\,D = (F_D, U, V_D)$, and let T be an arbitrary set of formulas in U. Denote by $T' = F$ the subset of all $f \in F_D$ which satisfy every formula from T. Correspondingly, $T'' = F'$ is the set of all $u \in U$ which satisfy every $f \in F$. The set T'' is a closure of the set T.

THEOREM 3.8 [BPl3]. *If T is a D-categorical set of formulas, then T'' is the filter generated by the set T.*

THEOREM 3.9. *A categorical set of formulas T in U gives a description of a state f for every model (D, Φ, f) of T.*

This theorem follows directly from the previous one. Indeed, if (D, Φ, f) is a model of a categorical set T, then T is also a D-categorical set. Hence, T generates the filter T''.

But $T'' = F'$ is the intersection of all $\operatorname{Ker}\hat{f}'$, $f' \in F$. From the categoricity of the set T follows that all models (D, Φ, f') are isomorphic to the model (D, Φ, f). Therefore, all $\operatorname{Ker}\hat{f}'$ coincide with $\operatorname{Ker}\hat{f}$. So, $T'' = \operatorname{Ker}\hat{f}$, and T generates $\operatorname{Ker}\hat{f}$.

It is well known that for every finite algebra D the corresponding D-categorical sets can be constructed. Let us give a simple example of such kind, which is based on the idea of pseudoidentities. We assume that the algebras U and V_D are HAs with equalities. A pseudoidentity is a formula of the kind

$$w_1 \equiv w_1' \vee \cdots \vee w_n \equiv w_n'.$$

Let us consider this formula as an element of algebra U. Pseudoidentities generalize identities, and they are associated with pseudovarieties.

Let now $D \in \Theta$ be a finite algebra, and let Θ_D be the pseudovariety, generated by it. This pseudovariety is determined by some finite set T_0 of pseudoidentities. Let n_i be a power of every D_i from $D = (D_i,\ i \in \Gamma)$. Consider a formula

$$u = \bigwedge_i \exists x_1^i \ldots x_{n_i}^i \big(\big(x_1^i \neq x_2^i\big) \wedge \big(x_1^i \neq x_3^i\big) \wedge \cdots \wedge \big(x_{n_i-1}^i \neq x_{n_i}^i\big)\big).$$

We take for T any set of formulas, containing this formula u and a set T_0. Such set T is D-categorical. Indeed, if a model (D_1, Φ, f) satisfies the set T, then it also satisfies T_0, and the algebra D_1 belongs to pseudovariety Θ_D. It is known (see, for example, [Mal3]), that D_1 is a homomorphic image of some subalgebra in D. Since D_1 satisfies also the formula u, then the last is possible only when D_1 and D are isomorphic.

3.2.3. *Introduction of an algebra D into the language and the algebra of queries.* Our next goal is to represent D as a syntactical object. It will allow

(1) given a formula u and a row μ, to construct a new formula u_μ such that the inclusion $\mu \in f * u$ is equivalent to that of $u_\mu \in \operatorname{Ker}\hat{f}$.

(2) given a state (D, Φ, f) with a finite D, to construct a categorical set of formulas, which this state satisfies.

We apply standard methods. Let $D = (D_i,\ i \in \Gamma)$ be determined by generators and defining relations. Fix the scheme of calculus, which includes a map $n : X \to \Gamma$, a variety Θ with a set of symbols of operations Ω, and a set of symbols of relations Φ. A set of generators we denote by $M = (M_i,\ i \in \Gamma)$ and let τ be a set of defining relations. For each $a \in M$ we take a variable $y = y_a$. The set M_i corresponds to the set of variables Y_i and we have $Y = (Y_i,\ i \in \Gamma)$. Let W_D be the free algebra in Θ over the given Y. The correspondence $y_a \to a$ gives an epimorphism $\nu : W_D \to D$. The kernel $\operatorname{Ker}\nu = \rho$ is generated by τ. Thus, we have an isomorphism $W_D/\rho \to D$.

For every $a \in M_i$, $i \in \Gamma$, we add to the set Ω symbols for the nullary operations ω_a for all $i \in \Gamma$. Denote the new set by Ω'. Then Θ' is the variety of Ω'-algebras, defined by the identities of the variety Θ and by the defining relations of the algebra D. Here, if $w(y_{a_1}, \ldots, y_{a_n}) = w'(y_{a_1}, \ldots, y_{a_n})$ is one of the defining relations, then we must write it in the form

$$w(\omega_{a_1}, \ldots, \omega_{a_n}) = w'(\omega_{a_1}, \ldots, \omega_{a_n}).$$

We have no variables here, and this equality is an identity.

Let W be the free algebra in Θ over $X = (X_i,\ i \in \Gamma)$, and let W' be the free algebra in Θ' over the same X. All ω_a are elements in W' and let D' be the subalgebra in W', generated by these ω_a. It can be proved that there exists a canonical isomorphism $\nu : D' \to D$. It can be also proved that W' is a free product of W and D' in Θ, i.e. $W' = W * D'$. Here D' is a syntactical copy of the algebra D.

It is easy to understand that an algebra $D_1 \in \Theta$ can be considered as an algebra in Θ' if and only if there is a homomorphism $h : D \to D_1$. The homomorphism h takes elements from D to constants in D_1. Different h define different Ω'-algebras over the given D_1.

In the old scheme we had a Halmos algebra U, and the new scheme gives rise to an algebra U'. There is the canonical injection $U \to U'$. The kernel of this injection can be well determined in the case when D is finitely defined in Θ. The algebra D can be taken as the algebra in variety Θ' under trivial h, and we can identify the sets of homomorphisms $\mathrm{Hom}(W, D)$ and $\mathrm{Hom}(W', D)$. Simultaneously we can identify the algebras V_D and V'_D. But V_D is considered in the old scheme with the semigroup $\mathrm{End}\, W$, and V'_D in the new scheme with the semigroup $\mathrm{End}\, W'$.

We now have two databases

$$(F_D, U, V_D) \quad \text{and} \quad \bigl(F_D, U', V'_D\bigr),$$

where the second one is in the scheme with the semigroup $\mathrm{End}\, W'$.

In the second DB the algebra D is introduced into the query algebra U'. Thus, on the one hand, the query algebra now depends on D, but on the other hand, this new language has many more possibilities with respect to the given D. As for the syntactical description of D, in fact it is contained in generators and defining relations.

3.2.4. *Reply to a query from the deductive point of view.* We first make some auxiliary remarks. Take a homomorphism $\mu : W \to D$, which we identify with $\mu : W' \to D$, an isomorphism $\nu : D' \to D$ and let $s = \nu^{-1}\mu : W' \to D'$. Since D' is a subalgebra in W', we have $s \in \mathrm{End}\, W'$ and $\mu(x) = \nu\nu^{-1}\mu(x) = \nu s(x)$. For every $\tau : W' \to D$, we have $\tau s(x) = \nu s(x)$, because $s(x) \in D'$ is a constant. So, $\mu(x) = \tau s(x)$. This is true for all $x \in X$, and $\mu = \tau s$ for each τ.

Given a model (D, Φ, f), where $D \in \Theta$, an algebra D being simultaneously an algebra in Θ', we have a homomorphism $\hat{f} : U' \to V_D = V'_D$. Every $u \in U$ we consider as an element in U', and the value of u in D is $\hat{f}(u) = f * u$. We are concerned with when $\mu \in f * u$.

Together with the element u we take also an element $u_\mu = su$. This u_μ we can interpretate as the result of substitution of the row μ in the formula u.

THEOREM 3.10. *The inclusion $\mu \in f * u$ holds if and only if $f * su = 1$.*

PROOF. Let $\mu \in f * u$. Take an arbitrary τ, $\tau s = \mu$. Then $\tau s \in f * u$ and $\tau \in s(f * u) = f * su$. Since τ is arbitrary, we have $f * su = 1$.

Now let $f * su = 1$. Then $\mu \in f * su = s(f * u)$ and $\mu s = \mu \in f * u$.

So we get that the set $f * u$ is the set of all μ, for which *the formula $u_\mu = su$ is derivable from some description of the state f.* □

3.2.5. *Description of a state.* Given a model (D, Φ, f), D finite, we construct a categorical set of formulas, describing the set f. For this purpose we use Θ'-logic, associated with the algebra D, and pass to the Halmos algebra U'.

For two distinct elements a and b of an algebra $D' \subset W'$ of the same sort, take an element $a \not\equiv b$ in U'. The set of all such $a \not\equiv b$ is denoted by T_1.

It is evident that each state $f \in F_D$ satisfies the set T_1. As before we assume that $D = (D_i,\ i \in \Gamma)$ and $|D_i| = n_i$. Take $D' = (D_i',\ i \in \Gamma)$ and let $c_1, c_2, \ldots, c_{n_i}$ be a list of all elements of D_i'. For each $i \in \Gamma$ we consider the axiom

$$\forall x\big((x \equiv c_1) \vee (x \equiv c_2) \vee \cdots \vee (x = c_{n_i})\big),$$

where x is a variable of the sort i. Collect all these axioms for all $i \in \Gamma$ and denote the set of such axioms by T_2.

It is easy to prove, that if the set $T \subset U'$ includes T_1 and T_2 then T is D-categorical. Take now the diagram of the model (D, Φ, f) [Mal3]. Consider formulas of the kind $\varphi(a_1, \ldots, a_n)$ with $\varphi \in \Phi$ and constants $a_1, \ldots, a_n$ arranged accordingly to the type of φ. We identify these formulas with the respective elements in U. The diagram T_3 of the state f includes the set T_1 and contains all formulas $\varphi(a_1, \ldots, a_n)$ satisfied on the model (D, Φ, f), as well as formulas $\neg\varphi(a_1, \ldots, a_n\rceil$ with $\varphi(a_1, \ldots, a_n)$ not satisfied in the state f.

Denote by T the union of the sets T_2 and T_3. It easy to prove that the set T is categorical and so T is the description of the state f.

It is clear, that the description of the state, just constructed, is directly connected with the source basic information, because the corresponding diagram is derived from the tables of the type $f(\varphi)$, $f \in F$, $\varphi \in \Phi$.

In various specific situations also other descriptions of states can be used. Sometimes we deal with a not complete description of a state. This means, that one takes a subset T_0 in the kernel $\operatorname{Ker} \hat{f}$ which does not necessarily generate the whole elementary theory. In this case we cannot get the full reply to a query (i.e. not the full set $f * u$).

We have connected the approach to DBs, based on model theory, and the deductive approach. In general, a deductive approach does not assume existence of the corresponding model of a database [TPlK].

3.2.6. *Deductive databases.* Our goal is to construct a new model of a DB, which we call *a deductive database*. We consider an extended language U'. The data algebra D, which is introduced into the language, can be not finite.

PROPOSITION 3.11. *Two states f_1 and f_2 coincide if and only if*

$$\operatorname{Ker} \hat{f}_1 = \operatorname{Ker} \hat{f}_2.$$

PROOF. Let $\operatorname{Ker} \hat{f}_1 = \operatorname{Ker} \hat{f}_2$. We have to check that $f_1 * u = f_2 * u$ holds for every $u \in U'$. We can restrict ourselves to basic queries u, which lie in U. Take such a u and let $\mu \in f_1 * u$. Find $s \in \operatorname{End} W'$ which is assigned to μ (see Theorem 3.10). Then $su \in \operatorname{Ker} \hat{f}_1 = \operatorname{Ker} \hat{f}_2$, and hence $\mu \in f_2 * u$. We have checked the inclusion $f_1 * u \subset f_2 * u$, and similarly we can get the opposite inclusion. Therefore, $f_1 * u = f_2 * u$ holds for all basic u. This means that $f_1 = f_2$. $\square$

It follows from this proposition, that the universal database $\operatorname{Atm} D = (F_D, U', V_D)$ can be represented by means of the Halmos algebra U'. Let us denote by $\operatorname{Spec} U'$ the set of all maximal filters in U'. The kernel $\operatorname{Ker} \hat{f}$ is a maximal filter in U' for every $f \in F_D$, since the image of homomorphism $\hat{f}$ is a simple algebra. Hence, we have a map $\operatorname{Ker} : F_D \to \operatorname{Spec} U'$, which is injective according to Proposition 3.11. The rows $\mu : W' \to D$ can be identified with the corresponding $s \in \operatorname{End} W'$. Given a query $u \in U'$, the reply $f * u$ to this query is the set of all s, for which $su \in \operatorname{Ker} \hat{f} = T$. In this sense we represent replies to queries through the algebra U'. Note that not the whole algebra V_D, but only a part, i.e all possible $f * u$, $f \in F_D$, is represented in U'.

So, we describe the database by means of the single algebra U'. Indeed, queries are elements of U', states are elements of $\operatorname{Spec} U'$, and replies have just been defined. A database, constructed in such a way, we will consider as a *deductive database* (compare [Sc]). The motivation for such a name is that all the components of a DB are described syntactically, and the reply is constructed by deductive means.

Having in mind the idea of a category of these new DBs, let us make further observations. Now, we intend to introduce the *category of deductive DBs*.

As usual, we fix Θ, X, W and Φ, and let U be a HA in this scheme. Take also $D \in \Theta$ and extend the scheme with its help, passing to $W' = W^D$ and $U' = U^D$. We have $W^D = W * D'$. A homomorphism $\delta : D_1 \to D_2$ in Θ, induces a homomorphism $\bar{\delta} : W^{D_1} \to W^{D_2}$, which is the identity on W.

A homomorphism $\bar{\delta}$ naturally induces a homomorphism of algebras of formulas $L\Phi W^{D_1} \to L\Phi W^{D_2}$. Let now $s \in \operatorname{End} W^{D_1}$ be given by the map $s : X \to W^{D_1}$. A map $s^{\bar{\delta}} : X \to W^{D_2}$ is defined by the rule $s^{\bar{\delta}}(x) = s(x)^{\bar{\delta}}$. It gives a homomorphism of semigroups

$$\tilde{\delta} : \operatorname{End} W^{D_1} \to \operatorname{End} W^{D_2}.$$

Simultaneously we get a homomorphism of Halmos algebras

$$\delta' : U^{D_1} \to U^{D_2}.$$

It is a homomorphism with varying semigroup, i.e. $\delta'(su) = \tilde{\delta}(s)\delta'(u)$. The homomorphism δ' turns out to be an epimorphism if $\delta : D_1 \to D_2$ is.

Given homomorphisms $\delta : D_1 \to D_2$ and $\zeta : U_1 \to U_2$ in the scheme of the variety Θ, we have also the canonical embeddings $U_1 \to U_1^{D_1}$ and $U_2 \to U_2^{D_2}$. The homomorphism $\tilde{\delta} : \operatorname{End} W^{D_1} \to \operatorname{End} W^{D_2}$ allows us to consider an algebra $U_2^{D_2}$ as a HA in the scheme $\operatorname{End} W^{D_1}$. Canonical embeddings give the opportunity to extend ζ to $\zeta' : U_1^{D_1} \to U_2^{D_2}$, with both algebras being considered in the scheme $\operatorname{End} W^{D_1}$. The same ζ' can be considered as a homomorphism of algebras with varying semigroup, i.e. $(su)^{\zeta'} = s^{\tilde{\delta}} u^{\zeta'}$. This leads us to a commutative diagram

$$\begin{tikzcd} U_1 \arrow[r, "\zeta"] \arrow[d] & U_2 \arrow[d] \\ U_1^{D_1} \arrow[r, "\zeta'"] & U_2^{D_2} \end{tikzcd}$$

with vertical arrows defined by the corresponding canonical embeddings. The homomorphism ζ' coincides with δ' for trivial ζ.

Let now K be a *category of pairs* (D, U), where $D \in \Theta$, and U is an algebra of calculus in HA_Θ. In these U only the set of symbols of relations Φ changes, while the variety Θ, the set of sorts Γ, and the set of variables X are permanent. *Morphisms* in K are pairs (δ, ζ), where $\zeta : U_1 \to U_2$ is a homomorphism of HAs and $\delta : D_1 \to D_2$ is a homomorphism of data algebras. Denote by K' the category of HAs of the kind U^D, in which algebras are considered as algebras with varying semigroups $\mathrm{End}\, W^D$. *Morphisms* in K' are pairs (δ, ζ), with $\zeta : U_1^{D_1} \to U_2^{D_2}$ a homomorphism of the corresponding quantifier algebras, and if $u \in U_1^{D_1}$ and $s \in \mathrm{End}\, W^{D_1}$, then $(su)^\zeta = s^{\tilde{\delta}} u^\zeta$. By a quantifier algebra we mean a HA without action of the semigroup End W. As we know, objects in K' have some deductive meaning. The above diagram defines a functor $K \to K'$. It assigns to an object (D, U) in K the object U^D, and a morphism in K is transformed to (δ, ζ'), defined by the diagram. This functor induces a functor for the corresponding categories of DBs, see Theorem 3.12.

Along with a category of pairs (D, U) we can consider a category of triples (D, U, F) which is naturally correlated with the category of passive DBs in the sense that Φ defines U. This category of triples we denote by $\mathrm{DB}^0(\Theta)$. The category of all U^D with distinguished maximal filters is denoted by $\mathrm{DB}^d(\Theta)$. This category can be well defined, and it can be treated as a *category of deductive DBs*. The transition $K \to K'$ leads to the following statement.

THEOREM 3.12. *There is a natural embedding functor for categories $DB^0(\Theta) \to DB^d(\Theta)$.*

Let us give a sketch of the proof. Morphisms in the category $\mathrm{DB}^0(\Theta)$ have the form $(\delta, \zeta, \alpha) : (D_2, U_2, F_2) \to (D_1, U_1, F_1)$, where $(\delta, \zeta) : (D_1, U_1) \to (D_2, U_2)$ is a morphism in K, and $\alpha : F_2 \to F_1$ is a map, such that

$$f^\alpha * u = \left(f * u^\zeta\right)^{\delta_*}, \quad f \in F_2,\ u \in U_1.$$

This condition means also, that $\mu \in f^\alpha * u \Leftrightarrow \mu\delta \in f * u^\zeta$, where $\mu \in \mathrm{Hom}(W, D_1)$, $\mu\delta \in \mathrm{Hom}(W, D_2)$.

Let us pass to the algebras $W^{D_1} = W * D_1'$ and $W^{D_2} = W * D_2'$, where D_1' is a syntactical copy of the algebra D_1 with isomorphism $\nu_1 : D_1' \to D_1$. The same for $\nu_2 : D_2' \to D_2$. Simultaneously we pass to $U_1^{D_1}$ and $U_2^{D_2}$.

The inclusion $\mu \in f^{\alpha} * u$ is equivalent to the inclusion $s_{\mu}u \in \operatorname{Ker} f^{\alpha}$, where $s_{\mu} = \nu_1^{-1}\mu$ and the kernel is calculated in the algebra $U_1^{D_1}$. Correspondingly,

$$\mu\delta \in f * u^{\zeta} \Leftrightarrow s_{\mu\delta}u^{\zeta} \in \operatorname{Ker} f.$$

Here $s_{\mu\delta} = \nu_2^{-1}(\mu\delta)$, and $\operatorname{Ker} f$ is evaluated in the algebra $U_2^{D_2}$. So, we have

$$s_{\mu}u \in \operatorname{Ker} f^{\alpha} \Leftrightarrow s_{\mu\delta}u^{\zeta} \in \operatorname{Ker} f.$$

Let us now investigate the element $s_{\mu\delta}u^{\zeta}$. First write down a commutative diagram

$$\begin{tikzcd} D_1 \arrow[r, "\nu_1^{-1}"] \arrow[d, "\delta"'] & D_1' \arrow[d, "\delta'"] \\ D_2 \arrow[r, "\nu_2^{-1}"] & D_2' \end{tikzcd}$$

Here $\delta' = \nu_1\delta\nu_2^{-1}$. For every $x \in X$

$$\begin{aligned} s_{\mu\delta}(x) &= \nu_2^{-1}(\mu\delta)(x) = \nu_2^{-1}\big(\delta(\mu(x)\big) \\ &= \delta'\big(\nu_1^{-1}\big(\mu(x)\big)\big) = \delta'\big(s_{\mu}(x)\big) = s_{\mu}(x)^{\tilde{\delta}} = s_{\mu}^{\tilde{\delta}}(x). \end{aligned}$$

Therefore, $s_{\mu\delta} = s_{\mu}^{\tilde{\delta}}$, $s_{\mu\delta}u^{\zeta} = s_{\mu}^{\tilde{\delta}}u^{\zeta} = (s_{\mu}u)^{\zeta'}$, where ζ' is constructed from ζ according to the transition $K \to K'$. Hence, $s_{\mu}u \in \operatorname{Ker} f^{\alpha} \Leftrightarrow (s_{\mu}u)^{\zeta'} \in \operatorname{Ker} f$.

Now we come to the point where we have to give the definition of the category $\mathrm{DB}^d(\Theta)$ in more detail. *An object of this category is a pair (M, U^D), where M is a set of maximal filters in U^D*, which are kernels of homomorphisms of the kind $U^D \to V_D$. Elements of M are regarded as states of "databases", and the $u \in U^D$ are considered to be queries. In order to speak of replies to queries, let us define a set $\operatorname{End}^0 W^D$, consisting of $s \in \operatorname{End} W^D$, such that $s(x)$ is always contained in the corresponding D'. A reply to a query u in a state $T \in M$ is a set of all $s \in \operatorname{End}^0 W^D$, for which $su \in T$. Let us recall once more, that the filter T can be replaced by the description T_0 and we speak of those s, for which su is derivable from T_0. Comparing it with the standard approach in DBs, we see that everything returns to the starting point.

Morphisms in $\mathrm{DB}^d(\Theta)$ have the form

$$(\beta, \delta, \zeta) : \big(M_2, U_2^{D_2}\big) \to \big(M_1, U_1^{D_1}\big),$$

where $\beta : M_2 \to M_1$ is a map of sets and (δ, ζ) defines a homomorphism of Halmos algebras $U_1^{D_1} \to U_2^{D_2}$ with varying semigroup $\operatorname{End} W^D$. The homomorphism $\delta : D_1 \to D_2$ defines $\tilde{\delta} : \operatorname{End} W^{D_1} \to \operatorname{End} W^{D_2}$. We assume also that β, δ and ζ are linked by the following condition: for every $T \in M_2$, $u \in U_1^{D_1}$ and $s \in \operatorname{End}^0 W^{D_1}$ an inclusion $su \in T^{\beta}$ is equivalent to the inclusion $(su)^{\zeta} = s^{\tilde{\delta}} u^{\zeta} \in T$. In other words, s belongs to the reply to a query u in the state T^{β} if and only if $s^{\tilde{\delta}} \in \mathrm{End}^0 W^{D_2}$ belongs to the reply to the query u^{ζ} in the state T. This ends the precise description of the category of deductive databases.

Finally, let us consider the transition

$$\mathrm{DB}^0(\Theta) \to \mathrm{DB}^d(\Theta).$$

Let a database (D, U, F) be an object in $\mathrm{DB}^0(\Theta)$. Passing to the algebra U^D, consider for every $f \in F$ the homomorphism $\hat{f} : U^D \to V_D$. Let $T = \mathrm{Ker}\,\hat{f}$. The set of all such T for all $f \in F$ defines $M = M(F)$. Thus, we have (M, U^D).

Let now a homomorphism of passive databases

$$(\delta, \zeta, \alpha) : (D_2, U_2, F_2) \to (D_1, U_1, F_1)$$

be given. A map $\alpha : F_2 \to F_1$ naturally defines $\beta : M_2 \to M_1$, and (δ, ζ) defines a homomorphism $(\delta, \zeta') : U_1^{D_1} \to U_2^{D_2}$. Now $(\beta, \delta, \zeta') : (M_2, U_2^{D_2}) \to (M_1, U_1^{D_1})$ is a morphism in the category $\mathrm{DB}^d(\Theta)$. So, we described the functor from $\mathrm{DB}^0(\Theta)$ to $\mathrm{DB}^d(\Theta)$.

We have finished the sketch to the proof of Theorem 3.12. This theorem can be used when we constructing a model of computations in a DB (see also [CM]).

Observe finally that we had already (2.2.4) *the transition from passive databases to active ones. Now we have constructed the transition from passive databases to deductive ones.*

3.3. *Categories and databases*

3.3.1. *Preview.* The study of DB problems from the point of view of categorical semantics has developed in numerous directions. There are categorical semantics of logic in DBs, categorical aspects in the theory of computations and programming, categorical understanding of various general problems in DBs. We draw attention on two special collection of papers [Cat,Appl], devoted to these topics, and also to the papers [Vi,Ge]. Nice research on categories in databases in general, were made by E. Beniaminov [Ben5].

Recall now the ties with categories already mentioned in this survey. First of all, this is the concept of an algebraic theory, as a category of special form. An algebra is regarded as a functor from this category into some other category, and homomorphisms of algebras are the natural transformations of such functors. An algebraic theory corresponds to a variety, which can be treated as a data type. In fact, algebraic theory itself can be regarded as a data type. Categorical definition of algebraic theory gives rise to the categorical description of first-order logic, which in its turn leads to the idea of relational algebra. In 3.3.2 we will consider DB models, based on this idea. Finally, DBs in some scheme constitute a category, depending on the given scheme. We will outline connections between the categories, which appear when the scheme is changed.

3.3.2. *Databases from the point of view of relational algebras.* As was pointed out, in algebraic logic there are three approaches to the algebraization of first-order logic. In this paper we used the approach based on polyadic Halmos algebras. The second possibility is to proceed from cylindric algebras, but this approach is less convenient. Now we will apply a third approach.

Let T be an algebraic theory, which is considered as a database scheme. Recall that (see 1.3.2) a data algebra D is an object in $\operatorname{Alg} T$ and a relational algebra is an object in $\operatorname{Rel} T$.

We consider $*$-automata of the type (F, Q, R), where Q and R are relational algebras. To every state $f \in F$ corresponds a homomorphism $\hat{f} : Q \to R$, which is a natural transformation of functors. Each (F, Q, R) is a system of Boolean automata $(F, Q(\tau), R(\tau))$ with one and the same F. Here $Q(\tau)$ and $R(\tau)$ are the Boolean algebras, τ is an object of the category T. In every such automaton the operation is defined through the corresponding homomorphisms of the Boolean algebras $f(\tau) : Q(\tau) \to R(\tau)$. All these Boolean automata are compatible with the morphisms of the category T and with the corresponding operations in relational algebras.

Let us define the *universal automaton in the category of relational algebras*

$$\operatorname{Atm}(D) = (F_D, U, V_D).$$

We have to assume that to the scheme T the set of symbols of relations Φ and the set of variables X with $n : X \to \Gamma$ are added. Recall that the set of sorts Γ is used already in the definition of the category T. In fact, the set $X = (X_i,\ i \in \Gamma)$ also sits in T, if take into account the correspondence between the category T and the variety $\Theta = \operatorname{Alg}(T)$. We need that all X_i have to be countable.

An algebra D is a functor $D : T \to \mathrm{Set}$ and the relational algebra V_D is defined by

$$V_D = (\mathrm{Sub} \circ D) : T \to \mathrm{Bool}.$$

The definition of the algebra U is more difficult. The algebra U is connected with the calculus and for the construction of the relational algebra U one has to get the "relational" analogue of the calculus. But we choose another way and use Theorem 1.10 from Section 1. Given a scheme, take first the Halmos algebra U specialized in the variety Θ. By the theorem mentioned above, to the Halmos algebra U corresponds some relational algebra $T \to \mathrm{Bool}$, which is also denoted by U.

For F_D one can take the set $\operatorname{Hom}(U, V_D)$, which appears as a set of natural transformations. Note here, that the relational algebra V_D could be also constructed proceeding from the Halmos algebra V_D and in this case the whole database (F_D, U, V_D) is defined in the similar way using Theorem 1.10.

Let us pay attention to one fact, connected with the application of Theorem 1.10. We speak about Boolean algebras $U(\tau)$ and $V(\tau)$. With the object τ some finite set $Y = (Y_i,\ i \in \Gamma) \subset X$ is associated. Let $\overline{Y} = X \setminus Y$. Then $U(\tau) = \exists(\overline{Y})U$ and $V_D(\tau) = \exists(\overline{Y})V_D$, where the U and V_D on right-hand sides are HAs. The algebras $\exists(\overline{Y})U$ and $\exists(\overline{Y})V_D$ are also HAs, but in the scheme, where X is replaced by the set Y. If $W(Y)$ is the free algebra over Y in Θ, then the semigroup $\operatorname{End} W(Y)$ acts on these algebras. The algebra $\exists(\overline{Y})V_D$ can be constructed as the algebra V_D in the new scheme. The same procedure for the algebra of calculus cannot be done. However, in the algebra $\exists(\overline{Y})U$ there is a subalgebra in the new scheme, generated by all elementary formulas. It is "similar" to U in the new scheme, but it is less than the initial $\exists(\overline{Y})U$. Thus, we avoid all the difficulties which are associated with infinity of the sets X_i.

Let us return to the universal relational database $\mathrm{Atm}(D) = (F_D, U, V_D)$. It is a system of Boolean automata $(F_D, U(\tau), V_D(\tau))$.

An arbitrary relational DB is an automaton (F, Q, R), *which is considered together with a representation*

$$\rho = (\alpha, \beta, \gamma) : (F, Q, R) \to (F_D, U, V_D),$$

where $\alpha : F \to F_D$ *is a mapping,* $\beta : U \to Q$ *and* $\gamma : R \to V_D$ *are homomorphisms, i.e. natural transformations of relational algebras.*

The definition of a DB in the language of relational algebras has some advantages in comparison with the approach based on HAs. One of them is that every possible query is in fact an element of some small Boolean algebra $U(\tau)$ and the reply is calculated in the corresponding $V_D(\tau)$. In this case, notation of a query $u \in U(\tau)$ can be adjusted to manipulating in other similar Boolean algebras. Various α_* and α^* (see Section 1) are used along with Boolean operations. In practice, a query represented by more natural first-order language, should be rewritten by means of operations of relational algebra. The operations of relational algebra are easier in practical realization, and this is one more advantage of the approach based on relational algebras.

For each given theory T one can consider a category of relational DBs, depending on T, and connected with the similar category of Halmos DBs. Denote this category by $\mathrm{DB}(T)$. On the other hand, if $\psi : T \to T'$ is a morphism of theories, then it links the categories of DBs:

$$\psi^* : \mathrm{DB}(T') \to \mathrm{DB}(T).$$

We will not consider this connection in more detail. Note only the particular case, when for the corresponding Θ and Θ' there is an inclusion $\Theta' \subset \Theta$ and the morphism ψ is connected with this inclusion.

3.3.3. *Changing the set of variables.* In the scheme of a DB we will preserve the variety Θ and will change the set of variables X. Let us examine how this change influences the Halmos algebras U and V_D (see also 3.3.2). Let X be a set, Y a subset, $W = W(X)$ be the free algebra over X in Θ, and $W(Y)$ the analogous algebra over Y, $W(Y) \subset W(X)$. We have $\mathrm{Hom}(W(X), D)$, $\mathrm{Hom}(W(Y), D)$, $D \in \Theta$. To every $\mu \in \mathrm{Hom}(W(X), D)$ corresponds its restriction to $W(Y)$, denoted by $\mu_Y = \mu^\alpha$. It is clear that the map $\alpha : \mathrm{Hom}(W(X), D) \to \mathrm{Hom}(W(Y), D)$ is a surjection. Let, further, $M_D(X) = \mathrm{Sub}(\mathrm{Hom}(W(X), D))$. $M_D(Y)$ is defined similarly. A monomorphism of Boolean algebras $\alpha_* : M_D(Y) \to M_D(X)$ corresponds to the surjection α. For any X and Y these algebras are simultaneously algebras in the variety HA_Θ, but, generally, in different schemes, because X and Y are different. A direct verification shows that α_* is a homomorphism of HAs, considered in different schemes. It is also easily checked that $\Delta A = \Delta\alpha_* A$

holds for every $A \in M_D(Y)$, ΔA being the support of an element A. This implies that the homomorphism α_* induces a homomorphism

$$\alpha_*: V_D(Y) \to V_D(X).$$

Let now $\overline{Y} = X \setminus Y$. Denote by $\exists(\overline{Y})V_D(X)$ a set of all elements $\exists(\overline{Y})A$, where $A \in V_D(X)$. Then $\exists(\overline{Y})V_D(X)$ is a subalgebra in $V_D(X)$, considered in the scheme of the set Y. It is straightforward that this subalgebra is the image of the homomorphism α_*.

We have seen how changing the set of variables X influences the Halmos algebra V_D. Now let us observe what happens with the algebra U of calculus. This algebra is constructed for $X = (X_i,\ i \in \Gamma)$ with infinite X_i. If Y is a subset in X, then it is possible to pass from U to $\exists(\overline{Y})U$, which is a subalgebra of U, considered in the scheme of the set Y.

We follow the methods used in the previous subsection. Consider a set $U_0 \subset U$ of basic elements with attributes from Y. Denote by $U(Y)$ a subalgebra in U in the scheme of the set Y, generated by U_0. It is evident that it is contained in $\exists(\overline{Y})U$ and we can show that these subalgebras are different in general. It is also clear that the algebra U is a union of the various $U(Y)$ with finite Y.

3.3.4. *Generalizations.* Generalizations can be carried out by taking instead of the category Set some topos, say the topos of fuzzy sets. Then we get DBs with fuzzy information. It is possible also to consider other categories L, which take into account, in particular, modalities, and to make generalizations in the spirit of the paper of Z. Diskin [Dis1]. If in a DB there is changing of states, regulated by some semigroup, then such DB is a dynamical one (see [BPl3] for details).

3.3.5. *Computations in databases.* Computation of a reply to a query is one of the main aims of databases, and therefore DBs are often considered as one of the directions of theoretical programming. Intending to link models and computations, we use ideas of constructive algebras [BPl3], to which research of A. Malcev [Mal2], Ju. Ershov [Er], S. Goncharov [Gon], etc., are devoted. These ideas are also applicable in DBs, however up to the moment there is no stable definition of a constructive DB. We can distinguish in DBs two levels of constructivization. On one of them it is supposed that the main operations allow computable realization, while on the other level the corresponding procedures of computations are directly pointed out. The functioning of a DB assumes the second level. For example, finite DBs are obviously constructive, but they have to be provided with algorithms and programs.

In the category of (finite) DBs we can use also the general abstract ideas of computability. One such abstract idea is based on the categorical notion of comonad, which is dual to that of monad. Let us give definitions ([Man,MacL,AL] et al.).

Given a category $\mathbb{C}$, a comonad over $\mathbb{C}$ is a triple (T, ε, δ), with an endofunctor $T : \mathbb{C} \to \mathbb{C}$ and natural transformations of functors $\varepsilon : T \to I$ and $\delta : T \to T^2$, where I is the identity functor.

For every object A from $\mathbb{C}$, the *object $T(A)$ is an object of computations in A*, and the transition

$$\varepsilon_A : T(A) \to A$$

connects *computations* with their *results*. The functor T^2 *represents computations over computations*, δ connects T and T^2. These three components of a comonad have to satisfy the commutative diagrams:

$$
\begin{tikzcd}
T(A) & T^2(A) \arrow[l, "\varepsilon_{T(A)}"'] \arrow[r, "T(\varepsilon_A)"] & T(A) \\
& T(A) \arrow[ul, "\mathrm{Id}_{T(A)}"] \arrow[u, "\delta_A"'] \arrow[ur, "\mathrm{Id}_{T(A)}"'] &
\end{tikzcd}
\qquad
\begin{tikzcd}
T(A) \arrow[r, "\delta_A"] \arrow[d, "\delta_A"'] & T^2(A) \arrow[d, "\delta_{T(A)}"] \\
T^2(A) \arrow[r, "T(\delta_A)"'] & T^3(A)
\end{tikzcd}
$$

Monads (T, η, μ) are defined dually. The natural transformations in monads are $\eta : I \to T$ and $\mu : T^2 \to T$. The definition of comonad is given here in the semantics of an abstract theory of computations (for details and references see in [BG,El,RB]).

A monad in $\mathbb{C}$ allows us to speak of algebras over objects from $\mathbb{C}$. If for $\mathbb{C}$ we take the category Set or a category of Γ-sort sets Set^Γ, then we obtain another interpretation of the notion of algebra and another view on varieties of algebras (see details in [MacL]). Monads and comonads are closely related to composition of adjoint functors.

The use of the notion of comonad as an abstract idea of computations has an essential peculiarity. For every object A from $\mathbb{C}$ the corresponding object of computations $T(A)$ is also an object in $\mathbb{C}$, or else we cannot speak of T^2, T^3, etc.

Not every category $\mathbb{C}$ is well adapted to this abstract concept of computations and not every abstract idea of computations leads to specific calculations. In the last section we will consider how to construct monads and comonads in the category of databases.

As we have mentioned, our main goal is to *connect a DB model with a model of computations*. The problem is reduced, in particular, to the question how to organize an object of computations in the given DB in such a way that procedures of computations in the DB become in some sense, elements of the given object, or they become morphisms which are associated with this object. It would be nice if *computations in a DB could be represented as some other DB*.

There is another approach to computations in a DB. We can consider a category $\mathbb{C}$, which is also connected with the notion of DBs, such that there exists a canonical functor T from the category of DBs under consideration into $\mathbb{C}$. As for the category $\mathbb{C}$, we assume that computations in the form of comonad T_1 are well organized in it. The composition of functors TT_1, hopefully, constructs objects of computations in DBs.

It is not clear yet how it can be realized, and this problem is still open. Probably for the category $\mathbb{C}$ we can take the category of deductive databases $\mathrm{DB}^d(\Theta)$, considered in 3.2.6 (see Theorem 3.12). The main tool in deductive databases is logical programming. It would be beautiful to associate logical programming with some comonad in the category $\mathrm{DB}^d(\Theta)$.

3.3.6. *Computations in finite databases.* In 3.3.3 we saw that the algebra U of calculus is the union of the algebras $U(Y)$, with finite Y. Here we consider computations in finite DBs and assume that the data algebra D and the set of symbols or relations Φ are finite. Then the set of states F_D is also finite. Take a query $u \in U$ and choose a finite set Y, containing all attributes, such that $u \in U(Y)$. To this Y corresponds a finite Halmos algebra $V_D(Y)$ in the scheme of the set Y. Besides that, for every f there holds $f * u = \hat{f}(u) \in V_D(Y)$. A query u is expressed explicitly by basic queries $u_0 \in U_0$. A value $\hat{f}(u_0)$ is known for every u_0. Hence, we can compute $\hat{f}(u)$, using the fact that $V_D(Y)$ is finite.

There is another approach to the computation of a reply to a query, based on deductive methods. In order to use them, let us note that for a finite model (D, Φ, f) in an extended scheme U', a finite description can be constructed. Such a description can be reduced to one element $v \in U'$ with empty support. As we know, an inclusion $\mu \in f * u$ holds if and only if the element su is derivable from the description v. But this is equivalent to the inequality $v \leqslant su$, or, what is the same, to $v(su) = v$. It is likely that this material can be somehow connected with comonads in their applications to the category of finite DBs.

3.4. *Monads and comonads in the category of databases*

3.4.1. *Additional information on monads and comonads.* It has been already mentioned that the notions of monad and comonad are associated with the adjunction of a pair of functors. Basing us on this, we can build examples and analyze these notions in more detail.

If $\mathbb{C}$ is a category and X and Y are its objects, then the system of morphisms from X to Y is denoted by $\mathbb{C}(X, Y)$. Let now

$$F : \mathbb{C}_1 \to \mathbb{C}_2 \quad \text{and} \quad G : \mathbb{C}_2 \to \mathbb{C}_1$$

be a pair of functors. *An adjunction of F and G assumes that for every pair of objects X in $\mathbb{C}_1$ and A in $\mathbb{C}_2$ there is a bijection*

$$\varphi : \mathbb{C}_2\big(F(X), A\big) \to \mathbb{C}_1\big(X, G(A)\big).$$

This bijection should be compatible with morphisms in $\mathbb{C}_1$ and $\mathbb{C}_2$ (i.e. be functorial).

Let us consider the products

$$T_1 = GF : \mathbb{C}_1 \to \mathbb{C}_1 \quad \text{and} \quad T_2 = FG : \mathbb{C}_2 \to \mathbb{C}_2.$$

The first product leads to a monad, while the second one gives rise to a comonad. It is left to define the corresponding natural transformations. Taking for A an object of the form $F(X)$, we get $\varphi : \mathbb{C}_2(F(X), F(X)) \to \mathbb{C}_1(X, T_1(X))$ and set

$$\varphi(1_{F(X)}) = \eta_X : X \to T_1(X).$$

Then $\eta : I_{\mathbb{C}_1} \to T_1$ is a natural transformation. Take further for X an object $G(A)$. Then we have $\varphi : \mathbb{C}_2(T_2(A), A) \to \mathbb{C}_1(G(A), G(A))$. Setting $\varphi^{-1}(1_{G(A)}) = \varepsilon_A : T_2(A) \to A$, we get a natural transformation $\varepsilon : T_2 \to I_{\mathbb{C}_2}$. Using η and $X = G(A)$, we have

$$\eta_{G(A)} : G(A) \to GFG(A).$$

Application of the functor F gives

$$F(\eta_{G(A)}) : FG(A) \to FGFG(A).$$

Writing $\delta_A = F(\eta_{G(A)})$, we obtain

$$\delta_A : T_2(A) \to T_2^2(A).$$

This gives the natural transformation

$$\delta : T_2 \to T_2^2.$$

It can be proved that the triple $(T_2, \varepsilon, \delta)$ is a comonad in the category $\mathbb{C}_2$. The notation $\delta = F\eta G$ is also used. $\mu = G\varepsilon F$ is built similarly, and the triple (T_1, η, μ) is a monad in $\mathbb{C}_1$. It can also be proved that every monad and every comonad can be obtained in such a way (see [MacL]).

Let us give now a standard example of monad and comonad. Let $\mathbb{C}_1$ be a category of sets (or of many-sorted Γ-sets), $\mathbb{C}_2$ be a variety of algebras Θ with the same set of sorts Γ.

For every X from $\mathbb{C}_1$ denote by $F(X) = W(X)$ the free algebra over X in Θ. If, on the other side, A is an algebra in Θ, then denote by $G(A)$ the set, on which the algebra A is defined, i.e. G is a forgetful functor. The functors F and G are naturally adjoint.

According to the above procedure, we have a comonad (T, ε, δ) in the category $\mathbb{C}_2 = \mathbb{C} = \Theta$. For any algebra $A \in \Theta$ denote by $T(A) = W(A)$ the free Θ algebra over the set A. To the identity map $A \to A$ corresponds a homomorphism $\varepsilon_A : W(A) \to A$, and to the natural embedding $A \to W(A)$ corresponds a homomorphism $\delta_A : W(A) \to W(W(A))$. This gives $\varepsilon : I \to T$ and $\delta : T \to T^2$.

Using this comonad as a model of computations in the variety Θ, we postulate that the homomorphism $\varepsilon_A : W(A) \to A$ is always computable and hence all operations in algebras $A \in \Theta$ are computable.

A category of DBs is not a variety, but it can be associated with a variety of HAs or a variety of Boolean algebras. It might be possible to find some good adjunction of functors, leading to a model of computations. Besides that, one can associate a model of computations with an important construction of D. Scott [Sc].

One more construction which we want to discuss gives a new general view on models of computations. Let $\mathbb{C}$ be a category with a comonad (T, ε, δ). We will construct a new category $\mathbb{C}_T$ (H. Kleisli, see [Kl,MacL]), whose objects coincide with objects of the category $\mathbb{C}$. Morphisms from A to B in $\mathbb{C}_T$ are morphisms from $T(A)$ to B in $\mathbb{C}$. Composition of morphisms in $\mathbb{C}_T$ is defined in a special way, and the axioms of comonad mean that $\mathbb{C}_T$ is a category. *Morphisms* in $\mathbb{C}_T$ are treated as algorithms, realizing morphisms in $\mathbb{C}$. To make this theory closer to reality, we should force the object of computations TA, as well as morphisms $TA \to B$, to be constructive. With every category $\mathbb{C}$ a category of comonads, or in other words, a category of models of computations, can be associated (compare [RB] for the case of monads).

There is a similar construction for monads. A monad (T, η, μ) defines algebras over objects from $\mathbb{C}$. The corresponding category of algebras is denoted by $\mathbb{C}^T$. Returning to the example $F : \mathrm{Set}^{\Gamma} \to \Theta$ and $G : \Theta \to \mathrm{Set}^{\Gamma}$, note that $(T = GF, \eta, \mu)$ is a monad in Set^{Γ}, for which $(\mathrm{Set}^{\Gamma})^{T} = \Theta$, i.e. the monad reconstructs the variety Θ.

Comonad and monad are significant mathematical notions, but it is not clear yet, what the role they will play in real DBs. The rest of the section is devoted to variations on this topic.

3.4.2. *Databases in an arbitrary category* $\mathbb{C}$. From now on our aim is to give a general method of construction of monads and comonads in categories of DBs. First of all we generalize the notion of a DB. We have considered DBs as $*$-automata in an arbitrary variety of algebras $\mathcal{L}$ (5.1). Now, we replace $\mathcal{L}$ by an arbitrary category $\mathbb{C}$. Given $\mathbb{C}$, we introduce a new category $\mathrm{DB}(\mathbb{C})$, i.e. the category of DBs.[4] *Objects* of $\mathrm{DB}(\mathbb{C})$ have the form:

$$(F, Q, R),$$

where Q and R are objects in $\mathbb{C}$; Q is called an *object of queries*, R is an *object of replies*, F is a subset in $\mathrm{Hom}(Q, R) = \mathbb{C}(Q, R)$. *Morphisms* in this category have the form

$$\nu = (\nu_1, \nu_2, \nu_3) : (F', Q', R') \to (F, Q, R),$$

where $\nu_2 : Q' \to Q$ and $\nu_3 : R \to R'$ are morphisms in $\mathbb{C}$, $\nu_1 : F \to F'$ is a map. The arrow ν has the same direction as ν_2, while ν_1 and ν_3 have the opposite direction. Besides that, we assume that for each $f \in F$ the following commutative diagram holds

$$\begin{tikzcd} Q' \arrow[r, "\nu_2"] \arrow[d, "f^{\nu_1}"'] & Q \arrow[d, "f"] \\ R' & R \arrow[l, "\nu_3"] \end{tikzcd}$$

These morphisms are called the morphisms of the second kind.

If $\mathbb{C}$ is the category of HAs or locally finite HAs then $\mathrm{DB}(\mathbb{C})$ is the category of active DBs with homomorphisms of the second kind. For $\mathbb{C}$ we can take also the category of relational algebras. In this case we can speak about the category of relational DBs.

One can also start with morphisms of the first kind. They have the form:

$$\nu = (\nu_1, \nu_2, \nu_3) : (F, Q, R) \to (F', Q', R'),$$

where all arrows have the same direction:

$$\nu_1 : F \to F', \qquad \nu_2 : Q \to Q', \qquad \nu_3 : R \to R',$$

[4] If we take here $\mathbb{C} = \Theta$ then the category $\mathrm{DB}(\Theta)$ differs from the that one, considered in 3.2.

with the commutative diagram

$$\begin{tikzcd} Q \arrow[r, "\nu_2"] \arrow[d, "f"'] & Q' \arrow[d, "f^{\nu_1}"] \\ R \arrow[r, "\nu_3"] & R' \end{tikzcd}$$

There are two other possibilities for arranging of directions of the arrows. In each of them we get different categories $\mathrm{DB}(\mathbb{C})$.

3.4.3. *Comonads in $\mathbb{C}$ and in $DB(\mathbb{C})$.* Let us consider $\mathrm{DB}(\mathbb{C})$ with morphisms of the second kind. For some $\beta : Q' \to Q$ and a database (F, Q, R) let us construct a special morphism of such kind. For every $f \in F$ we take $f^\alpha = \beta f : Q' \to R$, where $f^\alpha \in \mathrm{Hom}(Q', R)$. We have a map $\alpha : F \to \mathrm{Hom}(Q', R)$. Let $F' = \mathrm{Im}\,\alpha$, $F' \subset \mathrm{Hom}(Q', R)$. Then, there is a database (F', Q', R). Consider the triple $(\alpha, \beta, 1)$, where α is considered as a map $F \to F'$. The commutative diagram holds

$$\begin{tikzcd} Q' \arrow[r, "\beta"] \arrow[d, "\beta f = f^\alpha"'] & Q \arrow[d, "f"] \\ R & R \arrow[l, "1"'] \end{tikzcd}$$

Thus, the morphism $\beta : Q' \to Q$ gives the database (F', Q', R) and defines the morphism $(\alpha, \beta, 1) : (F', Q', R) \to (F, Q, R)$. Note, that if β is an epimorphism then $\alpha : F \to F'$ is a bijection.

Suppose now that in the category $\mathbb{C}$ there is a comonad (T, ε, δ), such that for each object Q the morphism $\varepsilon_Q : T(Q) \to Q$ is an epimorphism. Now we want to construct a comonad $(\bar{T}, \bar{\varepsilon}, \bar{\delta})$ in the category $\mathrm{DB}(\mathbb{C})$. Given a database (F, Q, R), define

$$\bar{T}(F, Q, R) = \big(F^T, T(Q), R\big).$$

Here, we start with the epimorphism $\varepsilon_Q : T(Q) \to Q$ and using $\varepsilon_Q = \beta$ we get $F^T = F'$. The corresponding $\alpha : F \to F'$ is denoted by $\bar{\varepsilon}_Q$ and then there is an epimorphism

$$(\bar{\varepsilon}_Q, \varepsilon_Q, 1) : \big(F^T, T(Q), R\big) \to (F, Q, R).$$

This epimorphism is taken as $\bar{\varepsilon}_{(F,Q,R)}$. We have $\bar{\varepsilon} : \bar{T} \to I$. Now we should define T as a functor. For every

$$\nu = (\nu_1, \nu_2, \nu_3) : (F', Q', R') \to (F, Q, R),$$

one needs to define $\bar{T}(\nu) = \bar{T}(F', Q', R') \to \bar{T}(F, Q, R)$. We set $\bar{T}(\nu) = (\nu_1^T, T(\nu_2), \nu_3)$, where ν_1^T is determined by the diagram

$$\begin{tikzcd} F^T & F \arrow[l, "\bar{\varepsilon}_Q"'] \arrow[d, "\nu_1"] \\ F'^T \arrow[from=u, "\nu_1^T"'] & F' \arrow[l, "\bar{\varepsilon}_{Q'}"'] \end{tikzcd}$$

Thus, $\nu_1^T = \bar{\varepsilon}_Q^{-1} \nu_1 \bar{\varepsilon}_{Q'}$.

It can be proved that $\bar{T}$ is an endofunctor in the category DB($\mathbb{C}$) and that $\bar{\varepsilon} : \bar{T} \to I$ is, indeed, a natural transformation of functors. For the definition of $\bar{\delta} : \bar{T} \to \bar{T}^2$, we take

$$\bar{T}(F, Q, R) = \left(F^T, T(Q), R\right),$$

$$\bar{T}^2(F, Q, R) = \left(\left(F^T\right)^T, T^2(Q), R\right).$$

Let us use $\delta_Q : T(Q) \to T^2(Q)$. Every element from F^T has the form $\varepsilon_Q f$, $f \in F$, and each element from $(F^T)^T$ is of the form $\varepsilon_{T(Q)} \varepsilon_Q f$. The morphism δ_Q defines a transition $\varepsilon_{T(Q)} \varepsilon_Q f \to \delta_Q \varepsilon_{T(Q)} \varepsilon_Q f$. It follows from the definition of the comonad (T, ε, δ) that $\delta_Q \varepsilon_{T(Q)} \varepsilon_Q = \varepsilon_Q$ always holds. Hence, the morphism δ_Q defines a map $(F^T)^T \to F^T$, denoted by $\bar{\delta}_Q$. The triple $(\bar{\delta}_Q, \delta_Q, 1)$ defines a morphism $\bar{T}(F, Q, R) \to \bar{T}^2(F, Q, R)$, denoted by $\bar{\delta}_{(F,Q,R)}$. Simultaneously we have $\bar{\delta} : \bar{T} \to \bar{T}^2$. It can be proved that $\bar{\delta}$ is a natural transformation of functors.

THEOREM 3.13. *The triple $(\bar{T}, \bar{\varepsilon}, \bar{\delta})$ is a comonad in the category of databases DB($\mathbb{C}$).*

The proof of this theorem uses the following observation. Starting with

$$\bar{\delta}_{(F,Q,R)} = (\bar{\delta}_Q, \delta_Q, 1) : \left(F^T, T(Q), R\right) \to \left(\left(F^T\right)^T, T^2(Q), R\right),$$

and applying the functor $\bar{T}$, we have

$$\left(\bar{\delta}_Q^T, T(\delta_Q), 1\right) : \left(\left(F^T\right)^T, T^2(Q), R\right) \to \left(\left(\left(F^T\right)^T\right)^T, T^3(Q), R\right).$$

It follows from the definition of the functor $\bar{T}$ that the equality

$$\bar{\delta}_Q^T = \bar{\varepsilon}_{T^2(Q)}^{-1} \bar{\delta}_Q \bar{\varepsilon}_{T(Q)}$$

holds. There remains some direct checking.

3.4.4. *An example of application of the general construction.* Up to now the category $\mathbb{C}$ was an arbitrary one. Now let us take for $\mathbb{C}$ the category of locally finite HAs, denoted by $lfHA_\Theta$, and consider the category of DBs for the category $lfHA_\Theta$. In order to construct a comonad in this category of DBs, we should first construct a comonad in the category $lfHA_\Theta$.

Recall that a DB's scheme consists of a set of sorts Γ, a variety Θ and a map $n : X \to \Gamma$. We consider various sets of symbols of relations Φ. With each Φ we associate an object in a new category *RSet*. The object, which corresponds to the set of symbols of relations Φ, consists of all elementary formulas of the form $\varphi(x_1, \ldots, x_n)$, where all x_i are distinct, $\varphi \in \Phi$. We assume also that to every φ uniquely corresponds a set $\{x_1, \ldots, x_n\}$. This new category denoted by *RSet* we call the category of sets of elementary formulas. The object in *RSet*, corresponding to Φ, we denote by $\tilde{\Phi}$.

The set X is assumed to be ordered and all x_i in the formula $\varphi(x_1, \ldots, x_n)$ are disposed according to the order in X. It also defines a type τ of a symbol of relation φ. If $u = \varphi(x_1, \ldots, x_n)$, then $\Delta u = \{x_1, \ldots, x_n\}$ is the support of the formula u. Morphisms in the category of sets of elementary formulas are maps of sets which are compatible with supports. Let $\tilde{\Phi}_1$ and $\tilde{\Phi}_2$ be two sets of elementary formulas. A map $\alpha : \tilde{\Phi}_1 \to \tilde{\Phi}_2$ is compatible with supports if for every $u \in \tilde{\Phi}$ there holds $\Delta\alpha(u) \subset \Delta u$. The support of a φ can be empty.

To define an adjunction in the two categories, we introduce two functors: F and G. To an object $\tilde{\Phi}$ from the category *RSet*, corresponds a HA of calculus, denoted by $F(\tilde{\Phi}) = U(\Phi)$, in the category $lfHA_\Theta$. To morphisms of the first category correspond homomorphisms of HAs, and F can be considered as a functor.

Let now Q be a locally finite HA. With every $q \in Q$ there are associated the symbol of a relation φ_q and a formula $\varphi_q(x_1, \ldots, x_n)$ with $\{x_1, \ldots, x_n\} = \Delta q$. The corresponding set of elementary formulas is denoted by $G(Q)$. G is a functor from $lfHA_\Theta$ to *RSet*. It is easy to check, that F and G are adjoint.

Indeed, if $\tilde{\Phi}$ is an object in *RSet* and Q is an object in $lfHA_\Theta$, then we have $\mathrm{Hom}(U(\tilde{\Phi}), Q)$ and $\mathrm{Hom}(\tilde{\Phi}, G(Q))$. There is a bijection between these two sets, which agrees with morphisms in the categories.

The product FG defines a comonad (T, ε, δ) in the category $lfHA_\Theta$ of locally finite HAs in the given scheme. If Q is an object of this category, then the morphism $\varepsilon_Q : T(Q) \to Q$ is an epimorphism. This makes it possible to construct a comonad in the corresponding category of DBs.

Given a database (F, Q, R) with R a subalgebra in V_D, every $f \in F$ is a homomorphism from Q to V_D. The homomorphism $\varepsilon_Q : T(Q) \to Q$ defines a DB $(F^T, T(Q), R)$. Elements from F^T have the form $\varepsilon_Q f$ and they are homomorphisms from $T(Q) = FG(Q)$ to V_D. The algebra $FG(Q)$ is an algebra of calculus, associated with elementary formulas $\varphi_q(x_1, \ldots, x_n)$. Hence, the elements of F^T are in one-to-one correspondence with some realizations of Φ in the algebra $D \in \Theta$. Besides that, it is easy to see that the state $\varepsilon_Q f \in F^T$ induces the initial state $f \in F$.

Remark here, that the category DB($\mathbb{C}$) differs from the category of active DBs from Section 2.2. There we considered triples (F, Q, R), in which F need not be a subset of $\mathrm{Hom}(Q, R)$, unlike what we are doing now. To every $f \in F$ there was uniquely assigned a homomorphism $\hat{f}: Q \to R$. The transition from f to $\hat{f}$ did not need to be an injection. It was an injection in the case of concrete DBs. It is clear, that this property is insignificant.

The comonad just constructed defines evaluations in DBs, however these evaluations are far from real ones, and they need additional tools. We mean first of all, computation of the homomorphism $\bar{\varepsilon}_{(F,Q,R)}$, or, which is the same, of the homomorphisms ε_Q and all $f \in F$.

3.4.5. *Another construction of comonads in databases.* Now we take the category $\mathrm{DB}(\mathbb{C})$ with morphisms of the third kind. They are defined as

$$\nu = (\nu_1, \nu_2, \nu_3) : (F', Q', R') \to (F, Q, R),$$

where $\nu_1 : F \to F'$, $\nu_2 : Q' \to Q$, $\nu_3 : R' \to R$, with the commutative diagram

$$\begin{tikzcd} Q' \arrow[r, "\nu_2"] \arrow[d, "f^{\nu_1}"'] & Q \arrow[d, "f"] \\ R' \arrow[r, "\nu_2"] & R \end{tikzcd}$$

Let (T, ε, δ) be a comonad in $\mathbb{C}$. Let us construct a comonad over (T, ε, δ) in the category $\mathrm{DB}(\mathbb{C})$. As earlier, denote it by $(\bar{T}, \bar{\varepsilon}, \bar{\delta})$. Take a database (F, Q, R). Given Q and R, consider the map $T_{Q,R} : \mathrm{Hom}(Q, R) \to \mathrm{Hom}(TQ, TR)$ defined by $T_{Q,R}(f) = T(f)$ if $f \in \mathrm{Hom}(Q, R)$. Let T_0 be the restriction of the map $T_{Q,R}$ to F, $T(F) = \mathrm{Im}\, T_0$, $T(F) \subset \mathrm{Hom}(TQ, TR)$. We have $T_0 : F \to T(F)$, and also a DB

$$\bar{T}(F, Q, R) = \big(T(F), TQ, TR\big).$$

From now on we assume that the functor T has the following property. If $f_1 : Q \to R$ and $f_2 : Q \to R$ are different morphisms, then $T(f_1) : TQ \to TR$ and $T(f_2) : TQ \to TR$ are also different. This property gives a bijection $T(F) \to F$.

Let us define $\bar{T}$ as a functor. Given a morphism of the third kind

$$\nu = (\nu_1, \nu_2, \nu_3) : (F', Q', R') \to (F, Q, R),$$

we have to define

$$\bar{T}(\nu) = (\nu_1', \nu_2', \nu_3') : \big(T(F'), TQ', TR'\big) \to \big(T(F), TQ, TR\big).$$

We define ν' according to the commutative diagram

$$\begin{tikzcd} F \arrow[r, "T_0"] \arrow[d, "\nu_1"'] & T(F) \arrow[d, "\nu_1'"] \\ F' \arrow[r, "T_0'"] & T(F') \end{tikzcd}$$

where T_0 and T_0' are bijections, and $\nu_1' = T_0^{-1}\nu_1 T_0'$. Setting $\bar{T}(\nu) = (\nu_1', \nu_2' = T(\nu_2), \nu_3' = T(\nu_3))$, we get a functor $\bar{T}$.

Now, use the natural map $\varepsilon : T \to I$. For each $f : Q \to R$ there is a commutative diagram

$$\begin{array}{ccc} TQ & \xrightarrow{\varepsilon_Q} & Q \\ {\scriptstyle T(f)}\downarrow & & \downarrow{\scriptstyle f} \\ TR & \xrightarrow{\varepsilon_R} & R \end{array}$$

In particular, this is true for $f \in F$. We have the morphism of the third kind

$$(T_0, \varepsilon_Q, \varepsilon_R) : \big(T(F), TQ, TR\big) \to (F, Q, R).$$

Denote it by $\bar{\varepsilon}_{(F,Q,R)}$. We get a natural transformation $\bar{\varepsilon} : \bar{T} \to I$.

The functor T can be applied also to $T(F)$. We obtain $T^2(F)$. The bijection $T^2(F) \to T(F)$, which is valid in this case, is denoted by T_1.

$\delta : T \to T^2$ is a natural transformation by the definition of comonad, hence there is a commutative diagram

$$\begin{array}{ccc} TQ & \xrightarrow{\delta_Q} & T^2Q \\ {\scriptstyle T(f)}\downarrow & & \downarrow{\scriptstyle T^2(f)} \\ TR & \xrightarrow{\delta_R} & T^2R \end{array}$$

for every $f : Q \to R$.

If $f \in F$, then $T(f) \in T(F)$, $T^2(f) \in T^2(F)$ and $T_1(T^2(f)) = T(f)$. This diagram means also, that we have a morphism of the third kind

$$(T_1, \delta_Q, \delta_R) : \big(T(F), TQ, TR\big) \to \big(T^2(F), T^2Q, T^2R\big),$$

which could be rewritten as

$$\bar{\delta}_{(F,Q,R)} : \bar{T}(F, Q, R) \to \bar{T}^2(F, Q, R).$$

It can be checked that $\bar{\delta} : \bar{T} \to \bar{T}^2$ is a natural transformation of functors.

This leads to the following result, which is valid under the restriction mentioned on T.

THEOREM 3.14. *The triple $(\bar{T}, \bar{\varepsilon}, \bar{\delta})$ is a comonad in the category of databases $DB(\mathbb{C})$ with morphisms of the third kind.*

3.4.6. *Monads in $\mathbb{C}$ and $DB(\mathbb{C})$.* Now the category $\mathrm{DB}(\mathbb{C})$ is a category with morphisms of the first kind. Let (T, η, μ) be a monad in $\mathbb{C}$. Let us construct a monad $(\bar{T}, \bar{\eta}, \bar{\mu})$ in the category $\mathrm{DB}(\mathbb{C})$ over (T, η, μ). We assume that the functor T satisfies the same condition, as in 3.4.5. If (F, Q, R) is an object in $\mathrm{DB}(\mathbb{C})$, then we have a bijection $T_0 : F \to T(F)$. We set

$$\bar{T}(F, Q, R) = \bigl(T(F), TQ, TR\bigr).$$

Given a morphism

$$\nu = (\nu_1, \nu_2, \nu_3) : (F, Q, R) \to (F', Q', R'),$$

we have

$$\bar{T}(\nu) = \bigl(\nu_1', T(\nu_2), T(\nu_3)\bigr) : \bar{T}(F, Q, R) \to \bar{T}(F', Q', R'),$$

where ν_1' is defined, as earlier, by a commutative diagram. For every $f : Q \to R$

$$\begin{tikzcd} Q \arrow[r, "\eta_Q"] \arrow[d, "f"'] & TQ \arrow[d, "T(f)=T_0(f)"] \\ R \arrow[r, "\eta_R"] & TR \end{tikzcd}$$

This diagram means that, defining $\bar{\eta}_{(F,Q,R)} = (T_0, \eta_Q, \eta_R)$, we get a homomorphism of the first kind

$$\bar{\eta}_{(F,Q,R)} : (F, Q, R) \to \bar{T}(F, Q, R) = \bigl(T(F), TQ, TR\bigr).$$

The morphism $\bar{\eta} : I \to \bar{T}$ is a natural transformation.

We should also define $\bar{\mu} : \bar{T}^2 \to \bar{T}$. As in the previous subsection, we take $T_1 : T^2(F) \to T(F)$. Let

$$\bar{\mu}_{(F,Q,R)} = (T_1, \mu_Q, \mu_R) : \bar{T}^2(F, Q, R) \to \bar{T}(F, Q, R).$$

It is easily verified that $\bar{\mu}$ is a natural transformation.

THEOREM 3.15. *The triple $(\bar{T}, \bar{\eta}, \bar{\mu})$ is a monad in the category of databases $DB(\mathbb{C})$ with morphisms of the first kind.*

As we already know, comonads are associated with evaluations, while monads enrich objects of categories. In particular, monads in the category of DBs lead to enrichment of DB structure. Let us consider the connection between monads and algebras (see details in [MacL]), and then make some remarks on applications to DBs.

Let $\mathbb{C}$ be a category and (T, η, μ) be a monad in $\mathbb{C}$. A T-algebra in $\mathbb{C}$ is a pair (A, h), where A is an object in $\mathbb{C}$ and $h: TA \to A$ is some morphism, which is regarded as a representation. Besides that, two commutative diagrams should hold:

$$\begin{tikzcd} A \arrow[r, "\eta_A"] \arrow[dr, "\mathrm{Id}"'] & TA \arrow[d, "h"] \\ & A \end{tikzcd} \qquad \begin{tikzcd} T^2A \arrow[r, "T(h)"] \arrow[d, "\mu_A"'] & T(A) \arrow[d, "h"] \\ T(A) \arrow[r, "h"] & A \end{tikzcd}$$

A morphism $f:(A,h) \to (A',h')$ is an arrow $f: A \to A'$ in $\mathbb{C}$ with a commutative diagram

$$\begin{tikzcd} TA \arrow[r, "h"] \arrow[d, "T(f)"'] & A \arrow[d, "f"] \\ TA' \arrow[r, "h'"] & A' \end{tikzcd}$$

An adjunction, defining the source monad (T, η, μ), is connected with the category $\mathbb{C}^T$ of T-algebras and the category $\mathbb{C}$. This construction can be applied, in particular, to Example 3.4.4. The product GF from this example gives a monad (T, η, μ) in the category *RSet*.

Then we have

$$(RSet)^T = lfHA_{(\cdot)}.$$

The precise meaning of this statement is the following. The category $RSet^T$ is equivalent to a subcategory in $lfHA_{(\cdot)}$, and each locally finite algebra H from $\mathrm{HA}_{(\cdot)}$ up to an isomorphism can be realised in this subcategory.

Let us now pass to the category $\mathrm{DB}(\mathbb{C})$ with the monad $(\bar{T}, \bar{\eta}, \bar{\mu})$. Consider an "algebra" $((F, Q, R), h)$ for a DB (F, Q, R). Here

$$h = (h_1, h_2, h_3): \big(T(F), TQ, TR\big) \to (F, Q, R)$$

is a morphism of DBs, satisfying the necessary conditions. It follows from these conditions that the map $h_1: T(F) \to F$ is inverse to the bijection $T_0: F \to T(F)$. In particular, it means that h_1 is defined uniquely by the functor T and the map h_1 is one and the same for all h for the given (F, Q, R).

The algebra $((F, Q, R), h)$ is an object of the category $\mathrm{DB}(\mathbb{C})^{\bar{T}}$. We want to represent it as a DB which is an object in the category $\mathrm{DB}(\mathbb{C}^T)$. For this aim let us give several propositions which can be checked directly.

PROPOSITION 3.16. *A morphism*

$$h = \big(h_1 = T_0^{-1}, h_2, h_3\big): \big(T(F), TQ, TR\big) \to (F, Q, R)$$

defines an algebra over database (F, Q, R) if and only if the morphisms $h_2 : TQ \to Q$ and $h_3 : TR \to R$ define algebras over Q and R respectively.

Note that, given the algebra $((F, Q, R), h)$, the elements of the set F are simultaneously morphisms from $\mathrm{Hom}((Q, h_2), (R, h_3))$. Indeed, since h is a morphism of DBs, we have commutative diagrams for every $\varphi \in T(F)$:

$$\begin{tikzcd} TQ \arrow[r, "h_2"] \arrow[d, "\varphi"'] & Q \arrow[d, "\varphi^{h_1}"] \\ TR \arrow[r, "h_3"] & R \end{tikzcd}$$

Let us take $\varphi = T(f) = T_0(f)$. Then $\varphi^{h_1} = f$ and the diagram can be rewritten as

$$\begin{tikzcd} TQ \arrow[r, "h_2"] \arrow[d, "T(f)"'] & Q \arrow[d, "f"] \\ TR \arrow[r, "h_3"] & R \end{tikzcd}$$

But this means that f defines a morphism $(Q, h_2) \to (R, h_3)$. Thus, the algebra $((F, Q, R), h)$ defines a DB $(F, (Q, h_2), (R, h_3))$. It is an object of the category $\mathrm{DB}(\mathbb{C}^T)$.

PROPOSITION 3.17. *Given a database $(F, (Q, h_2), (R, h_3))$, the triple $h = (h_1 = T_0^{-1}, h_2, h_3)$ is a morphism*

$$h : \bar{T}(F, Q, R) \to (F, Q, R).$$

By the previous proposition, this morphism defines an algebra $((F, Q, R), h)$.

PROPOSITION 3.18. *A morphism*

$$\nu = (\nu_1, \nu_2, \nu_3) : (F, Q, R) \to (F', Q', R')$$

defines a morphism of algebras $((F, Q, R), h) \to ((F', Q', R'), h')$, with $h = (h_1, h_2, h_3)$, $h' = (h'_1, h'_2, h'_3)$, if and only if this ν defines a morphism of databases

$$\big(F, (Q, h_2), (R, h_3)\big) \to \big(F', (Q', h'_2), (R', h'_3)\big).$$

All this means that the following theorem holds.

THEOREM 3.19. *Transitions of the type*

$$\big((F, Q, R), h\big) \to \big(F, (Q, h_2), (R, h_3)\big)$$

define an equivalence of categories $DB(\mathbb{C})^{\bar{T}}$ and $DB(\mathbb{C}^T)$.

The categories $DB(\mathbb{C}_T)$ and $DB(\mathbb{C})_{\bar{T}}$, where T and $\bar{T}$ are the corresponding comonads are similarly related.

3.4.7. *Remarks.* Let us first consider a special monad. Let $\mathbb{C}$ be a category in which the coproduct of every two objects is defined. For a fixed object B, let us define $T(A) = A * B$ for every A, with $*$ as a sign for the coproduct. T is an endofunctor of the category $\mathbb{C}$. Given A and B, we have morphisms $i_A : A \to T(A)$ and $i_B : B \to T(A)$. Setting $\eta_A = i_A$, we get that $\eta : I \to T$ is a natural transformation.

Let us pass to $T^2(A) = (A * B) * B$. Starting with the unity morphism e_{A*B} and i_B, define $\mu_A : T^2 A \to TA$. We have here a natural transformation $\mu : T^2 \to T$. It is verified that the triple (T, η, μ) is a monad in $\mathbb{C}$.

Let now a morphism $h : A * B \to A$ define a T-algebra (A, h). Then $i_A h = e_A$. Denote $h_1 = i_B h : B \to A$. Then h_1 and e_A define the initial h.

Given an arbitrary $h_1 : B \to A$, take $e_A : A \to A$. We have $h : A * B \to A$. It is checked that h defines T-algebra (A, h), which is defined also by the morphism $h_1 : B \to A$. Hence, T-algebras in $\mathbb{C}$ are "algebras" A, considered together with some morphism $h : B \to A$.

Let, in particular, Θ be a variety of algebras, considered as the category $\mathbb{C}$, and let D be an algebra in Θ. D plays the role of the object B. In this case the category $\mathbb{C}^T$ is the variety Θ', which was considered in Section 3.2 in the problem of introducing a data algebra into the logic of queries.

Let us now turn to other questions. Note first of all, that if $\mathbb{C}$ is a category and (T, η, μ) is a monad in $\mathbb{C}$, then not always an object A from $\mathbb{C}$ allows T-enrichment. For example, a structure of Boolean algebra can not be defined on every set, and not each semigroup can be simultaneously considered as a ring. Not every HA can be regarded as an algebra with equalities. However, objects of the kind TA always allow the desired enrichment. One can proceed from $h = \mu_A : T(TA) \to TA$. An algebra (TA, μ_A) is a T-algebra, which is free over A, like semigroup ring over some semigroup. Given T-algebra (A, h), the morphism $h : TA \to A$ defines a morphism $(TA, \mu_A) \to (A, h)$. If an object A doesn't allow T-enrichment, then one can put a question about embedding of A into some B, which allows T-structure. Embedding can be understood here as embedding as a subobject or quotient-object. Transition from A to TA is such an example. These ideas can be used when we want to embed a HA into a HA with equalities. The same considerations can be applied to DBs.

Returning to the monad defined by $TA = A * B$, we can note now that an object $A * B$ is simultaneously a T-algebra, free over A. Recall in this connection that a free Θ'-algebra W', considered in Section 3.2, was represented as $W' = W * D$ with the free Θ-algebra W.

Conclusion

As we have seen, algebra and categories play a significant part in the theory of DBs. On the other hand, algebra itself is enriched by new substantial structures.

The role of algebraic logic, associated with a variety Θ, is outstanding. In this way it founds numerous application in algebra.

References

[ABCD] S. Ageshin, I. Beylin, B. Cadish and Z. Diskin, *Outline of AGO: Algebraic graph-oriented approach to specification and design of data management systems*, Frame Inform Systems, Riga, Report FIS/DBDL 9401 (1994).

[Ag] V.N. Agafonov, *Abstract data types*, Data in Programming Languages, Mir Publ., Moscow (1982) (in Russian).

[AL] A. Asperti and G. Longo, *Category Theory for the Working Computer Scientist*, The MIT Press, Cambridge, MA (1991).

[AN] H. Andreka and I. Nemeti, *Importance of universal algebra for computer science*, Universal Algebra and Links with Logic, Algebra, Combinatorics and Computer Science, Berlin (1984), 204–215.

[ANS] H. Andreka, I. Nemeti and I. Sain, *Abstract model theory approach to algebraic logic*, Preprint (1992).

[Appl] *Applications of Categories in Computer Science*, London Math. Soc. Lect. Notes Series, Vol. 177, Cambridge University Press (1991).

[Ban] F. Bancilhon, *On the completeness of query language for relational databases*, Lecture Notes in Comput. Sci., Vol. 64, Springer, Berlin (1978), 112–123.

[BB] G. Birkhoff and T.C. Bartee, *Modern Applied Algebra*, McGraw-Hill (1974).

[Ben1] E.M. Beniaminov, *The algebraic structure of relational models of databases*, NTI (Ser. 2) **9** (1980), 23–25 (in Russian).

[Ben2] E.M. Beniaminov, *Galois theory of complete relational subalgebras of algebras of relations, logical structures, symmetry*, NTI (Ser. 2) (1980) (in Russian).

[Ben3] E.M. Beniaminov, *On the role of symmetry in relational models of databases and in logical structures*, NTI (Ser. 2) **5** (1984), 17–25 (in Russian).

[Ben4] E.M. Beniaminov, *An algebraic approach to models of relational databases*, Semiotika i Informatika **14** (1979), 44–80 (in Russian).

[Ben5] E.M. Beniaminov, *Foundations of the categorical approach to knowledge representation*, Izv. Akad. Nauk SSSR (Ser. Tekhn. Kibern.) (1988), 21–33 (in Russian).

[BG] S. Brookes and S. Geva, *Computational comonads and intensional semantics*, Applications of Categories in Comp. Sci., Cambridge Univ. Press (1991), 1–44.

[BK] V.B. Borschev and M.V. Khomyakov, *On informational equivalence of databases*, NTI (Ser. 2) **7** (1979), 14–21 (in Russian).

[BL] G. Birkhoff and J. Lipson, *Heterogeneous algebras*, J. Comb. Theory **8** (1) (1970), 115–133.

[BMSU] C. Beeri, A. Mendelzon, Y. Sagiv and J. Ullman, *Equivalence of relational database schemes*, ACM (1979), 319–329.

[BN] H. Biller and E. Neuhold, *Semantics of data bases: the semantics of data models*, Information Systems **3** (1978), 11–30.

[Bo] S.N. Bojko, *Galois theory of databases*, DBMS and Environment Package. Problems of Design and Application, Riga (1985), 43–45 (in Russian).

[Bor] V.B. Borschev, *A logical approach to description of relational databases*, Semiotica i Informatica **16** (1980), 78–122 (in Russian).

[BP] W.J. Blok and D. Pigozzi, *Algebraizable logics*, Memoirs Amer. Math. Soc. **77** (396) (1989).

[BPl1] B.I. Plotkin, *Models and databases*, Trudy VCAN Gruz. SSR **21** (2) (1982), 50–78 (in Russian).

[BPl2] B.I. Plotkin, *Algebraic model of database*, Latv. Math. Ezhegodnik **27** (1983), 216–232 (in Russian).

[BPl3] B.I. Plotkin, *Universal Algebra Algebraic Logic and Databases*, Kluwer Academic Publishers (1994).

[BPl4] B.I. Plotkin, *Galois theory of databases*, Lect. Notes in Math., Vol. 1352 (1988), 147–162.

[BPl5] B.I. Plotkin, *Algebraic logic, varieties of algebras, and algebraic varieties*, Jerusalem, Preprint (1995).

[Br1] S.A. Borkin, *Data model equivalence*, Proc. of 4th Conf. on the Very Large Data Bases, IEEE, NJ (1978), 526–534.

[Br2] S.A. Borkin, *Data Models: A Semantic Approach for Database Systems*, The MIT Press (1980).

[BW] S.L. Bloom and E.G. Wagner, *Many-sorted theories and their algebras with some applications to data tipes*, Algebraic Methods in Semantics, Cambridge Univ. Press (1985), 133–168.

[Cat] *Category Theory and Computer Programming*, Lecture Notes in Comput. Sci., Vol. 240, Springer, Berlin (1986).

[Ci1] J. Cirulis, Manuscript, Riga, Latvian University (December, 1993).

[Ci2] J. Cirulis, *An algebraization of first order logic with terms*, Algebraic Logic (Proc. Conf. Budapest 1988), North-Holland, Amsterdam (1991), 125–146.

[Ci3] J. Cirulis, *Superdiagonals of universal algebras*, Acta Univ. Latviensis **576** (1992), 29–36.

[Ci4] J. Cirulis, *Abstract algebras of finitary relations: several non-traditional axiomatizations*, Acta Univ. Latviensis **595** (1994), 23–48.

[CM] U. Clocksin and K. Mellish, *Programming in PROLOG*, Springer (1984).

[Co1] E.F. Codd, *A relational model of data for large shared data banks*, Comm. ACM **13** (6) (1970), 377–387.

[Co2] E.F. Codd, *Normalized data base structure: A brief tutorial*, Proc. of 1971 ACM–SIGFIDET Workshop on Data Description. Access and Control, ACM, New York (1971), 1–17.

[Co3] E.F. Codd, *A database sublanguage founded on the relational calculus*, Proc. of 1971 ACM–SIGFIDET Workshop on Data Description. Access and Control, ACM, New York (1971), 35–68.

[Co4] E.F. Codd, *Further normalization of data base relational model*, Data Base Systems, Prentice-Hall, New York (1972), 33–64.

[Co5] E.F. Codd, *Relational database. A practical foundation for productivity*, Comm. ACM **25** (1) (1982), 109–117.

[D] Y. Kambayashi, *Database. A Bibliography*, Vols. 1, 2, Comp. Sci. Press (1981).

[Da] C.J. Date, *An Introduction to Database Systems*, Vols. 1, 2, 2nd ed., Addison-Wesley (1983).

[Dg] A. Diagneault, *On automorphisms of polyadic algebras*, Trans. Amer. Math. Soc. **112** (1964), 84–130.

[Dis1] Z. Diskin, *Poliadic algebras for the nonclassical logics*, Riga, Manuscript (1991) (unpublished).

[Dis2] Z. Diskin, *Abstract queries, schema transformations and algebraic theories; An application of categorical algebra to database theory*, Frame Inform. Systems, Riga, Report FIS/DBDL 9302 (1993).

[Dis3] Z. Diskin, *A unifying approach to algebraization of different logical systems*, Frame Inform. Systems, Riga, Report FIS/DBDL 9403 (1994).

[Dis4] Z. Diskin, *When is a semantically defined logic algebraizable*, Acta Univ. Latviensis **595** (1994), 57–82.

[DC] Z. Diskin and B. Cadish, *An application of categorical algebra to databases*, Acta Univ. Latviensis **595** (1994), 83–97.

[El] C.C. Elgot, *Monadic computation and iterative algebraic theories*, Proc. Logic Colloquim '73, North-Holland (1975).

[Er] Ju.L. Ershov, *Decision Problems and Constructive Models*, Nauka, Moscow (1980).

[Fe] N. Feldman, *Cylindric algebras with terms*, J. Symb. Logic **55** (1990), 854–866.

[Ga] J. Galler, *Cylindric and polyadic algebras*, Proc. Amer. Math. Soc. **8** (1957), 176–183.

[Ge] J. Georgescu, *A categorical approach to knowledge-based systems*, Comput. Artificial Int. (1) (1984), 105–113.

[GB1] J.A. Goguen and R.M. Burstall, *Introducing institutions*, Lecture Notes in Comput. Sci., Vol. 164, Springer, Berlin (1984), 221–256.

[GB2] J.A. Goguen and R.M. Burstall, *Institutions: Abstract model theory for specification and programming*, J. ACM **39** (1) (1992), 95–146.

[Go] J.A. Goguen, *A junction between computer science and category theory I, II*, IBM Thomas J. Watson Res. Center, Report RC 6908 (1976).

[GoTW] J.A. Goguen, J.W. Thatcher and E.G. Wagner, *An initial algebra approach to the specification, correctness and implementation of absract data types*, Current Trends in Programming Methology, Prentice-Hall (1978), 80–149.

[GoTWW] J.A. Goguen, J.W. Thatcher, E.G. Wagner and J.B. Wright, *Introduction to categories, algebraic theories and algebras*, IBM Res. Report RC 5369 (1975).

[Gol] R. Goldblatt, *Topoi – The Categorical Analysis of Logic*, North-Holland (1979).

[Gon] S.S. Goncharov, *Countable Boolean Algebras*, Nauka, Novosibirsk (1988) (in Russian).

[GPl1] G.D. Plotkin, *Call-by name, call-by value and λ-calculus*, Theoretical Computer Science **1** (1975), 125–159.

[GPl2] G.D. Plotkin, *LCF considered as programming language*, Theoretical Computer Science **5** (3) (1977), 223–256.

[GPl3] G.D. Plotkin, *A structural approach to operational semantics*, Tech. Dep. Aarchus University, Daimi FN-19 Report (1981).

[Gr] P.M. Gray, *Logic, Algebras and Databases*, Halsted Press, Chichester (1984).

[GS] C. Gunter and D. Scott, *Semantic domains*, Handbook of Theoretical Computer Science, Vol. B: Formal Models and Semantics, MIT Press/Elsevier (1990).

[Gue] I. Guessarian, *Algebraic Semantics*, Lecture Notes in Comput. Sci., Vol. 99, Springer, Berlin (1981).

[Gur1] Ju. Gurevich, *Toward logic tailored for computational complexity*, Lecture Notes in Math., Vol. 1104, Springer (1984), 175–216.

[Gur2] Ju. Gurevich, *Logic and the challenge of computer science*, Trends in Theoretical Computer Science, Computer Science Press (1989), 1–57.

[Hal] P.R. Halmos, *Algebraic Logic*, Chelsea, New York (1962).

[Hig] P.I. Higgins, *Algebras with a scheme of operators*, Math. Nachr. **27** (1–2) (1963), 115–132.

[HMT] L. Henkin, J.D. Monk and A. Tarski, *Cylindric Algebras*, North-Holland Publ. Co. (1971, 1985).

[IL] T. Imielinski and W. Lipski, *The relational model of data and cylindric algebras*, J. Comput. Syst. Sci. **28** (1984), 80–102.

[Ja] B.E. Jacobs, *On database logic*, J. ACM **29** (2) (1982), 310–332.

[Jo] P.T. Johnstone, *Topos Theory*, Academic Press (1977).

[Ka] P.C. Kanellakis, *Elements of relational database theory*, Handbook of Theoretical Computer Science, Vol. 2, Elsevier (1990), 1073–1156.

[Kl] H. Kleisli, *Every standard consrtuction is induced by a pair of adjoint functors*, Proc. Amer. Math. Soc. **16** (1965), 544–546.

[KR] A. Kock and G. Reyes, *Doctrines in cathegorical logic*, Handbook of Mathematical Logic, North-Holland (1977).

[Kr] M.I. Krasner, *Généralization et analoques de la théorie de Galois*, Comptes Rendus de congréss de la Victorie de l'Ass. Frac. pour l'Avancem. Sci. (1945), 54–58.

[LS] J. Lambek and B. Scott, *Introduction to Higher Order Categorical Logic*, Cambridge Univ. Press, Cambridge, MA (1980).

[Law1] F.W. Lawvere, *Functorial semantics of algebraic theories*, Proc. Nat. Acad. Sci. **21** (1963), 1–23.

[Law2] F.W. Lawvere, *Functorial semantics of algebraic theories*, Proc. Nat. Acad. Sci. **50** (1963), 869–872.

[Law3] F.W. Lawvere, *Some algebraic problems in the context of functorial semantics of algebraic theories*, Rep. Midwest Category Seminar II **13** (1968), 41–61.

[MacL] S. MacLane, *Categories for the Working Mathematician*, Springer (1971).

[MR] M. Makkai and G. Reyes, *First Order Categorical Logic*, Lecture Notes in Math., Vol. 611, Springer (1977).

[Mal1] A.I. Maltsev, *Model correspondences*, Izv. AN SSSR (Ser. Math.) **23** (3) (1959), 313–336 (in Russian).

[Mal2] A.I. Maltsev, *Constructive algebras*, Russian Math. Surv. **16** (3) (1961), 3–60.

[Mal3] A.I. Maltsev, *Algebraic Systems*, North-Holland (1973).

[Man] E.G. Manes, *Algebraic Theories*, Springer (1976).

[Mar] J. Martin, *Computer Data-Base Organization*, Prentice-Hall (1977).

[MP] E.S. Maftsir and B.I. Plotkin, *Automorphism group of database*, Ukrainskii Mat. Zh. **40** (3) (1988), 335–345.

[Maf] E.S. Maftsir, *Galois connections in databases*, Latv. Mat. Ezhegodnik **33** (1989), 90–100 (in Russian).

[Ma] D. Maier, *The Theory of Relational Databases*, Comp. Sci. Press (1983).

[MG] J. Meseguer and J.A. Goguen, *Initiality, induction, and computability*, Algebraic Methods in Semantics, Cambridge Univ. Press (1985), 459–541.

[MSS] A. Melton, D. Schmidt and G. Strecker, *Galois connections and computer science applications*, Lecture Notes in Comput. Sci., Vol. 240, Springer, Berlin (1986), 299–312.

[Ne] I. Nemeti, *Algebraizations of quatifier logics (overwiew)*, Studia Logica **50** (3/4) (1991), 485–569.

[Ner] A. Nerode, *Some lectures on modal logic*, Logic, Algebra and Computations, Springer (1991), 281–335.

[Ni] M. Nivat, *On the interpretation of recursive polyadic program schemes*, Symp. Mathematica, Rome **15** (1975), 255–281.

[Ol] F.J. Oles, *Type algebras, functor categories and block structure*, Algebraic Methods in Semantics, Cambridge Univ. Press (1985), 543–573.

[PGG] B. Plotkin, L. Greenglaz and A. Gvaramija, *Algebraic Structures in Automata and Databases Theory*, World Sci., Singapore (1992).

[RB] D.E. Rydeheard and R.M. Burstall, *Monads and theories: a survey for computation*, Algebraic Methods in Semantics, Cambridge Univ. Press (1985), 575–605.

[Ri] C. Ricardo, *Database Systems*, Maxwell Macmillan (1990).

[Ro] S.M. Rosenberg, *Specialized relational and Halmos algebras*, Latv. Mat. Ezhegodnik **34** (1993), 219–229 (in Russian).

[Sc] D. Scott, *Domains for denominational semantics*, Lecture Notes in Comput. Sci., Vol. 140, Springer, Berlin (1982), 577–613.

[Sy] *Database and Knowledge Base Management Systems*, Moscow, Finansi i Statistika (1981) (in Russian).

[TPl1] T. Plotkin, *Equivalent transformations of relational databases*, Latv. Mat. Ezhegodnik **29** (1985), 137–150 (in Russian).

[TPl2] T. Plotkin, *Algebraic logic in the problem of databases equivalence*, Logic Colloquium '94, Clermont-Ferrand, France (1994), 104.

[TPlK] T. Plotkin and S. Kraus, *Categoricity of a set of axioms and databases' states description*, Logic Colloquium '94, Clermont-Ferrand, France (1994), 105.

[TPPK] T. Plotkin, S. Kraus and B. Plotkin, *Problems of equivalence, categoricity of axioms and states description in databases*, Studia Logica **61** (1998), 321–340.

[Ts1] M.Sh. Tsalenko, *Relational models of databases*, Algoritmy Economicheskich Zadach (Moskow), Part I **9** (1977), 18–36, Part II **10** (1977) 16–29 (in Russian).

[Ts2] M.Sh. Tsalenko, *The main tasks of database theory*, NTI (Ser. 2) **3** (1983), 1–7 (in Russian).

[Ts3] M.Sh. Tsalenko, *Modeling Semantics in Data Bases*, Moskow, Nauka (1989) (in Russian).

[Ulm] J.D. Ullman, *Principles of Database and Knowledge-Base Systems*, Computer Science Press (1988).

[Vi] S. Vickers, *Geometric theories and databases*, Applications of Categories, Cambridge Univ. Press (1991), 288–315.

[Vo1] N.D. Volkov, *Transfer from relational algebras to Halmos algebras*, Algebra and Discrete Math., Latv. Univ., Riga (1986), 43–53.

[Vo2] N.D. Volkov, *On equivalence between categories of Halmos algebras and of relational algebras*, Latv. Mat. Ezhegodnik **34** (1993), 171–180.

[WBT] E.G. Wagner, S.L. Bloom and J.W. Thatcher, *Why algebraic theories?*, Algebraic Methods in Semantics, Cambridge Univ. Press (1985), 607–634.

[Za] A.V. Zamulin, *Data Types in Programming Languages and Databases*, Nauka, Novosibirsk (1987) (in Russian).

Section 2B
Homological Algebra. Cohomology. Cohomological Methods in Algebra. Homotopical Algebra

This Page Intentionally Left Blank

Homology for the Algebras of Analysis*

A.Ya. Helemskii

Department of Mechanics and Mathematics, Moscow State University, Moscow, 119899 GSP Russia

Contents

*This work was partially supported by the Grant 93-01-00156 from the Russian Foundation of Fundamental Investigations.

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

Preface

The area indicated by the title (another appellation is topological homology) deals with homological properties of various classes of topological, notably Banach and operator algebras. Of course, the important part of the subject concerns classes of algebras, described in general topologico-algebraic terms. But the main emphasis of the theory and, let us say, its real fun, lies in the study of homological characteristics of algebras, which are useful in diverse branches of modern analysis: operator theory, harmonic analysis, complex analysis, function spaces theory,... Hence the words "algebras of analysis".

Topological homology arose from several independent explorations. Initially, it appeared as a result of the consideration, in the Banach algebraic context, of problems concerning extensions and derivations (= crossed homomorphisms); later the problems concerning perturbations (or deformations) joined. The work on these problems was stimulated by earlier investigations of similar problems in abstract algebra. However, it began in earnest only when the inner development of functional analysis made the study of these problems unavoidable.

The pioneering work on topological homology is the paper of Kamowitz [127], who introduced cohomology groups of Banach algebras and applied them to extensions of these algebras; nevertheless, the need to study extensions in operator theory was demonstrated earlier in the Dunford theory of spectral operators [42]. The people who came to homology, starting from the study of derivations (notably Kadison and Ringrose [121,122]), certainly had in mind possible applications to automorphism groups of operator algebras, so important to quantum physics (cf. [13,21,142]). Some trends in the theory of operator algebras, also connected with the requirements of quantum physics (cf. [120]), eventually led to the investigation of perturbation problems [112,176,19].

A new source of homology, which had no predecessors in abstract algebra, was the multi-operator spectral theory, where long-standing problems in this area were solved by Taylor. It was he who showed that homology provides a proper language to define the "right" concept of a joint spectrum of several commuting operators on a Banach space [206] and used homological methods to construct a holomorphic calculus of such a system of operators on this spectrum [208,207]. These achievements required the development of the homology theory of topological algebras outside the framework of Banach structures [209]. (In the Banach context the basic homological notions were already defined and studied [79]). Afterwards homological methods in the multioperator spectral theory received a further considerable development in the work of Putinar and Eschmeier and Putinar (see, e.g., [174,48] and their joint forthcoming book).[1]

Finally, new facts came to light, which increased the interest in topological homology and accelerated its development. It was found that many basic notions in various branches of analysis and topology can be adequately expressed and sometimes better understood in homological terms. The first result of this kind was the interpretation of the topological concept of paracompactness in terms of the Banach version of projectivity [81]. But far greater resonance was caused by Johnson's discovery of the class of amenable Banach algebras [109] (which now, in retrospective view, appears to be the most important of the

[1] *Added in proof*: now it has appeared.

existing classes of "homologically best" algebras). It turned out, that the traditional, i.e. group-theoretic amenability, as well as certain important properties of operator algebras, can be treated as special cases of Johnson amenability, to the essential benefit to all areas concerned.

At the moment,[2] a considerable number of analytic and topological concepts is known to be essentially homological. Perhaps, the reader would be interested to look at a part of this rather long list. It includes:

- bounded approximate identity [102],
- the property of a locally compact group to be compact [88]; amenable [109],
- the property of a C*-algebra to be nuclear [29,75]; to have a discrete primitive spectrum and finite-dimensional irreducible representations [84,193],
- the property of a von Neumann algebra to be hyperfinite [29,91]; atomic [100],
- the property of a locally compact topological space to be discrete [89]; metrizable [190, 139]; paracompact [81],
- the approximation property of a Banach space [195],
- the property of a maximal ideal in a commutative Banach algebra to have a neighbourhood which is an analytic disk [171]; to belong to a fiber, which is such a disk (idem).

To conclude our introductory remarks, let us say a few words of an informal character about the intrinsic contents of topological homology. Its principal concepts are, as a rule, suitable topological versions of well-known notions of abstract homological algebra. The novelty is that algebras and modules bear a Banach (or, say, polynormed) structure, and relevant maps, notably module morphisms, are required to be continuous. Rather innocent in appearance, these changes make so strong an impact on the theory, that they completely transform its. Now the values of homological invariants depend on those properties of objects which appearance belong to the synthetic language of functional analysis. Moreover, new phenomena arise, which have no analogue in abstract algebra and sometimes are even contrary to what algebraists could expect. Among these, we can indicate the existence of "forbidden values" for some homological characteristics: for example, the global dimension of a commutative and semisimple Banach algebra is never equal to one, unlike the case of an abstract algebra of the same class [85]. As another illustration, well-known topological algebras, like $\mathbf{C}^M$ or the distribution algebra on a compact Lie group, can be infinite-dimensional, and nevertheless have trivial cohomology [209,208]. On the other hand, homological invariants sometimes behave better (more regularly) in topological, rather than in "pure" homology. Examples are provided by so-called "additivity formulae" for homological dimensions of tensor products of Banach algebras. The most advanced of these is, at the moment, the recently proved identity $\text{w.db}(A \mathbin{\widehat{\otimes}} B) = \text{w.db}\, A + \text{w.db}\, B$, for the weak (= flat) bidimension of unital Banach algebras (Selivanov).

It is my pleasure to thank, on these pages, Yu.O. Golovin and Yu.V. Selivanov who read the manuscript and gave a lot of valuable comments.

---

[2] Spring 1995.

Introduction

Our presentation of the fundamentals of topological homology begins with a "Part 0", containing necessary preparatory material on Banach and more general topological modules. Then the bulk of the exposition follows, which is divided in two parts. Part 1 is concerned with various concepts of what is homologically best Banach (or topological) module and algebra, and also with the closely connected topics of derivations and module extensions. From the point of view of the principal homology functors, here only one-dimensional cohomology and one-dimensional Ext-spaces are involved. At the same time, it is the most developed part of the whole theory, especially in the section concerning amenability and its variations.

The second part is devoted to cohomology and some other derived functors of arbitrary dimension, and also to closely connected numerical characteristics of modules and algebras – homological dimensions.

In case we wish to state several parallel definitions or results in a single phrase, we use the brackets $\langle\ \rangle$ and the demarcation sign $|$; thus "$\langle A \mid B \rangle$ is a $\langle$contractible $|$ amenable$\rangle$ algebra" means that A is a contractible, and B is an amenable algebra.

The reader, wishing to get a more detailed acquaintance with the area, is invited to consult the monographs [109,89], and also special chapters in the textbooks [11,96,159] (the latter book is to be published soon). The subject is also the main topic of the articles [86–88], and of the surveys [92,90,93,114,180,181]. Some special questions are reflected in the recent monograph [204].

* * *

Now we fix a number of terms and notation which we use throughout our exposition. To begin with, all the underlying linear spaces of objects (algebras, modules etc.) are always considered over the complex field **C**.

Our algebras are not, generally speaking, assumed to have an identity. (Recall that in analysis the adjoining of an identity is often a rather painful operation, and many respectable algebras (e.g., $C_0(\Omega)$ and $\mathcal{K}(E)$ below) do not have the luxury of being unital.) The abbreviation "b.a.i." always means "bounded approximate identity" (see the monograph [39] that is wholly dedicated to this extremely important concept).

The symbol $\langle \widehat{\otimes} \mid \overline{\otimes} \rangle$ denotes the complete $\langle$projective$|$inductive$\rangle$ topological tensor product (cf. [71]). The projective tensor product of Banach spaces is always considered as a Banach space with respect to the projective tensor product of given norms.

Certainly the reader knows what a *Banach algebra* is. A (general) *topological algebra* is a topological linear space equipped with a separately continuous multiplication. A $\langle \widehat{\otimes} \mid \overline{\otimes} \rangle$*-algebra* (cf. [209]) is a complete Hausdorff polynormed (= locally convex) algebra with a $\langle$jointly$|$separately$\rangle$ continuous multiplication. A *Fréchet algebra* is a metrizable $\widehat{\otimes}$- (or, which is now equivalent, $\overline{\otimes}$-) algebra. An *Arens–Michael algebra* is a locally multiplicatively-convex $\widehat{\otimes}$-algebra (cf. [89]).

For a given topological algebra A, A_+ is its *unitization algebra*, and A^{op} is its *opposite* (= reversed) *algebra*. For a $\langle \widehat{\otimes} \mid \overline{\otimes} \rangle$-algebra A, A^{env} is its enveloping $\langle \widehat{\otimes} \mid \overline{\otimes} \rangle$-algebra, that is $\langle A_+ \widehat{\otimes} A_+ \mid A_+ \overline{\otimes} A_+ \rangle$ with the multiplication, well-defined by $(a \otimes b)(c \otimes d) =$

$ac \otimes db$. For the same A, the *product map* is the continuous operator $\pi : \langle A \widehat{\otimes} A \mid A \overline{\otimes} A \rangle \to A$, well-defined by assigning ab to $a \otimes b$. If A is a Banach algebra, the *dual product map* is the adjoint operator to π, that is $\pi^* : A^* \to (A \widehat{\otimes} A)^*$. The similar operators, obtained by substituting A_+ for A, are denoted by π_+ and, respectively, π_+^*.

For a subset M in a topological algebra, its *topological square* $\overline{M^2}$ is the closure of the linear span of the set $\{ab\colon\ a, b \in M\}$.

If A is a commutative Banach algebra, its (Gel'fand) *spectrum* is the set of modular maximal ideals of A, equipped with the Gel'fand topology. It is denoted by Ω, or $\Omega(A)$ if there is a danger of a misunderstanding. The value of the Gel'fand transform of an element $a \in A$ at a point $s \in \Omega$ is denoted by $a(s)$.

If Ω is a $\langle$compact | locally compact$\rangle$ topological space, $\langle \mathrm{C}(\Omega) \mid \mathrm{C}_0(\Omega) \rangle$ is the Banach algebra (or just space) of $\langle$all | all vanishing at infinity$\rangle$ continuous functions on Ω. The symbol $\langle c_0 \mid c \mid c_b \rangle$ denotes the Banach algebra (or space) of all $\langle$converging to zero | converging | bounded$\rangle$ sequences. The so-called disk-algebra, consisting of functions, defined and continuous in the closed unit disk $\mathbf{D} \subset \mathbf{C}$ and holomorphic in its interior, is denoted by $\mathcal{A}(\mathbf{D})$. Its obvious multi-dimensional generalization, the polydisk- (or, to be precise, the n-disk-) algebra, is denoted by $\mathcal{A}(\mathbf{D}^n)$. For Banach spaces E and F, $\mathcal{B}(E, F)$ is the Banach space of all continuous (= bounded) operators from E to F, with respect to the operator norm. For the same E, $\langle \mathcal{B}(E) \mid \mathcal{K}(E) \mid \mathcal{N}(E) \rangle$ is the Banach algebra (or space) of all $\langle$ continuous | compact | nuclear $\rangle$ operators on E with respect to the $\langle$ operator | operator | nuclear $\rangle$ norm. For a locally compact group G, the group L^1-algebra of G, that is the Banach space of Haar integrable functions on G with the convolution product, is denoted by $L^1(G)$.

For an $\langle$ infinitely smooth manifold $\mathcal{M}$ | Stein manifold $\mathcal{U}$ | arbitrary set M $\rangle$, $\langle C^\infty(\mathcal{M}) \mid \mathcal{O}(\mathcal{U}) \mid \mathbf{C}^M \rangle$ is the $\langle$ Fréchet | Fréchet | $\widehat{\otimes}$- $\rangle$ algebra of $\langle$ all infinitely smooth | all holomorphic | all $\rangle$ functions on $\langle \mathcal{M} \mid \mathcal{U} \mid M \rangle$ with pointwise multiplication; recall that the topology on $\langle C^\infty(\mathcal{M}) \mid \mathcal{O}(\mathcal{U}) \mid \mathbf{C}^M \rangle$ is that of the $\langle$ uniform convergence of functions and their partial derivatives on compact subsets of $\mathcal{M}$ | uniform convergence of functions on compact subsets of $\mathcal{U}$ | pointwise convergence of functions $\rangle$.

If A is an operator algebra in a Hilbert space H (that is, subalgebra of $\mathcal{B}(H)$), then $\operatorname{Lat} A$ denotes the lattice of subspaces of H, invariant under all operators in A. Recall that A is said to be *reflexive*, if it consists of all operators, which leave all elements of Lat A invariant. The most venerable operator algebras, *von Neumann algebras* (defined as self-adjoint and closed in the weak operator topology operator algebras in H, containing the identity operator $\mathbf{1}$), can be characterized as those reflexive algebras, for which $H_0 \in \operatorname{Lat} A$ implies $H_0^\perp \in \operatorname{Lat} A$. Apart from that traditional case, we shall consider homological properties of another class of reflexive algebras, this time non-selfadjoint. (This class is more recent and not so widely known, but it is steadily gaining popularity.) These are *CSL-algebras*, defined as reflexive algebras with mutually commuting projections on the spaces in $\operatorname{Lat} A$ ("CSL" stands for "Commutative Subspace Lattice"). CSL-algebra A is called *indecomposable*, if (contrary to what was said about von Neumann algebras), $H_0 \in \operatorname{Lat} A$, $H_0 \neq H$, $\{0\}$, implies $H_0^\perp \notin \operatorname{Lat} A$. (Such algebras play for general CSL-algebras the role of "elementary bricks", similar to that of factors in the theory of von Neumann algebras.)

A lattice of subspaces in H is called a *nest* if it is linearly ordered under inclusion. A reflexive algebra A, for which $\operatorname{Lat} A$ is a nest, is called a *nest algebra* (of the nest $\operatorname{Lat} A$).

Nest algebras form the distinguished and historically first explored (by Ringrose [179]) class of indecomposable CSL-algebras; see, e.g., [37]. A reflexive algebra is said to be of finite width, if it is the intersection of a finite family of nest algebras (cf. [54]).

Throughout all the text, the category of all (not necessarily Hausdorff) topological vector spaces and continuous linear operators is denoted by **TVS**. Its full subcategory, consisting of $\langle$ Banach | complete seminormed | Fréchet $\rangle$ spaces is denoted by $\langle$ **Ban**| $\langle$**Ban**$\rangle$ | **Fr**$\rangle$. The category of linear spaces and linear operators is denoted by **Lin**.

Note that almost all categories of our exposition, those just mentioned as well as more complicated categories which will appear later, have the following common feature: their objects are linear spaces and their morphisms are linear operators. Considering a diagram in such a category, we call it *exact sequence* (= exact complex), complex, morphism of complexes or, say, short exact sequence of complexes, if it satisfies the respective definitions as a diagram in **Lin**.

A subspace E_0 in a topological linear space E is called *complementable,* if E decomposes into a topological direct sum of E_0 and some other subspace. A subspace E_0 in a Banach space E is called *weakly complementable,* if the space $\{f \in E^*\colon\ f(x) = 0 \text{ for all } x \in E_0\}$ is complementable as a subspace in E^*. (Recall that $\langle c_0 \mid \mathcal{K}(H)\rangle$, where H is a Hilbert space, is weakly complementable but not complementable in $\langle c_b \mid \mathcal{B}(H)\rangle$.)

Finally, outer multiplications in modules are always denoted by the dot "$\cdot$".

0. Preparing the stage: Banach and topological modules

0.1. *Initial definitions*

DEFINITION 0.1.1. Let A be a topological algebra. A topological linear space X, equipped with the structure of a $\langle$ left | right | two-sided $\rangle$ A-module (in the pure algebraic sense) is called a *topological* $\langle$ *left* | *right* | *two-sided* $\rangle$ *A-module,* if the outer multiplication/s is/are separately continuous. If, in addition, A is an $\overline{\otimes}$-algebra and X is a complete polynormed space, X is called a *$\overline{\otimes}$-module over A* (of a respective type). If, in addition to this, A is an $\widehat{\otimes}$-algebra, and the outer multiplication/s is/are jointly continuous, X is called a *$\widehat{\otimes}$-module over A*. Finally, if A is a $\langle$ Fréchet | Banach $\rangle$ algebra and X is a $\langle$ Fréchet | Banach $\rangle$ space, X is called a $\langle$ *Fréchet* | *Banach* $\rangle$ *A-module.*

In what follows, two-sided modules are often called *bimodules.* The more general notion of a topological polymodule over several topological algebras is not considered here; see, e.g., [89].

Topological modules do not live in solitude; the following definition expresses their social relations.

DEFINITION 0.1.2. Let X and Y be topological $\langle$left | right | two-sided$\rangle$ A-modules. A map $\varphi : X \to Y$ is called a *morphism of topological* $\langle$*left*|*right*|*two-sided*$\rangle$ A-modules if it is a morphism of respective modules in the pure algebraic sense (i.e. it is an operator, satisfying the identity/ies $\langle \varphi(a \cdot x) = a \cdot \varphi(x) \mid \varphi(x \cdot a) = \varphi(x) \cdot a \mid \text{both}\rangle$), and it is *continuous.*

Among all various notions of topological modules, that of a Banach module appears, at least at the moment, to be most important. In terms of representation theory, the notion of a Banach left A-module is equivalent to that of continuous representation of A in the underlying space of X. Indeed, we can pass from the former to the latter and vice versa, using the relation $T_a x = a \cdot x$ where $T : a \mapsto T_a \in \mathcal{B}(X)$ is the relevant representation. Under this identification, the notion of morphism of left Banach modules corresponds to that of an intertwining operator. Bearing in mind such an identification, we shall speak later about the $\langle$ module | representation $\rangle$ *associated* with a given $\langle$ representation | module $\rangle$. (Something similar could be said about more general topological modules; however, outside the class of Banach spaces, one should be careful with the words "continuous representation", cf. [96].)

Topological modules, together with their morphisms, form a lot of various categories, depending on which algebraic type and which class of underlying spaces we wish to choose. Isomorphisms in these categories are called *topological isomorphisms*; thus the role of continuity is emphasized. (In the Banach context, an *isometric isomorphism* is an isomorphism, which is also an isometric map.) Most of categories, as is typical for functional analysis, are additive but not Abelian. In the present exposition, the categories of Banach $\langle$ left | right | two-sided $\rangle$ A-modules are of a special importance; they are denoted by $\langle$ **A-mod** | **mod-A** | **A-mod-A** $\rangle$ (of course, **0-mod**, where **0** is the zero algebra, is not other thing that **Ban**). Sometimes we shall use the same symbols for some larger categories of modules; say, **A-mod** can denote the category of left $\widehat{\otimes}$-modules over a $\widehat{\otimes}$-algebra A; but in these cases we shall give a special warning.

Every topological right A-module can be considered as a topological left A^{op}-module in the obvious way. Every $\langle\, \widehat{\otimes} \mid \overline{\otimes}\, \rangle$-bimodule over a $\langle\, \widehat{\otimes} \mid \overline{\otimes}\, \rangle$-algebra A can be considered as a (unital) left $\langle\, \widehat{\otimes} \mid \overline{\otimes}\, \rangle$-module over a respective version of the enveloping algebra A^{env} (see Introduction); the outer multiplication in the left module is well defined by $(a \otimes b) \cdot x = a \cdot x \cdot b$. Obviously, these identifications lead to identifications of respective categories; in particular, **A-mod-A** is isomorphic to the full subcategory of unital modules in $\mathbf{A}^{\mathrm{env}}$**-mod**.

Combining, in the obvious way, the appropriate concepts for modules in abstract algebra and for topological linear spaces, we are able to speak, in the context of topological modules, about submodules, quotient modules, Cartesian products and direct sums. Note that a closed submodule of a Banach module, the quotient module of a Banach module with respect to a closed submodule, and the direct sum of two Banach modules are again Banach modules; norms are provided by the respective constructions in **Ban**.

REMARK. In the framework of Fréchet (in particular, Banach) algebras and modules, the outer multiplications are automatically jointly continuous whenever they are supposed to be separately continuous; see, e.g., [188]. In more general cases one must distinguish these two types of continuity, in particular, discern $\overline{\otimes}$- and $\widehat{\otimes}$-modules. As to other naturally arising continuity requirements concerning outer multiplications, the most interesting are, perhaps, hypocontinuity for general polynormed modules and so-called strict continuity for Banach modules (the latter, by definition, respects the injective tensor product "$\widehat{\widehat{\otimes}}$"). We do not consider these types of modules here; see [209] for the former and [139] for the latter.

0.2. *Examples*

We begin with several simple examples of a general character; they are just standard algebraic examples with the suitable topological modifications.

Let A be a topological algebra. Then every topological algebra B, containing A, is a topological left, as well as right and two-sided, A-module; the only required thing is that the natural embedding A into B must be continuous. Further, every $\langle$ left | right | two-sided $\rangle$ ideal in B is a topological $\langle$ left | right | two-sided $\rangle$ A-module. Of course, in both statements we can everywhere replace the word "topological" by "$\overline{\otimes}$-", "$\widehat{\otimes}$-", "Fréchet" or "Banach"; the only obvious novelty is that ideals now are supposed to be closed. Most often, the role of B is played by A itself, by A_+ or, in the Banach case, by the double centralizer algebra $\mathcal{D}(A)$ [107]. If A is a Banach operator algebra, acting on a Banach space E, we can fairly choose B as $\mathcal{B}(E)$; thus, in particular, we get the Banach A-bimodules $\mathcal{B}(E)$ and $\mathcal{K}(E)$. If, in addition, the norm in A is the operator norm, then $\mathcal{N}(E)$ is also a Banach A-bimodule.

Apart from submodules of B, one can consider its quotient modules. The case of $B = A_+$ is of a special interest.

PROPOSITION 0.2.1. *Let A be a Fréchet algebra. Then every cyclic* (in the purely algebraic sense) *Fréchet left A-module is topologically isomorphic to some quotient left A-module of A_+ (or just of A, if our algebra and module are unital).*

The proof combines the well known algebraic argument with the open mapping theorem.

As a particular case, the Fréchet A-bimodule A_+, being considered as a left, obviously cyclic and unital, A^{env}-module (see above), can be identified with the quotient A^{env}-module $A^{\mathrm{env}}/I_\Delta$. Here I_Δ is the left ideal in A^{env}, which is the kernel of the product map $\pi_+ : A_+ \widehat{\otimes} A_+ \to A_+$ (see Introduction); it is called the *diagonal* (or augmenting) ideal. As we shall show later, this module is extremely important in homology theory.

The following class of modules also deserves special attention.

DEFINITION 0.2.1. Let A be a Banach algebra, E be a Banach space. Then the Banach $\langle$ left | right | two-sided $\rangle$ A-module $\langle A_+ \widehat{\otimes} E \mid E \widehat{\otimes} A_+ \mid A_+ \widehat{\otimes} E \widehat{\otimes} A_+ \rangle$, with the outer multiplication/s, well-defined by $\langle a \cdot (b \otimes x) = ab \otimes x \mid (x \otimes a) \cdot b = x \otimes ab \mid a \cdot (b \otimes x \otimes c) = ab \otimes x \otimes c$ and $(a \otimes x \otimes b) \cdot c = a \otimes x \otimes bc \rangle$, is called a *free Banach $\langle$ left | right | two-sided $\rangle$ A-module with the basic space E.*

Free Fréchet, $\widehat{\otimes}$- and $\overline{\otimes}$-modules are defined similarly. (In the latter case one should replace the symbol "$\widehat{\otimes}$" by "$\overline{\otimes}$".)

It could be shown that this is a proper topological version of a notion of a free module in abstract algebra. In particular, our free (bi)modules have a parallel characterization in categorical terms; see, e.g., [89].

For any algebra A, the complex plane can be considered as a left A-module with respect to zero outer multiplication. Such a module is denoted by $\mathbf{C}_\infty$; obviously it is isomorphic to A_+/A.

Now we proceed to some concrete modules over algebras, serving in various areas of analysis. (In fact, analysis swarms with modules not less than algebras; one should only call things by their names.)

In the next two examples A is a commutative Banach algebra with Gel'fand spectrum Ω.

EXAMPLE 0.2.1. Let μ be a probability measure on Ω. Then the Banach space $L^p(\Omega, \mu)$, $1 \leqslant p \leqslant \infty$, is a Banach left A-module with the outer multiplication defined by $[a \cdot x](s) = a(s)x(s)$, $a \in A$, $x \in L^p(\Omega, \mu)$, $s \in \Omega$. In the simplest case of a measure, concentrated at a single point, say t, we get a one-dimensional module denoted by $\mathbf{C}_t$. Modules of that type will be called *point modules*. (Note that the map $t \mapsto \mathbf{C}_t$ is a bijection between Ω and the set of isomorphism classes of non-trivial one-dimensional A-modules.) As another simple particular case, one can consider the C[0, 1]-modules $L^p[0, 1]$, $1 \leqslant p \leqslant \infty$, with the pointwise outer multiplication.

EXAMPLE 0.2.2. Let $\underline{t} = \{t_n;\ n = 1, 2, \ldots\}$ be a sequence of points of Ω. Then the space c_{b} (see Introduction) is a Banach left A-module with respect to the operation $[a \cdot \xi]_n = a(t_n)\xi_n$, $\xi = \{\xi_n;\ n = 1, 2, \ldots\} \in c_{\mathrm{b}}$. This module is denoted by $c_{\mathrm{b}}(\underline{t})$; it is extremely useful in study of the homological dimensions of Banach algebras (see 2.5.2).

EXAMPLE 0.2.3. Let A be a Banach operator algebra acting on a Banach space E. Then E is a Banach left A-module with $a \cdot x$ defined as the action of the operator a on the vector x. This particular A-module is called *natural* or spatial.

EXAMPLE 0.2.4. Let G be a locally compact group with the Haar measure m. Then the complex plane $\mathbf{C}$ is a left $L^1(G)$-module with

$$a \cdot z = \left(\int_G a(t)\,\mathrm{d}m(t)\right) z; \quad z \in \mathbf{C}.$$

This module, denoted by $\mathbf{C}_{\mathrm{aug}}$, is called the *augmenting* $L^1(G)$*-module.* It can be identified with the quotient module $L^1(G)/I_{\mathrm{aug}}$, where

$$I_{\mathrm{aug}} = \left\{a\colon \int_G a(t)\,\mathrm{d}m(t) = 0\right\}$$

is so-called *augmenting ideal* in $L^1(G)$.

Outside the scope of Banach algebras, the following two classes of topological modules are worthy to be singled out.

EXAMPLE 0.2.5. Let $T_1, \ldots, T_n$ be an n-tuple of mutually commuting bounded operators on a Banach space E. Then E becomes a left topological module over the Fréchet algebra $\mathcal{O}(\mathbf{C}^n)$ (of entire functions on $\mathbf{C}^n$; cf. Introduction) with $w \cdot x$ well-defined as

$$\sum_{j_1, \cdots, j_n = 0}^{\infty} c_{j_1 \ldots j_n} T_1^{j_1} \ldots T_n^{j_n}(x), \quad x \in E,$$

where $\sum_{j_1,\dots,j_n=0}^{\infty} c_{j_1\dots j_n} z_1^{j_1}\dots z_n^{j_n}$ is the power series for $w = w(z_1,\dots,z_n) \in \mathcal{O}(\mathbf{C}^n)$. In such a way n-tuples of commuting bounded operators in Banach spaces can be identified with Banach left modules over $\mathcal{O}(\mathbf{C}^n)$, and vice versa. These modules play an outstanding role in Taylor's theory of the multioperator holomorphic calculus (cf. Preface).

EXAMPLE 0.2.6. Let $\mathcal{M}$ be an infinitely smooth manifold, $C^{\infty}(\mathcal{M})$ the Fréchet algebra of infinitely smooth functions on $\mathcal{M}$ (cf. Introduction), and $\mathcal{U}$ an open set in $\mathcal{M}$. Then the Fréchet space $C^{\infty}(\mathcal{U})$ (with the similarly defined topology) is a Fréchet left module over $C^{\infty}(\mathcal{M})$ with respect to the pointwise outer multiplication.

0.3. *Dual and normal modules*

We return to general definitions. It was observed (mainly by Kadison and Ringrose [121] and Johnson [109]) that, inside the class of all Banach modules, there exists a certain subclass with some additional topological properties. These properties make the work with such modules much easier, than with general modules.

DEFINITION 0.3.1. A Banach $\langle$ left | right | two-sided $\rangle$ module X over a Banach algebra A is called a *dual module* (of the respective type), if its underlying space is, up to a topological isomorphism, dual to some another Banach space, and for every $a \in A$ the operator/s $\langle x \mapsto a\cdot x \mid x \mapsto x\cdot a \mid \text{both}\rangle$ is/are continuous in the respective weak* topology of X.

There is a characterization of dual modules, presenting them in a more constructive manner. Let X_* be a Banach $\langle$ left | right | two-sided $\rangle$ A-module. It is easy to see that the dual Banach space $(X_*)^*$ is a Banach $\langle$ left | right | two-sided $\rangle$ A-module with the operation/s defined by $\langle \langle x\cdot a, x_*\rangle = \langle x, a\cdot x_*\rangle \mid \langle a\cdot x, x_*\rangle = \langle x, x_*\cdot a\rangle \mid \text{both} \rangle$, $a \in A, x_* \in X_*$, $x \in (X_*)^*$. This new module (of the respective type) is called the *dual to* (the module) X_*, whereas the latter is called, in this context, the *predual to the module* $(X_*)^*$.

PROPOSITION 0.3.1. *A Banach A-module* (of an arbitrary type) *is dual iff, up to an isometric isomorphism, it has a predual Banach A-module.*

The proof uses the fact that a weak* continuous operator between dual Banach spaces is always an adjoint to some other operator.

The apparent advantage of dual modules is that, by virtue of the Banach–Alaoglu theorem, their unit balls are weakly* compact, so one can use various tricks based on this property (see, e.g., [109,180]).

Let us present several illustrations. First of all, every reflexive, as a Banach space, module is, of course, dual. The $C[0,1]$-module $L^p[0,1]$, $1 \leqslant p \leqslant \infty$, is dual iff $p > 1$. The modules $c_b(\underline{t})$ from Example 0.2.2 are dual, whatever $\underline{t}$ is chosen.

Finally, take an operator algebra A, acting on a Hilbert space H, and the Banach A-bimodules $\mathcal{B}(H)$, $\mathcal{K}(H)$ and $\mathcal{N}(H)$ with the outer multiplications provided by the operator composition. Then, as Banach spaces, $\mathcal{B}(H) = (\mathcal{N}(H))^*$ and $\mathcal{N}(H) = (\mathcal{K}(H))^*$ ("Schatten–von Neumann duality"), and the same is true in the context of Banach

A-bimodules. Thus $\mathcal{B}(H)$ and $\mathcal{N}(H)$ are dual bimodules, whereas $\mathcal{K}(H)$, lacking even a predual Banach space, is not.

Now let us concentrate on the just considered case of an operator algebra acting on a Hilbert space. As it turns out (cf. [121,98]), in this case the class of dual A-modules contains a substantial proper subclass with still better properties.

The already mentioned duality between $\mathcal{N}(H)$ and $\mathcal{B}(H)$ enables us to equip A with the new topology, inherited from the weak* topology of $\mathcal{B}(H)$. This is the so-called *ultraweak* (or Dixmier) topology.

In what follows, an operator from A to some dual Banach space $E = (E_*)^*$ is called *normal*, if it is continuous with respect to the ultraweak topology of A and the weak* topology of E; in particular, a normal functional on A is an ultraweakly continuous functional.

Let A be as above, and $X = (X_*)^*$ be a dual Banach $\langle$ left | right $\rangle$ A-module. Then the subset of X consisting of those x, for which the operator $A \to X : a \mapsto \langle a \cdot x \mid x \cdot a \rangle$ is normal, form a closed submodule X, called the $\langle$ left | right $\rangle$ *normal part* of X and denoted by $\langle X_l \mid X_r \rangle$. If X is a dual Banach A-bimodule, X_l and X_r are both subbimodules of X; the closed subbimodule $X_l \cap X_r$ is called the (just) *normal part of* X, and it is denoted by X_w.

DEFINITION 0.3.2 (cf. [121]). With A as above, a dual Banach $\langle$ left | right | two-sided $\rangle$ A-module X is called *normal* if it coincides with its $\langle$ left normal | right normal | normal $\rangle$ part.

This concept is essentially important, when A is an operator C*-algebra (= norm closed *-subalgebra of $\mathcal{B}(H)$), especially when it is a von Neumann algebra (= ultraweakly closed *-subalgebra of $\mathcal{B}(H)$ with $\mathbf{1}$).

As a simple example, the natural (= spatial) left A-module H is normal, whatever operator algebra A, acting on H, we take. Further, if A is an operator C*-algebra, then every von Neumann algebra B such that $A \subseteq B \subseteq \mathcal{B}(H)$ is a normal A-bimodule. Indeed, the well known Sakai's characterisation of von Neumann algebras (see, e.g., [165]) easily implies that B is a dual A-bimodule: its predual is the closed subbimodule B_* in B^*, consisting of all normal functionals. The required continuity conditions of operators $a \mapsto a \cdot x \,/\, x \cdot a$, $x \in B$, follow from the separate ultraweak continuity of operator composition.

On the contrary, for an infinite-dimensional operator C*-algebra A, the A-bimodules A^* and A^{**}, being dual, are not normal. In particular, the normal part of A^* is its proper subbimodule, coinciding with the predual to A^-, the ultraweak closure of A in $\mathcal{B}(H)$. This bimodule, which will play a rather significant role in the sequel, is denoted by A^-_*.

0.4. *Constructions, involving morphism spaces and tensor products*

WARNING. Unless explicitly stated otherwise, we consider in this subsection only Banach algebras and modules.

Different categories of modules are connected by various functors, which enable us to pass from one kind of (bi)modules to another – or, in the simplest cases, to Banach spaces

(= **0**-modules). Two families of functors are most important: morphism functors and tensor product functors.

The simplest representative of the first family is just a specialization, to a category of Banach modules, of the abstract categorical morphism functor. Take, say, **A-mod**. Then, for any pair X, Y of its objects, the set ${}_Ah(X, Y)$ of morphisms from X to Y is a closed subspace in the Banach space $\mathcal{B}(X, Y)$, hence a Banach space itself. Further, if X is as before, and $\varphi: Y_1 \to Y_2$ is a morphism in **A-mod**, then the map

$$ {}_Ah(X, \varphi): {}_Ah(X, Y_1) \to {}_Ah(X, Y_2), \quad \psi \mapsto \varphi \circ \psi, $$

is a bounded operator. Thus, fixing X, we get a *covariant morphism functor*

$$ {}_Ah(X, ?): \mathbf{A\text{-}mod} \to \mathbf{Ban}: \; Y \mapsto {}_Ah(X, Y), \quad \varphi \mapsto {}_Ah(X, \varphi). $$

Similarly, after the fixation of the second module Y in ${}_Ah(X, Y)$ and the proper definition of the operator ${}_Ah(\psi, Y)$ for a morphism ψ, the relevant categorical construction leads to the *contravariant morphism functor* ${}_Ah(?, Y): \mathbf{A\text{-}mod} \to \mathbf{Ban}$.

Analogous functors with values in **Ban** can be introduced by replacing **A-mod** by **mod-A** and **A-mod-A**. In the respective notation, we shall use the symbols h_A for right modules and ${}_Ah_A$ for bimodules.

The second principal functor is based on the Banach version of the algebraic concept of the tensor product of modules. Let A be a Banach algebra; take this time modules of different types: $X \in \mathbf{mod\text{-}A}$ and $Y \in \mathbf{A\text{-}mod}$. Consider the projective tensor product $X \widehat{\otimes} Y$ of the underlying Banach spaces of these modules and denote by L the closure of the linear span of all possible elements of the form $x \cdot a \otimes y - x \otimes a \cdot y$, $a \in A$, $x \in X$, $y \in Y$. Denote by $X \widehat{\otimes}_A Y$ the quotient space $X \widehat{\otimes} Y/L$.

DEFINITION 0.4.1. The Banach space $X \widehat{\otimes}_A Y$ is called the (projective) *tensor product of the Banach modules* X and Y (over A).

REMARK. As its pure algebraic prototype, the tensor product of the Banach modules could be also defined in terms of a suitable universal property; see, e.g., [89].

If $\varphi: X_1 \to X_2$ and $\psi: Y_1 \to Y_2$ are morphisms respectively in **mod-A** and **A-mod**, the bounded operator

$$ \varphi \mathbin{\widehat{\otimes}_A} \psi: X_1 \widehat{\otimes}_A X_2 \to Y_1 \widehat{\otimes}_A Y_2, $$

well-defined by

$$ (\varphi \mathbin{\widehat{\otimes}_A} \psi)(x \mathbin{\widehat{\otimes}_A} y) = \varphi(x) \mathbin{\widehat{\otimes}_A} \psi(y) $$

is called the *tensor product of these morphisms.* (Here symbols like $x \mathbin{\widehat{\otimes}_A} y$ denote, of course, cosets of the respective elementary tensors.) This construction enables us to introduce a new pair of functors, this time consisting only of covariant functors. Namely, fixing X, we get a functor

$$X\widehat{\otimes}_A ?: \mathbf{A\text{-}mod} \to \mathbf{Ban}, \quad Y \mapsto X \widehat{\otimes}_A Y, \ \psi \mapsto \mathbf{1}_X \widehat{\otimes}_A \psi,$$

whereas fixing Y, we get

$$? \widehat{\otimes}_A Y: \mathbf{mod\text{-}A} \to \mathbf{Ban}, \quad X \mapsto X \widehat{\otimes}_A Y, \ \varphi \mapsto \varphi \widehat{\otimes}_A \mathbf{1}_Y.$$

Both are called *tensor product functors.*

REMARK. It is easy to see, that each of the functors ${}_Ah(A_+, ?)$ and $A_+\widehat{\otimes}_A ?$ is, up to a natural equivalence of functors, no other thing than the forgetful functor from **A-mod** to **Ban**.

The functors, introduced above, make the given structure poorer, transforming modules to mere Banach spaces. There are other varieties of these functors, which preserve a given type of modules, change it or enrich an initial structure. All of these modifications could be obtained as special cases of some general construction, using the notion of a bimodule over two different algebras. The reader can find the relevant details, say, in [89]; here we just mention the most frequent particular cases. Only the association between objects will be indicated; its extension to morphisms will be always obvious from the context.

EXAMPLE 0.4.1. If $X \in \langle$ **A-mod** | **Ban** | **A-mod** $\rangle$ and $Y \in \langle$ **Ban** | **A-mod** | **A-mod** $\rangle$, then the space $\mathcal{B}(X, Y)$ is a Banach $\langle$ right | left | two-sided $\rangle$ A-module with the operation/s defined as $\langle [\varphi \cdot a](x) = \varphi(a \cdot x) \mid [a \cdot \varphi](x) = a \cdot \varphi(x) \mid \text{both}\rangle$, where $a \in A$, $\varphi \in \mathcal{B}(X, Y)$, $x \in X$. So, fixing, say, the second object Y, we get a contravariant functor from $\langle$ **A-mod** | **Ban** | **A-mod** $\rangle$ to $\langle$ **mod-A** | **A-mod** | **A-mod-A** $\rangle$. Note that the construction of dual module, so important in the previous subsection, is just a particular case of the latter construction, namely, when $Y = \mathbf{C} \in$ **Ban**. In this case, the respective functor is the so-called *star functor* $(^*)$: **A-mod**$\to$ **mod-A**.

EXAMPLE 0.4.2. If $X \in \langle$ **A-mod** | **Ban** | **A-mod** $\rangle$ and $Y \in \langle$ **Ban** | **mod-A** | **mod-A** $\rangle$, then the tensor product of spaces (not of modules!) $X \widehat{\otimes} Y$ is a Banach $\langle$ left | right | two-sided $\rangle$ A-module with the operation/s, well defined by $\langle a \cdot (x \otimes y) = (a \cdot x) \otimes y \mid (x \otimes y) \cdot a = x \otimes (y \cdot a) \mid \text{both}\rangle$. Accordingly, fixing, say, X, we get a covariant functor from $\langle$ **Ban** | **mod-A** | **mod-A** $\rangle$ to $\langle$ **A-mod** | **mod-A** | **A-mod-A** $\rangle$. In particular, taking A_+ as X and **Ban** as the domain category, we obtain the construction of free Banach left A-modules, already discussed earlier (see Definition 0.2.1).

EXAMPLE 0.4.3. If $X \in$ **A-mod-A** and $Y \in$ **A-mod**, then the space $X \widehat{\otimes}_A Y$, where X is considered as a right module, is a Banach left A-module with the operation, well-defined by

$$a \cdot (x \widehat{\otimes}_A y) = (a \cdot x) \widehat{\otimes}_A y.$$

Thus, fixing X, we get a covariant functor from **A-mod** to the same category.

All hitherto considered functors are certainly additive.

Tensor multiplication by cyclic modules and, under certain conditions, by ideals has a rather transparent description.

Let M be a subset in A_+, and let $X \in$ **A-mod**. Denote by $\overline{M \cdot X}$ the closure in X of the linear span of the set $\{a \cdot x \colon a \in M,\ x \in X\}$. Obviously, $\overline{I \cdot X}$ is a closed submodule in X provided I is a left ideal in A_+. The closed submodule $\overline{A \cdot X}$ in X is called the *non-degenerate part of* X, and X itself is called *non-degenerate* if it coincides with its non-degenerate part (or, equivalently, if its associated representation is non-degenerate in the sense of representation theory). Similarly, non-degenerate parts of right modules and non-degenerate right modules are defined. Of course, all these notions can be obviously extended to general topological modules over topological algebras.

PROPOSITION 0.4.1. *Let I be a closed $\langle$right | two-sided$\rangle$ ideal in A_+, and let $X \in$* **A-mod**. *Then there exists an isomorphism*

$$\alpha : A_+/I \mathbin{\widehat{\otimes}_A} X \to X/\overline{I \cdot X}$$

in $\langle$ **Ban** | **A-mod** *$\rangle$, well-defined by*

$$(a + I) \mathbin{\widehat{\otimes}_A} x \mapsto a \cdot x + \overline{I \cdot X}.$$

PROPOSITION 0.4.2 (cf. [178] and [200]). *Let I be a closed $\langle$right | two-sided$\rangle$ ideal in A_+. Then the following conditions are equivalent:*

(i) *I has a left bounded approximate identity,*

(ii) *for any $X \in$* **A-mod** *there exists an isomorphism $\sigma : I \mathbin{\widehat{\otimes}_A} X \to \overline{I \cdot X}$ in $\langle$***Ban** | **A-mod***$\rangle$, well-defined by $a \mathbin{\widehat{\otimes}_A} x \mapsto a \cdot x$.*

The tensor product functors are connected with the morphism functors by some important identities, which are versions of so-called "adjoint associativity" or "freezing isomorphism". The simplest of these versions can be presented by the formula

$$\mathcal{B}\big(X \mathbin{\widehat{\otimes}_A} Y, E\big) \simeq h_A\big(X, \mathcal{B}(Y, E)\big)$$

which should be understood as a natural equivalence of the indicated functors of three variables. As to more general forms of such a formula see, e.g., [89].

REMARK. Outside the framework of Banach structures, the question of what is an appropriate topology on morphism spaces (to begin with, on ${}_Ah(X, Y)$ for Fréchet modules X and Y) is rather delicate. There are a number of situations when it still has no satisfactory answer. (See [209] for details of the relevant discussion.) Therefore, for categories of non-Banach modules, it makes sense to consider versions of morphism functors with values in categories of pure algebraic modules (or, in simplest cases, of linear spaces). On the contrary, the constructions, based on tensor products, can be well extended, with suitable modifications, to various categories of polynormed modules. In particular, Definition 0.4.1, Examples 0.4.2 and 0.4.3, and Proposition 0.4.1 could be repeated, word by word, for categories of Fréchet modules. Details can be found in [89].

Tensor products of Banach modules play a major role in the study of topological versions of such an important algebraic concept as Morita-equivalence; see Grønbaeck [67,68].

1. Homologically best modules and algebras ("one-dimensional theory")

1.1. *The three pillars: Projectivity, injectivity, flatness*

Topological homology starts with the definition of "homologically best" Banach and more general topological modules. Their prototypes were introduced by the founding fathers of homological algebra [16,149]. Cartan, Eilenberg and MacLane presented us with three distinguished classes of modules, which together form a foundation of all algebraic homology theory. Here we give the proper topological versions of these concepts.

As a preparatory notion, a morphism φ in an arbitrary category of topological modules is called *admissible* if its kernel and image both are complementable as subspaces, that is have topological linear complements in the respective spaces, and, moreover, if φ is an open map onto its image.[3] Thus, in reasonable categories (like, say, **A-mod**), where monomorphisms are injective maps and epimorphisms are maps with dense images, a $\langle$ *mono* | *epi* $\rangle$ *morphism is admissible iff it has a* $\langle$ *left* | *right* $\rangle$ *inverse map, which is a continuous linear operator.*

In the context of Banach modules, a morphism is called preadmissible if its dual (= adjoint) morphism is admissible. A sequence of $\langle$ topological | Banach $\rangle$ modules is called $\langle$ admissible | preadmissible $\rangle$ if its every morphism is $\langle$ admissible | preadmissible $\rangle$.

1.1.1. *Projective modules.* To avoid tiresome repetitions, we concentrate in the following definition on the (most important) case of Banach structures. We imply that, replacing "Banach" by "Fréchet", "$\widehat{\otimes}$-", "$\overline{\otimes}$-" or just "topological", we get the respective definitions for more general classes of modules.

DEFINITION 1.1.1 [79]. Let A be a Banach algebra. A Banach $\langle$ left | right | two-sided$\rangle$ A-module P is called *projective* if, for any Banach A-module X of the same type, every admissible epimorphism $\tau : X \to P$ is a retraction in $\langle$ **A-mod** | **mod-A** | **A-mod-A**$\rangle$ (that is, it has a right inverse morphism of Banach $\langle$ left | right | two-sided $\rangle$ A-modules).

REMARK. The presence of the admissibility condition signifies that the topological homology is a kind of a relative homology (cf. [104,47]). Without this requirement we would get too poor class of objects to develop a rich theory. Sometimes, however, it makes sense to permit the "test" epimorphisms to be just surjective. The resulting "strictly projective" modules, as well as similar versions of injective and flat modules of the next subsection, are also of some use, mostly in questions concerning cyclic modules [88].

[3] Note that the latter condition is satisfied automatically in categories of Fréchet modules.

For a time, we shall consider only Banach modules over a fixed Banach algebra A. Note, however, that the majority of results of a general character, like Propositions 1.1.1 and 1.1.2 below, have natural analogues for Fréchet, $\widehat{\otimes}$- and $\overline{\otimes}$-modules.

A retract (or, equivalently, a direct summand) of a projective module of any type is itself projective. Every free module (see Definition 0.2.1), notably A_+ as a left and right, and $A_+ \widehat{\otimes} A_+$ as a two-sided A-module, is projective. (If A has an identity, one can substitute it here for A_+.) If A has no right identity, then the left A-module $\mathbf{C}_\infty$, that is A_+/A (see 0.2), provides a simplest example of a non-projective module.

Every object in **Ban**, i.e. in **0-mod**, is certainly projective.

REMARK. The question about projectivity of A_+ as of an A-*bi*module, is much more complicated and will be discussed later, in Subsection 1.3.1. Now, to forestall events, we notice that it happens only in exceptional cases, and this is just what makes the subject of topological homology worthy to be studied.

There is a number of approaches to the notion of the projectivity. In particular, in order to check the property, it is sufficient to consider a single important morphism. Namely, for a given Banach left A-module (hence unital A_+-module) X we consider the free Banach left module with X as its basic space and introduce the morphism $\pi_X : A_+ \widehat{\otimes} X \to X$, well-defined by assigning $a \cdot x$ to $a \otimes x$. This morphism, which is obviously admissible, is called the *canonical projection* (for X).

PROPOSITION 1.1.1. *A Banach left A-module P is projective iff the canonical projection π_P is a retraction in* **A-mod**.

A similar characterization, in terms of proper versions of the notion of a canonical projection, is valid for right and two-sided modules.

As a corollary, we have the following intrinsic description of projective modules.

PROPOSITION 1.1.2. *A Banach A-module* (of any type) *is projective iff, in the respective category, it is a retract of a free Banach A-module.*

Every projective bimodule is projective as a left and as a right module. If P and Q are projective objects respectively in **A-mod** and **mod-A**, and $E \in$ **Ban**, then $\langle P \widehat{\otimes} E \mid E \widehat{\otimes} Q \mid P \widehat{\otimes} Q \rangle$ is projective in $\langle$ **A-mod** $\mid$ **mod-A** $\mid$ **A-mod-A** $\rangle$.

One of most typical problems of topological homology is as follows: which modules, belonging to this or that class, are projective? At first let us consider ideals, proper and non-proper alike.

Obviously, if a left ideal in A has a right identity, it is projective. There exists, however, a much less burdensome sufficient condition, which is stated in terms of the presence, in these ideals, of a right bounded approximate identity with some good special properties. This condition has a somewhat technical character, and we shall not formulate it here (see [96], Theorem VII.1.23). But it is worth notice, that the class of ideals with this property includes maximal ideals of uniform algebras, corresponding to peak points of their spectra [80] and left closed ideals in separable C*-algebras [145]. The "if" part of the following theorem is also a special case of that general assertion.

Let I be a closed ideal in a commutative Banach algebra A. Recall that the Gel'fand spectrum of I, as of a commutative Banach algebra, can be identified with the subset $\{s\colon \text{there exists } a \in I \text{ with } a(s) \neq 0\}$ of $\underline{\Omega}(A)$.

THEOREM 1.1.1 [81]. *With A and I as above, I is a projective Banach left A-module only if, and in the case of $A = C_0(\underline{\Omega}(A))$ also if its spectrum is paracompact. In particular, for a locally compact space Ω, the Banach algebra $C_0(\Omega)$ is a projective Banach left $C_0(\Omega)$-module iff Ω is paracompact.*

Thus $C_0(\Omega) \in \mathbf{C_0(\Omega)\text{-mod}}$ is always projective in the case of a metrizable Ω. On the other hand, of the $2^{\aleph}$ maximal ideals in $c_b \simeq C(\beta\mathbf{N})$, only those from the countable set of ideals identified with points of $\mathbf{N}$, are projective.

In all hitherto considered cases of ideals, the idea of the proof of the projectivity can be clarified in the "embrionic" case of the c_0-module c_0. For a given $a \in c_0$, consider a formal series

$$\sum_{n=1}^{\infty} ap^n \otimes p^n$$

in $c \widehat{\otimes} c_0$, where p^n are the "orts" $\{\delta_m^n\colon m = 1, 2, \ldots\}$. A known estimation of norms of so-called diagonal elements in projective tensor products (see, e.g., [89], II,2.6) provides that this series converges to some $\rho(a) \in c \widehat{\otimes} c_0$. Sending a to $\rho(a)$, we obtain a map, which turns out to be a morphism in $\mathbf{c_0\text{-mod}}$, right inverse to the canonical projection for c_0.

Apart from good topological properties of the spectrum, provided by the previous theorem, projective ideals in commutative Banach algebras possess a number of special analytic and geometric properties. Let us concentrate on the most transparent case of a maximal ideal I in a unital algebra A of that class with the spectrum Ω. In what follows, s is a point in Ω, representing I, $\partial\Omega$ is the Shilov boundary of Ω, and $\overline{I^2}$ is the topological square of I (cf. Introduction).

THEOREM 1.1.2 ([85,171] and [173] combined). *Let I be projective. Then*

(i) *If $s \in \partial\Omega$, then $\overline{I^2} = I$. Besides, there is a constant $C > 0$ such that, for any $t \in \Omega$ there exists $a \in I$ (i.e. $a \in A$ with $a(s) = 0$) such that $a(t) = 1$ and $\|a\| < C$.*

(ii) *If $s \notin \partial\Omega$, then* $\dim I/\overline{I^2} = 1$. *Besides, there exists a neighbourhood $\mathcal{U}$ of s in* (the Gel'fand topology of) *Ω which is an analytic disc.*

(Recall, that the latter means that there is a homeomorphism ω of $\mathbf{D}$ onto $\mathcal{U}$, where $\mathbf{D}$ is the open unit disc in $\mathbf{C}$, such that for every $a \in A$ the function $\omega\colon \mathbf{D} \to \mathbf{C}\colon z \mapsto a(\omega(z))$ is holomorphic.)

REMARK. The inequality $\dim I/\overline{I^2} \leqslant 1$ (but not the subsequent analytic assertion) can be extended from projective to flat modules which will soon appear (see Theorem 1.1.6 below).

As to ideals of group algebras, the following can be mentioned.

PROPOSITION 1.1.3. *For an arbitrary locally compact group G, $L^1(G)$ is a projective Banach left $L^1(G)$-module. If G is compact, then every closed left ideal in $L^1(G)$ is projective.*

The proof of the first, more elementary, assertion is based on the Grothendieck isomorphism $L^1(G) \widehat{\otimes} L^1(G) \simeq L^1(G \times G)$. The latter enables us to construct a right inverse to $\pi_{L^1(G)}$ as the map $\rho : L^1(G) \to L^1(G \times G)$ defined by $[\rho(a)](s,t) = \chi(t^{-1})a(st)$, where χ is the characteristic function of an arbitrary chosen compact set of the Haar measure 1 in G.

At the same time, for a very wide class of non-compact groups (and, perhaps, for all such groups) the augmenting ideal (see Example 0.2.4) in $L^1(G)$ is not projective. Some details will be given later, in 2.5.5; there it will be shown that this ideal is in fact homologically much worse than projective modules.

From ideals we turn to cyclic and more general finitely generated modules. Let X be a cyclic module A_+/I (cf. 0.2). If there exists another closed ideal, say J, in A_+ with $A_+ = I \oplus J$ (or, equivalently, I has a right identity), then X is obviously projective. *Whether the converse is always true?* In abstract algebra such a question, because of its triviality, would sound strange, but in the context of Banach modules it is, at the moment, an open problem. Of course, if we know beforehand, that I is, as a subspace, complementable, the very definition of the projectivity immediately implies the affirmative answer, but what if not? Fortunately, the result is still valid, if we assume another condition, this time in terms of some "good geometry" of X itself, which is, as a whole, much milder and can be effectively checked. This is the famous approximation property, discovered, in its numerous guises, by Grothendieck [71]. The following result can be considered as a suitable topological version of the known algebraic "dual basis lemma" (cf. [50]).

THEOREM 1.1.3 (Selivanov [198]). *Let X be a finitely generated* (in the algebraic sense) *Banach left A-module, possessing, as a Banach space, the approximation property. Then it is projective iff there exists a finite set of elements $y_i \in X$, $1 \leqslant i \leqslant k$, and a set of morphisms $\varphi_i \in_A h(X, A_+)$, $1 \leqslant i \leqslant k$, such that every $x \in X$ has the form*

$$x = \sum_{i=1}^{k} \varphi_i(x) \cdot y_i.$$

As a particular case, if $X = A_+/I$ is a one-generated (=cyclic) *Banach left A-module with the approximation property, then it is projective iff $A_+ = I \oplus J$ for some closed left ideal J in A_+.*

Now let E be the natural module over an operator Banach algebra A (see Example 0.2.3). It will be convenient to say that A is *spatially projective* if E is a projective Banach module.

Which operator algebras are spatially projective? A rather simple sufficient condition is as follows. Recall that a *column of rank one operators* on E is a subset in $\mathcal{B}(E)$, consisting, for a fixed $f \in E^*$, $f \neq 0$, of all $x \otimes f$; $x \in E$; the latter notation, as it is usual, is taken for the operator $y \mapsto \langle f, y \rangle x$.

PROPOSITION 1.1.4. *If an operator algebra contains a column of rank one operators, then it is spatially projective.*

This condition is certainly satisfied for the standard triple of algebras $\mathcal{B}(E)$, $\mathcal{K}(E)$ and $\mathcal{N}(E)$, and also for many other operator algebras, selfadjoint and non-selfadjoint alike, e.g., for all algebras which are cones, in the sense of Gilfeather and Smith [56], of other algebras. For a certain important class of operator algebras in Hilbert space this sufficient condition turns out to be a criterion (cf. Theorem 1.1.8 below).

In the proof (the idea originates in [125]), one chooses $x_0 \in E$ with $\langle f, x_0 \rangle = 1$ and considers the maps $\tau : A_+ \to E$: $a + \lambda e \mapsto a \cdot x_0 + \lambda x_0$, $a \in A$, $\lambda \in \mathbf{C}$ and $\rho : E \to A_+$: $y \mapsto y \otimes f$. We see that both are morphisms in **A-mod**, and the equation $\tau \circ \rho = \mathbf{1}_E$ represents E as a retract of a free Banach left A-module A_+.

Within the traditional class of von Neumann algebras, the spatially projective algebras admit a full description:[4]

THEOREM 1.1.4 [97]. *A von Neumann algebra A is spatially projective iff the following two conditions are satisfied:*

(i) *A decomposes into a direct sum of a family of type I factors* (or, equivalently, A is *atomic*, that is, generated by its minimal projections),

(ii) *this decomposition does not contain any infinite factor having infinite commutant.*

Thus we see, in particular, that none of von Neumann algebras of types II and III, and none of infinite von Neumann algebras of type I in the standard form are spatially projective. (In fact, it is the case of type I factor in the standard form which turns out to be most troublesome in the course of the proof.)

As an application of this theorem, we present a homological characterization of some important classes of von Neumann algebras. Recall that a module over an involutive algebra associated (cf. 0.1) with an involutive representation of this algebra in a Hilbert space, is called a *Hilbert* (or *star*) module. (Their intrinsic definition is given, e.g., in [96].) It is known that a Hilbert module over a von Neumann algebra is normal iff its associated representation is ultraweakly continuous (or, equivalently, its image is again a von Neumann algebra; cf. [123]).

PROPOSITION 1.1.5 [100]. *Let A be a von Neumann algebra. Then*

(i) *A has minimal projections iff there exists at least one non-zero normal Hilbert A-module, which is projective,*

(ii) *A is atomic iff there exists at least one non-zero normal Hilbert A-module, which is projective and faithful,*

(iii) *A is atomic and finite iff all normal Hilbert A-modules are projective.*

In particular, a given factor A is of type I iff there exists a non-zero normal Hilbert A-module which is projective, and it is of type I and finite (that is, isomorphic to a complete matrix algebra) iff all normal Hilbert A-modules are projective.

[4] *Added in proof*: now all spatially projective operator C^*-algebras and, moreover, all projective Hilbert modules over C^*-algebras are described.

Outside the scope of Banach structures, we know much less about projective modules. Here is one of their useful classes.

PROPOSITION 1.1.6 (Ogneva [154]). *Let $\mathcal{M}$ be an infinitely smooth manifold, and $\mathcal{U}$ be its open subset. Then the Fréchet left $C^\infty(\mathcal{M})$-module $C^\infty(\mathcal{U})$ (see Example* 0.2.6) *is projective.*

1.1.2. *Injective and flat modules.* Again, unless the contrary is explicitly stated, A is a fixed Banach algebra. By reversing arrows in Definition 1.1.1, we get

DEFINITION 1.1.2. A Banach $\langle$ left | right | two-sided $\rangle$ A-module J is called *injective*, if, for any Banach A-module X of the respective category, every admissible monomorphism $i : J \to X$ is a coretraction in $\langle$ **A-mod** | **mod-A** | **A-mod-A** $\rangle$ (that is, it has a left inverse morphism of Banach $\langle$ left | right | two-sided $\rangle$ A-modules.

The general statements about projective Banach modules, formulated at the beginning of the previous subsection, have natural analogues for injective modules. As to details, see, e.g., [89]. Here we only mention that the role of free $\langle$left | right | two-sided$\rangle$ modules now passes to so-called cofree modules, having the form $\langle \mathcal{B}(A_+, E) \mid \mathcal{B}(A_+, E) \mid \mathcal{B}(A_+ \widehat{\otimes} A_+, E)\rangle$ with the outer multiplication/s defined by $\langle [a \cdot \varphi](b) = \varphi(ba) \mid [\varphi \cdot a](b) = \varphi(ab) \mid [a \cdot \varphi](b \otimes c) = \varphi(b \otimes ca)$ and $[\varphi \cdot a](b \otimes c) = \varphi(ab \otimes c)\rangle$. In particular, putting E as $\mathbf{C}$, we get $\langle A_+^* \mid A_+^* \mid (A_+ \widehat{\otimes} A_+)^*\rangle$ (or $\langle A^* \mid A^* \mid (A \widehat{\otimes} A)^*\rangle$, if A has an identity) as a simplest example of an injective Banach $\langle$ left | right | two-sided $\rangle$ A-module.

Thus, for a unital A, the A-bimodule $(A \times A)^*$ of bounded bilinear functionals on A with the operations defined by $[a \cdot f](b, c) = f(b, ca)$ and $[f \cdot a](b, c) = f(ab, c)$ is, being isomorphic to $(A \widehat{\otimes} A)^*$, injective. Somewhat less trivial is that, in the case of a von Neumann algebra, the closed subbimodule $(A \times A)_*$ in $(A \times A)^*$, consisting of separately ultraweakly continuous bilinear functionals, is also injective [91].

(As to the A-bimodule A_+^*, now let us just say that "it is much more often injective, than A_+ itself is projective"; nevertheless this happens only when A is sufficiently nice. About such algebras, well-known under the name of amenable, much will be said in Section 1.3.)

REMARK. The definition of the injectivity obviously can be carried over to more general classes of topological modules. However, the symmetry between the "projective" and the "injective" constructions and results fails to be preserved outside the scope of Banach structures. The matter of fact is that the notion of a cofree module, unlike that of a free module, loses its sense: for a non-Banach A the space $\mathcal{B}(A_+, E)$, even A_+^* for a Fréchet algebra A, has no natural topological module structure (see [209] for a more detailed discussion). This is the main reason, why we know so little about injective non-Banach modules. As to open problems, the following seems rather provocative: *is it true that for every Fréchet algebra A there exists at least one non-zero injective Fréchet A-module?* As to concrete algebras, the Fréchet algebra $\mathcal{O}(\mathbf{D})$ of holomorphic functions on $\mathbf{D}$ appears to be most suspicious: so far, no injective $\mathcal{O}(\mathbf{D})$-module, save zero, has been found.[5]

[5] Now we know, thanks to Pirkovskii that a lot of topological algebras, including $\mathcal{O}(\mathbf{D})$, has no non-zero injective modules at all.

We return to Banach modules. The properties of projectivity and injectivity are connected by the following important assertion.

PROPOSITION 1.1.7. *A Banach A-module* (of any type), *dual to a projective A-module, is injective.*

Indeed, if the diagram $F \underset{\rho}{\overset{\tau}{\rightleftarrows}} P$ presents P as a retract of a free module F, then the "adjoint" diagram $F^* \underset{\rho^*}{\overset{\tau^*}{\leftrightarrows}} P^*$ presents P^* as a retract of F^*. The latter, as is easy to check, is cofree and hence injective.

However, the predual to an injective dual module is not bound to be projective. Before displaying a variety of examples, let us make an observation.

PROPOSITION 1.1.8. *The two following properties of a Banach left A-module Y are equivalent:*

(i) *for every exact sequence $\mathcal{X}$ in* **mod-A** *with admissible morphisms the sequence* $\mathcal{X} \mathbin{\widehat{\otimes}}_A Y$ (that is, the result of the applying the functor $? \mathbin{\widehat{\otimes}}_A Y$ to $\mathcal{X}$) *is exact,*

(ii) *the dual Banach right A-module Y^* is injective.*

DEFINITION 1.1.3. A Banach left A-module is called *flat*, if it has the indicated (equivalent) properties. A Banach $\langle$ right | two-sided $\rangle$ A-module is called *flat*, if it is flat as a Banach left $\langle A^{\mathrm{op}} \mid A^{\mathrm{env}} \rangle$-module.

REMARK. The definition of the flatness in the terms of tensor products, being suitably modified, can be transferred to Fréchet, $\widehat{\otimes}$- and $\overline{\otimes}$-modules. At the same time, the alternative definition in terms of dual modules, so useful in the realm of Banach structures, looses its sense (cf. the previous remark).

Now our point is that in analysis, as well as in abstract algebra, we observe the existence of two concepts of "what is homologically best": the more rigid requirement of the projectivity, and the more liberal and flexible requirement of the flatness. The interplay of both is a profound and ubiquitous phenomenon, affecting analysis not less than algebra. Let us proceed to the comparative picture of various manifestations and guises of both properties.

We begin with a pair of results, which are most important to applications and belong to the general theory of Banach algebras. Here the difference between both properties manifests itself as the difference between algebras possessing an identity and algebras, possessing a mere bounded approximate identity (abbreviated further to b.a.i.).

To speak informally, b.a.i.'s naturally appear in homological considerations for two reasons. First, their presence allows to employ the Rieffel's interpretation of the Banach space $X \mathbin{\widehat{\otimes}}_A I$ (see Proposition 0.4.2 above) and so to work with the definition of the flatness, based on tensor products. Secondly, it is known that A has a right b.a.i. iff its bidual algebra A^{**} (with the Arens multiplication) has a "real" right identity [11]. The latter fact, together with some of its modifications, gives the possibility to prove or disprove the existence of some morphisms between dual modules, which help to check the second definition of flatness. Using either of these approaches, one can establish the following two things concerning ideals and cyclic modules. (Parts "(i)" of both statements were already mentioned earlier, and we repeat them to provide a background to more substantial parts "(ii)".)

PROPOSITION 1.1.9. *Let I be a closed left ideal in A_+. Then*
(i) *it is projective provided it has a right identity,*
(ii) [82] *it is flat provided it has a right b.a.i.*

THEOREM 1.1.5. *Let X be a Banach cyclic left A-module A_+/I, where A and I are as above. Then*
(i) *X is projective provided I has a right identity. Moreover, if I is a complemented subspace in A_+* (see Introduction), *then the converse is also true;*
(ii) ([82] and [201] combined) *X is flat provided I has a right b.a.i. Moreover, if I is a weakly complemented subspace in A_+* (see idem), *then the converse is also true.*

Since it is known that left ideals in C*-algebras certainly have right b.a.i., both results immediately imply

COROLLARY 1.1.1. *All closed left ideals of, and all cyclic modules over C*-algebras are flat. In particular, every operator C*-algebra A, possessing an (algebraically) cyclic vector, is spatially flat (i.e. the relevant Hilbert space is flat as a natural A-module).*

The most important special case of Theorem 1.1.5, namely that of A-bimodule (= cyclic A^{env}-module) A_+, will be discussed at length in Section 1.3.

We pass from the general theory of Banach algebras to traditional areas of its applications. First, notice that, as a particular case of Proposition 1.1.9, *every closed ideal in* $C_0(\Omega)$, for any locally compact space Ω, *is flat.* Comparing this with Theorem 1.1.1, we can conclude that, from the point of view of general topology, the difference between the flatness and the projectivity shows itself as the difference between arbitrary and paracompact locally compact spaces.

In harmonic analysis the same homological difference takes the form of a difference between amenable and compact groups:

PROPOSITION 1.1.10 [89]. *Let G be a locally compact group, $\mathbf{C}$ be the augmenting $L^1(G)$-module* (see Example 0.2.4). *Then*
(i) *$\mathbf{C}$ is projective iff G is compact,*
(ii) *$\mathbf{C}$ is flat iff G is amenable.*

(One will draw the same conclusion about that difference from the more important Theorem 1.3.6 and Proposition 1.3.7 below.)

Now recall the properties of projective maximal ideals presented in Theorem 1.1.2. The "flat" version of the statement, concerning the analytic structure, is

THEOREM 1.1.6 (Pugach, cf. [170,172]). *Let A, Ω and $\partial\Omega$ be as in Theorem* 1.1.2, *I be a maximal ideal in A, represented by a point $s \in \Omega$. Let, further, I be a flat Banach A-module. Then* $\dim I/\overline{I^2} \leqslant 1$. *Moreover, if $I \neq \overline{I^2}$, then there exists a subset V_s in Ω which* *is an analytic disk and forms a neighbourhood of s in the Gleason topology* (*that is, in the topology of Ω as of a subset in A^* with its norm topology*).

Thus, in both cases the point s belongs to some analytic disk. However, in terms of the most natural topology on Ω, that is the Gel'fand topology, the projectivity provides that this disk forms the whole neighbourhood of s, whereas the flatness only ensures that it is situated, so to speak, like a fiber.

EXAMPLE 1.1.1. Consider the known "cylinder algebra" consisting of functions, defined and continuous on the cylinder $\Omega = \overline{\mathbf{D}} \times [0, 1]$ and, for every fixed $t \in [0, 1]$, holomorphic on the "fiber" $\{(z, t)\colon z \in \mathbf{D}\}$. The spectrum of this algebra is Ω, and $\partial\Omega$ is $\mathbf{T} \times [0, 1]$. Here points of $\partial\Omega$ represent projective maximal ideals whereas remaining points, i.e. points of "fibers", correspond to flat non-projective maximal ideals.

Now turn to von Neumann algebras. We have seen that to require their spatial projectivity is to make high demands of their structure. The following result shows that the requirements of the spatial flatness and injectivity (that is, of the respective properties of the natural module) are much less restrictive.

THEOREM 1.1.7 [98]. *Every hyperfinite von Neumann algebra is spatially flat and spatially injective.*

(As to numerous equivalent definitions of hyperfinite algebras, including the celebrated "Connes injectivity", see, e.g., [28]. Recall that all von Neumann algebras of type I, as well as various algebras of other types, are hyperfinite.)

Another result concerns an important class of non-selfadjoint algebras.

THEOREM 1.1.8 (Golovin [60]). *Let A be an indecomposable CSL-algebra* (see Introduction), *acting on a Hilbert space H. Then the following conditions are equivalent:*

(i) *A is spatially projective,*
(ii) *the closure of the algebraic sum* (*i.e. the lattice-theoretic union*) *of all spaces in* Lat A *other than H, does not coincide with H,*
(iii) *A contains a column of rank-one operators.*

Again, the conditions of the spatial flatness are much more liberal.

THEOREM 1.1.9 (Golovin [62]). *Let A be as above. If it is spatially $\langle$ flat $|$ injective $\rangle$ then the $\langle$ algebraic sum $|$ intersection $\rangle$ of every finite set of spaces in* Lat A *other than $\langle H \mid \{0\} \rangle$ does not coincide with $\langle H \mid \{0\} \rangle$, and this necessary condition of the spatial $\langle$ flatness $|$ injectivity $\rangle$ is also sufficient provided A is of finite width* (see Introduction).

As an illuminating special case, we have

COROLLARY 1.1.2 [59,60]. *Let A be a nest algebra. Then*

(i) *A is spatially projective iff H has an immediate predecessor in* Lat A,
(ii) *A is always spatially flat and injective.*

1.2. *Derivations and module extensions: One-dimensional homology theory*

The contents of Section 1.1 may suggest a philosophy of what is a "homologically good" algebra: it is an algebra with a reasonably wide class of projective, or at least flat modules. This section, however, begins with a concept, which seems at first sight to be completely unrelated to such properties of modules: a derivation. The latter notion gives rise to another, apparently different approach to "good" algebras: loosely speaking, these are algebras A with a reasonably wide class of bimodules X such that all derivations of A with values in X are inner (have, in a sense, a trivial structure). As a matter of fact, both of these approaches are intimately connected and give the same classes of "homologically best" algebras. This connection will be established in Subsection 1.2.2 with the help of the notion of a module extension, which is naturally interesting in its own right.

1.2.1. *Derivations and one-dimensional cohomology.* Here is the obvious topological version of a well known algebraic notion.

DEFINITION 1.2.1. Let A be a topological algebra, X be a topological A-bimodule. A continuous operator $D: A \to X$ is called a *derivation of A with values in* X, if it satisfies the "Leibnitz identity" $D(ab) = a \cdot D(b) + D(a) \cdot b$.

Derivations are also called *one-dimensional cocycles of A with coefficients in X*; this terminology will be clarified in the second part of our exposition. With fixed A and X, derivations form a linear space (a closed subspace of $\mathcal{B}(A, X)$ in the context of Banach structures), denoted by $Z^1(A, X)$.

REMARK. We stress that our derivations are always supposed to be continuous. However, it should be mentioned that in a number of well defined situations the continuity of derivations may be deduced from their purely algebraic properties. The wide circle of important problems, concerning the phenomenon of this "automatic continuity" we regretfully leave outside the scope of our exposition; see, e.g., [35,158] and, especially, the forthcoming book of Dales [36].

There are sound reasons, why the study of derivations is important, and we shall gradually clear them up. Let us begin with the "classical" case $X = A$. The relevant derivations turn out to be closely connected with automorphisms of a given algebra and thus help to know how rich is its stock of intrinsic symmetries.

THEOREM 1.2.1 (cf. [111]). *Let A be a Banach algebra. A continuous operator* $\mathcal{U}: A \to A$ *is an automorphism of A if, and in the case when its spectrum belongs to* $\{z \in \mathbf{C}: -\frac{2\pi}{3} < \arg z < \frac{2\pi}{3}\}$ *only if,* $\mathcal{U}$ *has the form* e^D *for some derivation* $D: A \to A$.

Note that the value $\frac{2\pi}{3}$ in this assertion is the best possible [111].

COROLLARY 1.2.1. *If an automorphism of A satisfies the indicated condition on its spectrum, then it belongs to some one-parameter group of automorphisms of A.*

(For the physical meaning of such groups – they provide a model for the development over time of systems in quantum mechanics – see, e.g., [13].)

It is easy to see that $Z^1(A, A)$ is a Lie algebra with the standard Lie bracket $[D_1, D_2] = D_1 D_2 - D_2 D_1$. Theorem 1.2.1 can be interpreted as the assertion that this is the Lie algebra of the automorphism group of A.

In a sense, the presence of meaningful derivations of a Banach algebra with values in itself is a "non-commutative" property:

THEOREM 1.2.2 (Singer and Wermer [205]). *The image of any derivation of a commutative Banach algebra A with values in A lies in the (Jacobson) radical of a given algebra.*[6] *As a consequence, semisimple commutative Banach algebras have no non-zero derivations.*

(The usual derivation in $C^\infty(\mathbf{R})$ tells us that this theorem can not be extended outside the scope of Banach structures.)

There is another, less sophisticated connection between derivations and homomorphisms, this time concerning arbitrary bimodules. If X is a topological bimodule over a topological algebra A, the topological direct sum $A \oplus X$ is evidently a topological algebra with the multiplication $(a, x)(b, y) = (ab, a \cdot y + x \cdot b)$. One can easily check

PROPOSITION 1.2.1. *The map, sending D to $\varphi: a \mapsto (a, D(a))$, is a bijection between $Z^1(A, X)$ and the set of continuous homomorphisms from A to $A \oplus X$, that are left inverse maps to the natural projection of $A \otimes X$ onto A.*

Now it is time to make a principal observation. The very definition of a bimodule already provides some stock of derivations:

DEFINITION 1.2.2. Let A be a topological algebra, X be a topological A-bimodule, $x \in X$. The map $D_x : A \to X$: $a \mapsto a \cdot x - x \cdot a$ is called an *inner derivation* of A with values in X, generated by x.

Inner derivations are also called *one-dimensional coboundaries* of A with coefficients in X. Of course, they are indeed derivations, and they form a linear subspace in $Z^1(A, X)$, denoted by $B^1(A, X)$. (It is not bound to be closed in the case of Banach structures.) Note that $B^1(A, A)$ is a Lie ideal in the Lie algebra $Z^1(A, A)$.

In the situation, described by Theorem 1.2.1, one can check that the map $D \mapsto e^D$ sends inner derivations to inner automorphisms (those, acting as $a \mapsto bab^{-1}$ for some fixed invertible $b \in A$). Therefore, if we somehow manage to prove that every $D \in Z^1(A, A)$ is inner, then we can describe all one-parameter automorphism groups of A: they are maps $\mathbf{R} \to \mathcal{B}(A)$: $t \mapsto U_t$, where U_t acts as $a \mapsto e^{tb} a e^{-tb}$ for some fixed $b \in A$.

This was one of many examples which show that it is important to know whether, for a given pair (A, X), every derivation from A to X is a priori inner. The following characteristic of such a pair shows, how abundant is the stock of outer (that is, not inner) derivations.

[6] In fact, the result is true even for discontinuous derivations (Thomas [211]).

DEFINITION 1.2.3. The quotient space $Z^1(A, X)/B^1(A, X)$ is called the *one-dimensional cohomology group of the* (topological) *algebra A with coefficients in X* and denoted by $\mathcal{H}^1(A, X)$.

(Cohomology groups of higher dimensions will be discussed in the second part of the present exposition.)

Here is another example of interpretations of one-dimensional cohomology groups, this time arising in the study of perturbations of operator algebras.

THEOREM 1.2.3 (Christensen [19]). *Let A be an operator C^*-algebra acting on a Hilbert space H. Then the following conditions are equivalent:*

(i) $\mathcal{H}^1(A, \mathcal{B}(H)) = 0$;

(ii) *there exists $k > 0$ such that for any $x \in \mathcal{B}(H)$ there exists z in the commutant of A such that $\|x - z\| \leqslant k \sup\{\|xy - yx\|\colon\ y \in A;\ \|y\| \leqslant 1\}$ (so, speaking informally, if x "gives small commutators" with elements of A, then it must be close to the commutant of A);*

(iii) *there exists $k > 0$ such that for each $*$-isomorphism α of A in $\mathcal{B}(H)$ such that $\|\alpha - \mathbf{1}_A\| < \frac{1}{k}$ there exists an unitary $v \in \mathcal{B}(H)$ such that α acts as $x \mapsto v^*xv$ and $\|\mathbf{1}_H - v\| \leqslant \sqrt{2}k\|\alpha - \mathbf{1}_A\|$ (so, if an $*$-isomorphism is close to the identity on A, then it is implemented by a unitary close to the identity operator).*

The greater part of results on one-dimensional cohomology will be presented a little bit later, when some general methods, connected with the material of the previous section, will be prepared. Now we shall mention several important facts, obtained by direct proof that a given derivation is inner (or by an explicit display of a counter-example). These results concern mostly operator algebras, and their proofs, sometimes rather sophisticated, rely heavily on their specific features. One of the most conspicuous is

THEOREM 1.2.4 (Kadison [119], Sakai [187]). *If A is a von Neumann algebra, then $\mathcal{H}^1(A, A) = 0$.*

Such a thing is obviously not true for general C^*-algebras. As an easiest example, $\mathcal{H}^1(\mathcal{K}(H), \mathcal{K}(H)) \neq 0$: indeed, every $a \in \mathcal{B}(H)$, which is not a compact perturbation of a scalar operator, generates the outer derivation $D: b \mapsto ab - ba$ in $\mathcal{K}(H)$. At the same time, as will be made clear in the next subsection, $\mathcal{H}^1(\mathcal{K}(H), \mathcal{B}(H)) = 0$. Now we cite the result, generalizing this last assertion.

THEOREM 1.2.5 (Akemann et al. [1] and Akemann and Pedersen [2] combined). *The following properties of a separable C^*-algebra A are equivalent:*

(i) *A is the direct sum of a C^*-algebra with continuous trace and some family of simple C^*-algebras,*

(ii) *$\mathcal{H}^1(A, \mathcal{D}(A)) = 0$ where $\mathcal{D}(A)$ is the double centralizer algebra of A.*

(Recall that if A is realized in a Hilbert space H, then $\mathcal{D}(A)$ can be identified with the subalgebra $\{a\colon ab, ba \in A$ for any $b \in A\}$ of $\mathcal{B}(H)$; see, e.g., [165].)

REMARK. Thus we see that, beginning with the study of derivations of A with values in A only, one inevitably comes to consider other bimodules.

One more non-trivial result, due to combined efforts of Johnson and Parrot [117] and Popa [168], asserts that $\mathcal{H}^1(A, \mathcal{K}(H)) = 0$ *for a von Neumann algebra* A, *acting on a Hilbert space* H (cf. Section 0.2). At the same time, the following problem still remains open: *is it true that* $\mathcal{H}^1(A, \mathcal{B}(H)) = 0$ *for every von Neumann algebra* (*or*, which is in fact equivalent, *for every operator* C*-*algebra*), *acting in* H? This problem was raised by Christensen [19], who showed its connection with some other problems of operator algebra and gave a positive answer in several important cases. Notably, the answer is yes, if A has a topologically cyclic vector. (Some other cases will be discussed in the next subsection).

Compared with the case of von Neumann algebras, the existing results on the derivations of another important class of reflexive algebras have, at least at the moment, a somehow more accomplished form:

THEOREM 1.2.6 (Christensen [18]). *Let* A *be a CSL-algebra, acting on* H. *Then for all ultraweakly closed algebras* B *such that* $A \subseteq B \subseteq \mathcal{B}(H)$, *the spaces* $\mathcal{H}^1(A, B)$ *coincide up to a topological isomorphism. Moreover, if* A *is a nest algebra, then* $\mathcal{H}^1(A, B) = 0$ *for all these* B.

(See Theorem 2.4.16 for a "higher-dimensional" extension of this result.)

However, a general CSL-algebras are not bound to have trivial one-dimensional cohomology:

EXAMPLE 1.2.1 (Gilfeather [53]; see also [56]). Consider four arbitrary non-zero Hilbert spaces H_k; $1 \leqslant k \leqslant 4$ (one can have, for example, $H_1 = \cdots = H_4 = \mathbf{C}$), and take

$$H = \bigoplus_{k=1}^{4} H_k.$$

Let A be the operator algebra on H consisting of operator represented by block-matrices

$$\begin{bmatrix} * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ * & * & * & 0 \\ * & * & 0 & * \end{bmatrix}$$

where "$*$" denotes an arbitrary bounded operator between respective direct summands of H. Then $\mathcal{H}^1(A, \mathcal{B}(H)) \neq 0$.

(Again, a more advanced result, generalizing this example, will be discussed later; see Corollary 2.4.4, such examples will be included in the general framework of the operator-theoretic construction of suspension.)

EXAMPLE 1.2.2 (cf. Gilfeather, Hopenwasser and Larson [54]). Consider a countable family H_k, $k = 1, 2, \ldots$, of arbitrary non-zero Hilbert spaces and take the orthogonal direct sum H of all of them. Let A be an operator algebra on H consisting of all operators a such that $a(H_k) \subseteq H_k$ for even k, $a(H_k) \subseteq H_{k-1} \oplus H_k \oplus H_{k+1}$ for odd $k > 1$, and $a(H_1) \subseteq H_1 \oplus H_2$. Then, for any B as in Theorem 1.2.6, $\mathcal{H}^1(A, B) \simeq l_\infty/s$, where s consists of such $\xi = (\xi_1, \ldots, \xi_n, \ldots) \in l_\infty$ that the sequence

$$\eta_n = \sum_{k=1}^{n} \xi_k; \quad n = 1, 2, \ldots$$

is bounded. Thus, in this case $\mathcal{H}^1(A, B)$ is infinite-dimensional.

The above-mentioned result of Popa has no "non-selfadjoint" version:

EXAMPLE 1.2.3 [54]. Let A be the same algebra as in the previous example, but with the additional assumption that all H_k are finite-dimensional. Then

$$\mathcal{H}^1\big(A, \mathcal{K}(H)\big) = l_\infty/s_0$$

where s_0 consists of such ξ that the sequence η_n (see above) tends to 0.

A recent paper of Bowling and Duncan [12] contains a lot of computations of spaces $\mathcal{H}^1(A, X)$, mostly with $X = A$ or A^*, for various semigroup algebras A.

1.2.2. *Module extensions and the space* Ext

DEFINITION 1.2.4. Let A be a topological algebra, X and Z be topological $\langle$ left | right | two-sided $\rangle$ A-modules; then a triple (Y, j, i) consisting of a topological $\langle$left | right | two-sided$\rangle$ A-module Y and admissible morphisms i and j such that the sequence

$$0 \leftarrow X \xleftarrow{j} Y \xleftarrow{i} Z \leftarrow 0$$

is exact, is called a *singular* (or admissible) *extension of X by Z*.

Thus, in the indicated situation Z can be identified (up to a module isomorphism) with a closed submodule of Y, and X with the quotient module Y/Z.

The simplest example is provided by the so-called direct sum extension, where Y is the direct sum of X and Z as of modules, and the relevant morphisms are the natural embedding and the quotient map. The typical question is whether, for given X and Z, there exist module extensions other than the direct sum extension (and, if the answer is yes, how to describe them). The appropriate terminology is as follows.

DEFINITION 1.2.5. Two singular extensions (Y_k, j_k, i_k), $k = 1, 2$, of X by Z are called *equivalent*, if there is a commutative diagram

$$\begin{tikzcd}
 & Y_1 \arrow[dl, "j_1"'] \arrow[dd, "\kappa"] & \\
X & & Z \arrow[ul, "i_1"'] \arrow[dl, "i_2"] \\
 & Y_2 \arrow[ul, "j_2"] &
\end{tikzcd}$$

in which κ is an isomorphism of topological (bi)modules. An extension is said to *split,* if it is equivalent to a direct sum extension.

The set of equivalence classes of singular extensions of X by Z is denoted by $\mathrm{Ext}(X, Z)$, or, more precisely, $\langle {}_A\mathrm{Ext}(X, Z) \mid \mathrm{Ext}_A(X, Z) \mid {}_A\mathrm{Ext}_A(X, Z) \rangle$, if we need to indicate that we consider just $\langle$left $\mid$ right $\mid$ two-sided$\rangle$ modules. It can be identified with some linear space with the zero vector corresponding to the class of split extensions. Indeed, restricting ourselves, say, to left modules, one can consider, originating from a given extension, the separately continuous bilinear operator $f : A \times X \to Z$, which sends a pair a, x to

$$i^{-1}\big(a \cdot \rho(x) - \rho(a \cdot x)\big),$$

where ρ is an arbitrarily chosen right inverse continuous operator to the (admissible!) morphism j. It is not difficult to see that f satisfy the identity $f(ab, x) = a \cdot f(b, x) + f(a, b \cdot x)$. One can prove that there is a natural bijection between $\mathrm{Ext}(X, Z)$ and the quotient space E_1/E_2 where E_1 is the space of such bilinear operators, and E_2 consists of those of the form $f(a, x) = a \cdot \varphi(x) - \varphi(a \cdot x)$ for some continuous operator $\varphi : X \to Z$.

It is obvious that an extension splits iff j has a right inverse morphism or, equivalently, i has a left inverse morphism (of topological modules). This observation, combined with main definitions of Section 1.1, immediately gives

PROPOSITION 1.2.2. *Let A be a topological algebra. A topological A-module* $\langle X|Z\rangle$ (of any type) *is* $\langle$*projective* $|$*injective*$\rangle$ *iff* $\mathrm{Ext}(X, Z) = 0$ *for every topological A-module* $\langle Z|X\rangle$.

Less obvious is

PROPOSITION 1.2.3. *Let A be a Banach algebra. Then a Banach A-module X* (of any type) *is flat iff* $\mathrm{Ext}(X, Z) = 0$ *for any dual Banach A-module Z.*

The proof uses a natural linear isomorphism between the spaces ${}_A\mathrm{Ext}(X, Z^*)$ and $\mathrm{Ext}_A(Z, X^*)$, which, in its turn, can be obtained from the description of Ext as of a quotient space, given above.

The following theorem establishes the close connection between derivations and module extensions, and lays a foundation for subsequent results on homologically trivial algebras.

THEOREM 1.2.7. *Let A be a topological algebra, X a topological A-bimodule. There exists a linear isomorphism between spaces* ${}_A\operatorname{Ext}_A(A_+, X)$ *and* $\mathcal{H}^1(A, X)$.

To prove it, one can observe that a bimodule extension (Y, j, i) of A_+ by X gives rise to trilinear operators of the form

$$f : A_+ \times A_+ \times A_+ \to X\colon\ (a, b, c) \mapsto i^{-1}\big(a \cdot \rho(b) \cdot c - \rho(abc)\big),$$

where ρ is any right inverse continuous operator to j. This construction leads to an isomorphism between the space ${}_A\operatorname{Ext}_A(A_+, X)$ and a certain quotient space of the space of all trilinear operators, obtained by this way. Further, such an f, in its turn, gives rise to a derivation

$$D : A_+ \to X\colon\ a \mapsto f(a, e, e) - f(e, a, e).$$

The map $f \mapsto D$ is a linear operator, which generates the desired isomorphism as an operator between respective quotient spaces.

In the scope of Banach structures, there are important cases, when the groups $\mathcal{H}^1(A, \cdot)$ are isomorphic to some one-sided (and not just to two-sided, as above) Ext spaces. The proof of the following theorem resembles that of the previous one, but the calculations are somewhat easier.

THEOREM 1.2.8. *Let A be a Banach algebra, X and Y be Banach left A-modules. There is a linear isomorphism between spaces* ${}_A\operatorname{Ext}(X, Y)$ *and* $\mathcal{H}^1(A, \mathcal{B}(X, Y))$.

(As to the bimodule structure in $\mathcal{B}(X, Y)$, see Section 0.4.)

Combining this with Proposition 1.2.2 and 1.2.3, we obtain

COROLLARY 1.2.2. *Let A be a Banach algebra. A Banach left A-module $\langle X | Y | X \rangle$ is $\langle$ projective | injective | flat $\rangle$ iff $\mathcal{H}^1(A, \mathcal{B}(X, Y)) = 0$ for $\langle$ any | any | any dual $\rangle$ Banach left A-module $\langle Y | X | Y \rangle$.*

In turn, this corollary, when combined with various results on homological properties of concrete modules, provides a lot of information about one-dimensional cohomology. As an important example, Proposition 1.1.4 implies

THEOREM 1.2.9. *Let A be an operator algebra, acting on a Banach space E. If it contains a column of rank-one operators* (as, e.g., $\mathcal{B}(E)$, $\mathcal{K}(E)$ and $\mathcal{N}(E)$) *then every derivation of A with values in the A-bimodule $\mathcal{B}(E)$ $(= \mathcal{B}(E, E))$ is inner.*

In the same way, the results on spatial flatness mentioned in Theorems 1.1.7 and 1.1.9, and also in Corollaries 1.1.1 and 1.1.2 imply that all derivations of relevant algebras with values in $\mathcal{B}(H)$ are inner. (For direct proofs of such an assertion for some of these algebras see, e.g., [122,140,153,54].)

1.3. *Amenable algebras and their species*

Now we are prepared to discuss what is a "homologically best" algebra in real earnest.

1.3.1. *Contractible algebras.* We begin with an approach, which at first sight seems to be most natural.

THEOREM 1.3.1. *The following properties of a Banach algebra A are equivalent:*

(i) $\mathcal{H}^1(A, X) = 0$ *for any Banach A-bimodule X* (in other words, every derivation of A with values in any Banach A-bimodule is inner).

(ii) *The Banach A-bimodule A_+ is projective.*

(iii) *A has an identity, and the Banach A-bimodule A is projective.*

(iv) *Every Banach module over A, two-sided as well as one-sided, is projective.*

(Characterizations, involving cohomology groups of higher dimensions, will be presented later; see Theorem 2.4.20(i) below.)

As to the proof, Theorem 1.2.7, together with Proposition 1.2.2, imply (i) $\Leftrightarrow$ (ii). If A-bimodule A_+ is projective, the same is true for the A^{env}-bimodule A^{env}; therefore (iv) follows from (ii).

DEFINITION 1.3.1. A Banach algebra A is called *contractible* (or sometimes supra-amenable) if it has the (equivalent) properties indicated in Theorem 1.3.1.

(We say "contractible" because, as we shall see in the second part, the cohomology of such an algebra vanish in all positive dimensions; so, the algebra behaves like a contractible polyhedron in topology. The term "supra-amenable" will be justified after our soon acquaintance with "genuine" amenable algebras.)

The following criterion provides an effective tool to check the property. To formulate it, note that the product map $\langle \pi \mid \pi_+ \rangle$ from $\langle A \widehat{\otimes} A \mid A_+ \widehat{\otimes} A_+ \rangle$ to $\langle A \mid A_+ \rangle$, is a morphism in **A-mod-A**, and therefore it will be referred as the *product morphism.*

PROPOSITION 1.3.1. *The following properties of a Banach algebra A are equivalent to its contractibility:*

(i) *The product morphism $\pi_+ : A_+ \widehat{\otimes} A_+ \to A_+$ has a right inverse in* **A-mod-A**,

(ii) *A has an identity, and the product morphism $\pi : A \widehat{\otimes} A \to A$ has a right inverse in* **A-mod-A**.

The definition of a contractible algebra, together with Theorem 1.3.1 and the previous proposition, can be generalized, with obvious modifications, to the Fréchet, $\widehat{\otimes}$- and $\overline{\otimes}$-algebras. As to general topological algebras (where there is no proper concept of a topological tensor product), we know only the equivalence of properties (i) and (ii) in Theorem 1.3.1. So, one can define the contractibility with the help of these properties.

The definition of a contractible algebra is attractive by its simplicity. However, its requirements are extremely severe and distinguish too scanty a class of algebras. As a matter of fact, at the moment we know of no example of a contractible Banach algebra different from a semisimple finite-dimensional algebra (that is, from a direct sum of several complete matrix algebras). In more detail, let us consider the following obvious hierarchy of possible properties of a Banach algebra A:

(1) A is semisimple and finite-dimensional
$\Rightarrow$ (2) A is contractible
$\Rightarrow$ (3) every left Banach A-module is projective
$\Rightarrow$ (4) every irreducible (in the algebraic sense) left Banach A-module is projective.

Is it true, that all these logical arrows, or at least some of them, can be reversed? So far the answer is unknown. The question, in its full generality, appears to be fairly complicated because of some notorious "pathological features of Banach space geometry" (cf. the similar discussion in Section 1.1).

REMARK. The answer to the pure algebraic prototype of this question is well known: (1) $\Leftrightarrow$ (2), (3) $\Leftrightarrow$ (4), but (2) $\not\Leftarrow$ (3). The non-contractible algebra with projective modules is the same algebra, which is well known to have no Banach algebra norm: the field of rational functions.

Nevertheless, the answer to the given question is positive if we put in some additional assumptions. For example, it can be easily obtained if we know that A is commutative [79] or $A = L^1(G)$. The most advanced result of this kind is, perhaps, as follows.

THEOREM 1.3.2 (Selivanov [192]). *Suppose that either every irreducible Banach left A-module or $A/\operatorname{Rad} A$ has, as a Banach space, the approximation property* (cf. Theorem 1.1.3). *Then properties* (1)–(4) *of A are equivalent.*

Since every irreducible module over a C^*-algebra is isomorphic in **Ban** to a Hilbert space and hence certainly has the approximation property, we get

COROLLARY 1.3.1. *For any C^*-algebra A, the properties* (1)–(4) *are equivalent.*

Coming to the end of the discussion about the contractibility property, we should mention that outside the framework of Banach structures the situation unexpectedly changes: new contractible algebras appear. For an arbitrary set M, we denote by $\mathbf{C}^M$ (cf. Introduction) the algebra of all functions on M with the topology of pointwise convergence.

THEOREM 1.3.3 (Taylor [209]; complete proof see [89]). *Let A be a commutative Arens–Michael algebra* (see Introduction). *Then its following properties are equivalent:*

(i) *A is contractible,*
(ii) *every $\widehat{\otimes}$-module over A is projective,*
(iii) *A is topologically isomorphic to $\mathbf{C}^M$ for some* (perhaps, infinite!) *set M.*

Note that the contractibility of $\mathbf{C}^M$ easily follows from Proposition 1.3.1 and the known topological algebra isomorphism $\mathbf{C}^M \widehat{\otimes} \mathbf{C}^M \simeq \mathbf{C}^{M\times M}$.

There is a natural conjecture that a general Arens–Michael algebra is contractible iff it is topologically isomorphic to a Cartesian product of some (perhaps, infinite) family of complete matrix algebras.[7] Here the "if" part is indeed true, and it is not difficult to prove it; however, the converse is established, at the moment, only under some additional assumptions [199]. Outside the class of Arens–Michael algebras, the $\overline{\otimes}$-algebra $\mathcal{E}'(G)$ of distributions on a compact Lie group G, is also contractible [208].

1.3.2. *Amenable algebras: Johnson's and Connes' versions.* The following result provides, in a sense, the most optimal of all concepts of homologically best algebras.

THEOREM 1.3.4 (cf. [88]). *The following properties of a Banach algebra A are equivalent:*

(i) $\mathcal{H}^1(A, X) = 0$ *for any dual Banach A-bimodule X;*
(ii) *the Banach A-bimodule A_+ is flat;*
(iii) *A has a bounded approximate identity, and the Banach A-bimodule A is flat;*
(iv) *every Banach module over A, two-sided as well as one-sided, is flat.*

(Again, some other equivalent properties will be presented later, in Theorem 2.4.20(ii) and (iii).)

The proof resembles that of Theorem 1.3.1, but now we use Proposition 1.2.3 instead of 1.2.2; (iv) follows from (ii) because the flatness of $A_+ \in$ **A-mod-A** implies the same property of $A^{\mathrm{env}} \in$ $\mathbf{A}^{\mathrm{env}}$**-mod-**$\mathbf{A}^{\mathrm{env}}$.

Now we introduce the class of algebras, discovered by Johnson [109], who adopted the definition based on derivations. (Flat bimodules cropped up in [83].)

DEFINITION 1.3.2. A Banach algebra A is called *amenable* (or, for precision, *Johnson amenable*) if it has the properties indicated in Theorem 1.3.4.

(The choice of the term "amenable" by Johnson will be clarified later, in Theorem 1.3.6.)

Here is a useful tool to check the property (cf. Proposition 1.3.1).

PROPOSITION 1.3.2. *The following properties of a Banach algebra A are equivalent to its amenability:*

(i) *the dual product morphism $\pi_+^* : A_+^* \to (A_+ \widehat{\otimes} A_+)^*$ has a left inverse in* **A-mod-A**,
(ii) *A has a b.a.i., and the the dual product morphism $\pi^* : A^* \to (A \widehat{\otimes} A)^*$ has a left inverse in* **A-mod-A**.

Amenable algebras have a complete description in terms of approximate identities. Applying the criteria, established in Theorem 1.1.5, to the A^{env}-module A_+, we get

THEOREM 1.3.5 (cf. [102]). *Let A be a Banach algebra, I_Δ be the diagonal ideal in A^{env}* (see 0.2). *Then*

(i) *A is contractible iff I_Δ has a right identity,*
(ii) *A is amenable iff I_Δ has a right b.a.i.*

[7] *Added in proof*: now we know that it is true for Fréchet algebras with the approximation property (Selivanov) and for locally C*-algebras (Fragoulopoulou).

For characterizations of amenability, close to the latter and formulated in terms of so-called approximate and virtual diagonals, see in [108]. As to some other characterizations, see [109,88], and also Theorem 1.3.11 below.

Outside the scope of Banach structures, the original definition of amenability, using the notion of a dual module, looses its sense. Nevertheless, in the case of Fréchet algebras and, more generally, whenever one can speak about flat modules (cf. Subsection 1.1.2), amenable algebras could be defined with the help of any of the properties (ii)–(iv) indicated in Theorem 1.3.4. However, it is not clear at the moment, whether these properties, together with the property, indicated in Theorem 1.3.5(ii), still remain equivalent.

Let no return to Banach algebras and modules. Combining Theorems 1.3.4 and 1.1.5, we get

PROPOSITION 1.3.3. *Every weakly complemented* $\langle$*left* | *right* | *two-sided*$\rangle$ *ideal of an amenable Banach algebra has a* $\langle$*right* | *left* | *two-sided*$\rangle$ *b.a.i.*

The property of amenability is preserved under a number of constructions. This is the case, for example, with tensor products and with a certain "uniform" version of inductive limit. A quotient algebra of an amenable algebra and, more generally, an algebra with a dense continuous image of some amenable algebra is amenable itself; the same is true with a closed two-sided ideal of an amenable algebra, possessing a b.a.i. See [109,88] for proofs and further details.

Now let us consider what the requirement of amenability turns out to be in concrete classes of Banach algebras. To begin with, we explain the very term "amenable".

THEOREM 1.3.6 (Johnson [109]). *Let G be a locally compact group. Then the Banach algebra $L^1(G)$ is amenable iff G is an amenable group* (in the classical sense of harmonic analysis; see, e.g., [160]).

We give an outline of the proof, based on Proposition 1.3.3(ii) (cf. [102]; the original proof see [109] or [11]). Assume, just for the sake of transparency, that G is discrete. For $A = L^1(G)$, the bimodules A^* and $(A \widehat{\otimes} A)^*$ are $L^\infty(G)$ and $L^\infty(G \times G)$ with some suitable convolutions as outer multiplications, and the dual product morphism π^* takes $f \in L^\infty(G)$ to $u_f \in L^\infty(G \times G)$, where $u_f(s,t) = f(st)$. Thus $L^1(G)$ is amenable iff there exists a bimodule morphism $\zeta : L^\infty(G \times G) \to L^\infty(G)$ such that $\zeta(u_f) = f$ for any f. Now we see that if G is amenable, and $M : L^\infty(G) \to \mathbf{C}$ is a left invariant functional, then the map ζ defined by $[\zeta(u)](s) = M(h_s)$, $s \in G$, where $h_s(t) = u(st^{-1}, t)$ is the desired morphism. Conversely, if such a morphism ζ is given, we construct a left invariant functional as $M : L^\infty(G) \to \mathbf{C}$: $f \mapsto [\zeta(u)](e)$, where $u(s,t)$ is defined as $f(s)$, and e is an identity of G.

The theorem of Johnson, combined with Proposition 1.3.3, has applications to Banach space geometry: it enables one to construct new examples of non-complementable and even not weakly complementable subspaces. Indeed, the latter proposition means that every left ideal without a right b.a.i. in an amenable algebra certainly has no weak complement. To display such an ideal, we recall that the Wiener algebra $W = L^1(\mathbf{Z})$ and, more generally, every $L^1(G)$ for an Abelian non-compact G, has so-called Malliavin ideals (cf. [124]), which do not coincide with their topological squares and therefore can not have a b.a.i. On the other hand, it is well known that Abelian groups are amenable (cf. [160]). Therefore, by virtue of Theorem 1.3.6, we get

COROLLARY 1.3.2 [102]. *Let G be an Abelian non-compact group. Then Malliavin ideals, as subspaces in $L^1(G)$, are not weakly complementable.*

For the following traditional class of Banach algebras, the picture is also absolutely clear.

THEOREM 1.3.7 (Kadison and Ringrose [122], Johnson [109], and Sheinberg [203] combined). *A uniform Banach algebra with spectrum Ω is amenable iff it coincides with $C_0(\Omega)$.*

There are many ways to prove the amenability of $C_0(\Omega)$. For example, one can use the known fact that the diagonal ideal in A^{env} for $A = C_0(\Omega)$ has a b.a.i. [41], and apply Theorem 1.3.5(ii). (As to similar applications of the latter theorem, concerning weighted convolution algebras on groups and semigroups, see [64,65]).

And what does amenability mean for C*-algebras? This problem turned out to be a strong nut, and it was cracked by combined efforts of several authors, notably Connes [29] and Haagerup [75]. Here is the answer.

THEOREM 1.3.8. *A C*-algebra is amenable iff it is nuclear (or, equivalently, its enveloping von Neumann algebra is hyperfinite).*

This theorem contains, as special cases, several earlier results. In particular, all approximately finite-dimensional C*-algebras are amenable (Kadison and Ringrose [122]), and the same is true with GCR-, or tame, C*-algebras (Johnson [109]); both results imply that $\mathcal{K}(H)$ is certainly amenable. At the same time, the reduced C*-algebra of a discrete group G is amenable iff G is amenable as a group (Bunce [15]). The case of $\mathcal{B}(H)$ will be discussed later.

REMARK. As a version of amenability, Johnson [109] has introduced so-called strongly amenable C*-algebras in some specific terms of their derivations. This interesting and useful subclass of amenable C*-algebras is discussed, e.g., in [14,75,161].

As for standard operator algebras on Banach spaces, there is the following open problem: for which Banach spaces E is the algebra $\mathcal{K}(E)$ is amenable? This is the case with many "good" spaces like C[0, 1], l_p and $L^p[0,1]$, $1 \leqslant p \leqslant \infty$ (cf. [109]). However, it is not true for the spaces $E = l_p \oplus l_q$, $1 < p < q < \infty$, $p, q \neq 2$ [70]. As to $\mathcal{N}(E)$, it is never amenable when E is infinite-dimensional [69,195]. Finally, so far, there is no example of an infinite-dimensional space E with amenable $\mathcal{B}(E)$.

To conclude the discussion about Johnson amenability, note that amenable algebras are not bound to be semisimple: relevant examples can be found among quotient algebras of $L^1(G)$ for any Abelian non-compact G. On the other hand, at the moment it is unknown, whether amenable radical Banach algebras exist (cf. [34]).[8] Failing to have an identity, radical algebras can not be contractible (see Theorem 1.3.1(iii)); however, as the example of the Volterra algebra $L^1[0, 1]$ shows, it is not forbidden to them to have a b.a.i., and so Theorem 1.3.4(iii) can not be directly applied. As to the Volterra algebra itself, it is not amenable because it has complemented ideals without b.a.i. (In fact, it is not even weakly amenable in the sense of our future Definition 1.3.6 (cf.[128]).)

* * *

As it is now universally recognized, the requirement of (Johnson) amenability has an optimal degree of severity for Banach algebras in general and for C*-algebras in particular. However, for von Neumann algebras with their peculiar features this condition turns out to be too burdensome. Indeed, Theorem 1.3.8 immediately implies that already $\mathcal{B}(H)$, not being nuclear [215], is not Johnson amenable. At the same time, the following result suggests an appropriate modification of the latter property.

THEOREM 1.3.9 [91]. *Let A be an operator* C**-algebra* (in particular, von Neumann algebra). *Then the following properties are equivalent:*

(i) $\mathcal{H}^1(A, X) = 0$ *for every normal A-bimodule X,*

(ii) *the Banach A-bimodule A_*^-* (just A_* in the case of a von Neumann algebras; see Section 0.3) *is injective.*

(Later, in Theorem 2.4.26, we shall discuss quite a few other equivalent properties.)

DEFINITION 1.3.3. An operator C*-algebra is called Connes amenable if it has the properties indicated in Theorem 1.3.9.

(It was Connes [28], who distinguished von Neumann algebras with the first of indicated properties. Of course, he used another term: "algebras, amenable as von Neumann algebras". There are also other names for the notion, like "ultraweak amenability" [141] and "normal amenability" [43]).

Comparing Theorem 1.3.4 and 1.3.9, we see that the Johnson amenability of a Banach algebra A with b.a.i., say of a C*-algebra, means just that its dual bimodule A^* is injective, whereas the Connes amenability of a von Neumann algebra A means that its predual bimodule A_*, which is in fact a direct summand of A^*, is injective.

It is not clear at the moment, whether there exists a reasonable version of Connes amenability outside the class of operator C*-algebras, say for some non-selfadjoint operator algebras, acting on Hilbert spaces, or for Banach algebras, possessing a predual bimodule (like the measure algebra $M(G)$).

Connes amenability can be also formulated in terms of a certain "canonical" morphism, like the product morphism for contractibility and the dual product morphism for Johnson amenability (cf. Propositions 1.3.1 and 1.3.2). Denote by $(A \times A)_*$ the A-bimodule of separately ultraweakly continuous bilinear functionals on an operator C*-algebra A; recall

[8] *Added in proof*: C.C. Read has managed to construct an example of amenable radical Banach algebra.

(cf. 1.1.2) that it is a closed subbimodule in $(A \widehat{\otimes} A)^* \simeq (A \times A)^*$. Further, denote by $\pi_* : A_*^- \to (A \times A)_*$ the restriction of the dual product morphism π^*; it will be called the *predual product morphism.*

PROPOSITION 1.3.4 ([91], cf. also [161]). *An operator* C*-*algebra A is Connes amenable iff the predual product morphism* $\pi_* : A_*^- \to (A \times A)_*$ *has a left inverse in* **A-mod-A**.

(This assertion is close in spirit to an earlier criterion of Effros [44] formulated in terms of so-called normal virtual diagonals.)

Here is a characterization of Connes amenability in intrinsic terms of von Neumann algebras.

THEOREM 1.3.10 (Connes [29]). *A von Neumann algebra is Connes amenable iff it is hyperfinite* (or, equivalently, "Connes injective"; cf. [28]).

Thus hyperfiniteness is in fact a homological concept. Now we show that both of versions of the amenability are closely connected.

PROPOSITION 1.3.5 (Connes [29] and Haagerup [75] combined). *A* C*-*algebra A is Johnson amenable iff its enveloping von Neumann algebra* A^{**} *is Connes amenable.*

The idea of a short proof of the more complicated "if" part is as follows [90,93]. Proposition 1.3.4, applied to A^{**}, implies that $\pi_* : (A^{**})_* \to (A^{**} \times A^{**})_*$, which is just $\pi^* : A^* \to (A \times A)^*$ (cf. [116]) is a coretraction in $\mathbf{A^{**}}$**-mod-**$\mathbf{A^{**}}$ and hence in **A-mod-A**. Thus we get the sufficient condition of Johnson amenability, formulated in Proposition 1.3.2(ii).

There is also an illuminating criterion of both varieties of amenability, given in the spirit of the Hahn–Banach extension theorem. For a Banach algebra A and $X \in$ **A-mod-A**, a continuous functional f on X is said to be *central* if $f(a \cdot x) = f(x \cdot a)$ for any $a \in A$, $x \in X$.

THEOREM 1.3.11 (Lau [141]). *A* $\langle$*Banach* | *von Neumann*$\rangle$ *algebra A is* $\langle$ *Johnson* | *Connes* $\rangle$ *amenable iff, for any* $\langle$*Banach* | *normal*$\rangle$ *A-bimodule X and any Banach sub-A-bimodule Y of X, every central functional on Y has an extension to a central functional on X.*

Finally we note that in the class of C*-algebras both types of amenability can be characterized in terms of group-theoretic amenability.

THEOREM 1.3.12 ($\langle$Paterson [162] | de la Harpe [76]$\rangle$). *A* $\langle$C*-*algebra* | *von Neumann algebra* $\rangle$ *is* $\langle$*Johnson* | *Connes*$\rangle$ *amenable iff its unitary group, equipped with the relative* $\langle$*weak Banach space* | *ultraweak*$\rangle$ *topology, is amenable.*

1.3.3. *Biprojective, biflat and weakly amenable algebras.* One can obtain another substantial class of "good" algebras by choosing one more way to weaken the requirements of contractibility.

DEFINITION 1.3.4. A Banach algebra A is called *biprojective* if the Banach A-bimodule A is projective.

Biprojective Fréchet, $\widehat{\otimes}$-, $\overline{\otimes}$- and (general) topological algebras are defined similarly.

Of course, a unital biprojective algebra is just a contractible algebra (cf. Definition 1.3.1); however the class of all biprojective algebras is wider and more interesting. The following proposition, closely resembling Proposition 1.1.1, provides a standard method to check the property.

PROPOSITION 1.3.6. *A Banach algebra A is biprojective iff the product morphism $\pi : A \widehat{\otimes} A \to A$ has a right inverse in* **A-mod-A.**

This assertion has an obvious analogue for Fréchet, $\widehat{\otimes}$- and $\overline{\otimes}$-algebras (and for some other classes, equipped with a suitable notion of a topological tensor product).

The simplest example of a non-contractible biprojective algebra is perhaps l_1 with coordinatewise multiplication: the product morphism

$$\pi : l_1 \widehat{\otimes} l_1 \simeq L^1(\mathbf{N} \times \mathbf{N}) \to L^1(\mathbf{N}) = l_1$$

obviously has a right inverse which takes "row"$(\xi_1, \dots, \xi_n, \dots)$ to "matrix"

$$\begin{pmatrix} \xi_1 & 0 & \dots \\ 0 & \xi_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

REMARK. The initial stimulus to select biprojective algebras was the vanishing of some of their high-dimensional cohomology (see Theorem 2.4.21(i) below).

Unlike contractible algebras, non-unital biprojective algebras have non-projective modules, to begin with A_+/A. Nevertheless, if A is a biprojective Banach (or, say, $\widehat{\otimes}$-) algebra, then *every $\langle$ left $|$ right $\rangle$ A-module of the form $\langle A \widehat{\otimes}_A X \mid X \widehat{\otimes}_A A \rangle$, where X is a $\langle$ left $|$ right $\rangle$ A-module, is projective* (cf. 0.4). These modules, which are said to be *reduced,* form a rather large class: for example, by virtue of Proposition 0.4.2, it contains all nondegenerate modules over Banach algebras with b.a.i.

Turning to traditional classes of algebras, we begin with the assertion, which can be considered as a counterpoint to Theorem 1.3.6.

PROPOSITION 1.3.7 [88]. *Let G be a locally compact group. Then the Banach algebra $L^1(G)$ is biprojective iff G is compact.*

Since the "comultiplication"

$$\rho : L^1(G) \to L^1(G \times G) \simeq L^1(G) \widehat{\otimes} L^1(G), \quad a \mapsto u,$$

where $u(s,t) = a(st)$, is a right inverse to the product morphism in $\mathbf{L^1(G)\text{-mod-}L^1(G)}$, the "if" part is provided by Proposition 1.3.6. A possible way to prove the "only if" part is to observe that the augmenting $L^1(G)$-module $\mathbf{C}$, being non-degenerate, is projective, and then apply Proposition 1.1.10(i).

The following criterion requires a more sophisticated argument.

THEOREM 1.3.13 ([84] and [193] combined). *A C^*-algebra is biprojective iff it decomposes into the c_0-sum of a family of complete matrix algebras* (or, equivalently, *it has discrete primitive spectrum and finite-dimensional irreducible representations*).

COROLLARY 1.3.3 [89]. *The algebra $C_0(\Omega)$ is biprojective iff Ω has the discrete topology.*

As another particular case of the previous theorem,we see that $\mathcal{K}(H)$, being, as it was observed earlier, *amenable,* is *not biprojective* unless H is finite-dimensional. On the other hand, $\mathcal{N}(H)$, possessing no b.a.i., is *not amenable* and now we shall see that it *is biprojective.* To show this, let us consider a more general class of Banach algebras, which deserves special attention.

Let (E, F) be a dual pair of Banach spaces, that is a pair, equipped with a non-degenerate bilinear functional $\langle x, y\rangle$, $x \in E$, $y \in F$. In this situation $E \widehat{\otimes} F$ is a Banach algebra with the multiplication well-defined by $(x_1 \otimes y_1) \cdot (x_2 \otimes y_2) = \langle x_2, y_1\rangle(x_1 \otimes y_2)$. It is the so-called *tensor algebra, generated by the duality* $\langle\cdot, \cdot\rangle$. Further, there is a continuous homomorphism $\sigma : E \widehat{\otimes} F \to \mathcal{B}(E)$ well-defined by $[\sigma(x_1 \otimes y)](x_2) = \langle x_2, y\rangle x_1$, $x_1, x_2 \in E$, $y \in F$. Its image, denoted by $\mathcal{N}_F(E)$, is a Banach algebra with respect to the quotient norm of $E \widehat{\otimes} F/\operatorname{Ker}\sigma$; it is called the *algebra of nuclear operators on E with respect to the duality* $\langle\cdot, \cdot\rangle$.

The standard example is, of course, provided by the dual pair $\langle E, E^*\rangle$; $\langle x, f\rangle = f(x)$; in this case $\mathcal{N}_F(E)$ is just $\mathcal{N}(E)$.

If either E or F has the approximation property, then $\operatorname{Ker}\sigma = 0$ and hence $E \widehat{\otimes} F \simeq \mathcal{N}_F(E)$ (cf. [27]). (Recall that the coincidence of $E \widehat{\otimes} E^*$ with $\mathcal{N}(E)$ is just one of equivalent definitions of the approximation property; see, e.g., [71]).

PROPOSITION 1.3.8 [193]. *Every tensor algebra, generated by a duality, is biprojective.*

Indeed, if we fix $x_0 \in E$ and $y_0 \in F$ with $\langle x_0, y_0\rangle = 1$, then the operator $\rho : E \widehat{\otimes} F \to (E \widehat{\otimes} F) \widehat{\otimes} (E \widehat{\otimes} F)$, well-defined by assigning $(x \otimes y_0) \otimes (x_0 \otimes y)$ to $x \otimes y$, is a morphism of $(E \widehat{\otimes} F)$-bimodules, which is right inverse to the product morphism π.

COROLLARY 1.3.4. *If one of the spaces in a dual pair (E, F) has the approximation property, then $\mathcal{N}_F(E)$ is biprojective. In particular, if E has the approximation property (e.g., $E = H$), then $\mathcal{N}(E)$ is biprojective.*

As a matter of fact, tensor algebras give something more than just examples of biprojective algebras. They play the role of bricks, from which an arbitrary biprojective algebra within a rather wide class is built:

THEOREM 1.3.14 (Selivanov [193]). *Let A be a biprojective semisimple Banach algebra, possessing, as a Banach space, the approximation property. Then it decomposes into a topological direct sum of a certain family of tensor algebras, generated by dualities.*

(Note that under the condition of this theorem every of the mentioned tensor algebras coincides with the respective algebra of nuclear operators.)

REMARK. If taken for granted, this theorem provides a short proof of several classical results on representations of compact groups. Combining it with Proposition 1.3.7, we see that $L^1(G)$ for a compact G is a sum of tensor algebras. Since $L^1(G)$ has a b.a.i., all of these summands must have b.a.i. themselves. But the specifics of the norm in $E \widehat{\otimes} F$ (or, equivalently, of the nuclear norm in $\mathcal{N}_F(E)$) imply that this can happen only in the case when E is finite-dimensional, in other words, when $E \widehat{\otimes} F$ coincides with a complete matrix algebra.

Non-semisimple biprojective algebras also exist: the simplest is, perhaps, the algebra, consisting of 2×2 matrices of the form

$$\begin{pmatrix} \lambda & \mu \\ 0 & 0 \end{pmatrix}, \quad \lambda, \mu \in \mathbf{C}$$

(it will be of some use in the second part of our exposition). At the same time, it is still unknown, whether commutative non-semisimple biprojective Banach algebras, as well as radical biprojective Banach algebras, exist.

Now let us briefly discuss a natural class of algebras, containing both amenable and biprojective algebras.

DEFINITION 1.3.5. A Banach algebra A is called *biflat*, if the Banach A-bimodule A is flat.

This definition can be generalized to those classes of topological algebras where the notion of a flat (bi)module still has sense. However, the following criterion requires the specifics of Banach structures:

PROPOSITION 1.3.9. *A Banach algebra A is biflat iff the dual product morphism $\pi^* : A^* \to (A \widehat{\otimes} A)^*$ has a left inverse morphism in* **A-mod-A**.

Not every module over a biflat algebra is bound to be flat, but it is the case with reduced modules, defined above. Of course, a biflat algebra with a b.a.i. is just an amenable algebra.

Every biflat Banach algebra coincides with its topological square [88].

An interesting question of a somewhat vague character is as follows: how much wider is the class of biflat algebras than "combinations"of biprojective and amenable algebras? We know no example of a biflat algebra which is not obtained from the two latter classes with the help of some standard operations like direct sum, tensor product, etc.[9]

[9] Selivanov gave an example of such a biflat algebra: it is $\mathcal{K}(\mathcal{N}(H))$ where H is an infinite-dimensional Hilbert space.

The following assertion, connected with biflatness, demonstrates that the approximation property can be considered as a homological property:

THEOREM 1.3.15 (Selivanov [195]). *The following properties of a Banach space E are equivalent:*

(i) *the Banach algebra $\mathcal{N}(E)$ is biprojective,*
(ii) *the Banach algebra $\mathcal{N}(E)$ is biflat,*
(iii) *E has the approximation property.*

Recently the following class of Banach algebras, which is wider than that of biflat algebras, has drawn attention.

DEFINITION 1.3.6. A Banach algebra A is called *weakly amenable* if $\mathcal{H}^1(A, A^*) = 0$.

Every biflat Banach algebra is weakly amenable: for such an A, the Banach A-bimodule A^* is injective, and hence the space ${}_A\mathrm{Ext}_A(A_+, A^*)$ which is no other thing than $\mathcal{H}^1(A, A^*)$ (cf. Subsection 1.2.2), vanishes. The converse is not true: one can get a lot of relevant examples, combining either of Theorems 1.3.6 and 1.3.8 with Theorem 1.3.17 below. (One more example, due to Aristov [3] – l_2 with the coordinatewise multiplication – will be discussed in 2.6.1.)

In the beginning, weakly amenable algebras were introduced in the commutative case by Bade, Curtis and Dales [8] as those with the following property: $\mathcal{H}^1(A, X) = 0$ for all X satisfying the identity $a \cdot x = x \cdot a$ (this means just that every derivation of A with the values in such an X is zero). In the cited paper they established the equivalence of both definitions for commutative algebras. In the same commutative case, a weakly amenable algebra coincides with its topological square, any quotient algebra of a weakly amenable algebra is weakly amenable, and a closed ideal in a weakly amenable algebra is weakly amenable iff it coincides with its topological square (cf. [34]). The following criterion is curiously parallel to Theorem 1.3.5(ii).

THEOREM 1.3.16 (Grønbaek [66]). *A commutative Banach algebra A is weakly amenable iff the diagonal ideal I_Δ in A^{env} coincides with its topological square.*

(See Runde [186] for an interesting alternative proof.)

Two principal (and rather difficult) results, concerning concrete classes of algebras, we combine in the following

THEOREM 1.3.17.

(i) (Haagerup [75]) *Every* C*-*algebra is weakly amenable.*
(ii) (Johnson [115]) *for every locally compact group G, the Banach algebra $L^1(G)$ is weakly amenable.*

(A new elegant proof of Johnson's result was presented by Despic and Ghahramani [38].)

Amenable and weakly amenable algebras among Beurling and Lipschitz algebras are characterized by Bade, Curtis and Dales [8].

Contrary to the case of amenable algebras (cf. previous subsection), examples of weakly amenable radical Banach algebras are known. One can come across such algebras among quotient algebras of certain ideals in $L^1(G)$ for an Abelian non-compact G (Curtis [34]).

We conclude our exposition with a diagram, which shows relations between principal classes of algebras, discussed in this section.

$$
\begin{tikzcd}
 & \text{biprojectivity} \arrow[r, Rightarrow] & \text{biflatness} \arrow[r, Rightarrow] & \text{weak amenability} \\
\text{contractibility} \arrow[ur, Rightarrow] \arrow[dr, Rightarrow] & & & \\
 & \text{Johnson amenability} \arrow[uur, Rightarrow] \arrow[r, Rightarrow] & \text{Connes amenability} &
\end{tikzcd}
$$

None of these logical arrows can be reversed. The lower horizontal arrow has sense only for operator C*-algebras.

2. Derived functors and cohomology groups ("higher-dimensional theory")

2.1. *Complexes of topological modules and their (co)homology*

We begin with the following observation, which appears to be extremely useful in "higher-dimensional" topological homology. It is based on the Hahn–Banach theorem and the Banach open mapping theorem.

THEOREM 2.1.1 (cf. [109,132,89]). *Let $\mathcal{X}$ be a sequence of Banach spaces and continuous linear operators, and let $\mathcal{X}^*$ be the sequence, consisting of its dual spaces and operators. Then both sequences are simultaneously either exact or non-exact.*

The theorem is still valid for Fréchet spaces but ceases to be true for more general topological vector spaces. On the other hand, we can not replace the dual spaces $(\cdot)^*$ by the more general spaces $\mathcal{B}(\cdot, E)$ with $E \neq \mathbf{C}$; the reason is the existence in **Ban** of noncomplemented subspaces.

Let $\mathcal{K}$ be any of the categories, which we can come across in our exposition. According to what was said in the end of the Introduction, we always have a right to speak about the category of ⟨ chain | cochain ⟩ complexes over $\mathcal{K}$ and their morphisms; let us denote this category by $\langle \underline{\mathcal{K}} | \overline{\mathcal{K}} \rangle$. (In particular, we shall use notations like $\underline{\textbf{A-mod}}$ etc.) Every ⟨ co | contra ⟩variant functor $F : \mathcal{K}_1 \to \mathcal{K}_2$ between two such categories generates, in an obvious way, a ⟨ co | contra ⟩variant functor from $\mathcal{K}_1$ to $\langle \underline{\mathcal{K}_2} | \overline{\mathcal{K}_2} \rangle$; this functor is denoted by $\langle \underline{F} | \overline{F} \rangle$.

As an important example, let $(^*)$ be so-called *star functor* of the transfer to dual modules and morphisms. It can be defined on various categories of ⟨ Banach | topological ⟩ modules and takes values in corresponding categories of ⟨Banach | pure algebraic⟩ modules (cf. a particular case in 0.4). For a given complex $\mathcal{X}$ in the initial category, the complex $\overline{(^*)}\mathcal{X}$ is denoted more briefly by $\mathcal{X}^*$ and we call it the *dual complex to $\mathcal{X}$*.

The reader certainly knows what the functor of n-dimensional homology $H_n : \underline{\mathbf{Lin}} \to \mathbf{Lin}$, $n \in \mathbf{Z}$ is. Let us assume that a given complex $\mathcal{X} \in \underline{\mathbf{Lin}}$ is in fact a complex in the category **TVS** (see Introduction). Then $H_n(\mathcal{X})$, considered with the quotient topology, itself is an object of **TVS**. Moreover, if a morphism $\underline{\varphi} : \mathcal{X} \to \mathcal{Y}$ in $\underline{\mathbf{Lin}}$ is in fact a morphism in $\underline{\mathbf{TVS}}$, the operator $H_n(\varphi): H_n(\mathcal{X}) \to H_n(\mathcal{Y})$ is obviously continuous. Thus we can speak about the (similarly called and denoted) *n-dimensional homology functor* $H_n: \underline{\mathbf{TVS}} \to$ **TVS**. In the same way, dealing with cochain complexes instead of chain complexes, we come to the *n-dimensional cohomology functor* $H^n: \overline{\mathbf{TVS}} \to \mathbf{TVS}$. Note that for a complex in **Ban** its n-dimensional (co)homology is, for all n, a complete prenormed space; it is not bound to be Hausdorff (= normed) because the images of the relevant differentials are not bound to be closed. Thus we can speak about functors $H_n: \underline{\mathbf{Ban}} \to \langle \mathbf{Ban} \rangle$ and $H^n: \overline{\mathbf{Ban}} \to \langle \mathbf{Ban} \rangle$ (cf. Introduction).

Now we proceed to the famous "main lemma of homological algebra". In the functional-analytic context it acquires some topological additions which enable it to play the same outstanding role in topological, as well as in pure homology.

THEOREM 2.1.2. *Let* $0 \longleftarrow \mathcal{X} \xleftarrow{\varphi} \mathcal{Y} \xleftarrow{\psi} \mathcal{Z} \longleftarrow 0$ $(\underline{\mathcal{S}})$ *be a short exact sequence of chain complexes of topological vector spaces. Then there are linear operators* $\xi_n: H_n(\mathcal{X}) \to H_{n-1}(\mathcal{Z})$, $n \in \mathbf{Z}$, *such that the sequence*

$$\cdots \longleftarrow H_{n-1}(\mathcal{Z}) \xleftarrow{\xi_n} H_n(\mathcal{X}) \xleftarrow{H_n(\varphi)} H_n(\mathcal{Y}) \xleftarrow{H_n(\psi)} H_n(\mathcal{Z}) \xleftarrow{\xi_{n+1}} \cdots \quad \left(H\left(\underline{\mathcal{S}}\right)\right)$$

is exact. If, in addition, we are talking about Fréchet spaces, in particular, Banach spaces, then these operators are continuous.

$H\left(\underline{\mathcal{S}}\right)$ is called *the long exact homology sequence associated with* $\underline{\mathcal{S}}$, and the ξ_n are called the *connecting operators*.

With the help of this theorem and some devices, based on the Banach open mapping theorem, one can obtain

PROPOSITION 2.1.1 [94]. *Suppose that an operator in* $H(\underline{\mathcal{S}})$ *has a zero operator as its neighbour on the right. Then it is open. Moreover, if some operator has zeros as both of its neighbours, then it is a topological isomorphism.*

COROLLARY 2.1.1. *If* $\underline{\mathcal{S}}$ *consists of complexes of Fréchet spaces and one of its component complexes is exact, then all the nontrivial morphisms in* $H(\underline{\mathcal{S}})$ *are topological isomorphisms.*

We spoke about chain complexes; the "cochain" versions of these statements could be formulated in an obvious way. Besides, the statements could be formulated for modules instead of just spaces; see, e.g., [89].

2.2. *Resolutions*

2.2.1. *General definitions and standard resolutions.* Consider a certain object X in some reasonable category of topological modules. Be it projective or not, let us try to "express it by means of projective modules" with the help of the following procedure. At first we try to somehow represent X as a quotient module of some projective module so that the corresponding quotient map is an admissible morphism. Then, if successful, we represent the kernel of the latter morphism, in turn, as a quotient module of another projective module, again with an admissible quotient morphism; after this we do the same with the kernel of the latter and so on. The concept of a resolution that we are talking about provides a convenient "synchronous" way of writing down such a process. Here are the formal definitions.

We agree to identify a given module X with a chain, as well as with a cochain, complex in which it itself stands at the zero place and there are zeros at the others. Accordingly, if a $\langle$ chain | cochain $\rangle$ complex $\mathcal{Y} = \{Y, \langle d|\delta \rangle\}$ in our category is positive (that is $\langle Y_n | Y^n \rangle = 0$ when $n < 0$), then giving a morphism of complexes $\langle \varepsilon : \mathcal{Y} \to X | \eta : X \to \mathcal{Y} \rangle$ is equivalent to giving a complex

$$\left\langle 0 \longleftarrow X \xleftarrow{\varepsilon} Y_0 \xleftarrow{d_0} Y_1 \xleftarrow{d_1} \cdots \,\middle|\, 0 \longrightarrow X \xrightarrow{\eta} Y^0 \xrightarrow{\delta^0} Y^1 \xrightarrow{\delta^1} \cdots \right\rangle$$

briefly denoted by $\langle 0 \leftarrow X \leftarrow \mathcal{Y} | 0 \to X \to \mathcal{Y} \rangle$. In this case we say that we are given a *complex* $\langle$ *over* | *under* $\rangle$ X. The *length* of such a complex is the smallest n such that $\langle Y_k \mid Y^k \rangle = 0$ for $k > n$ or ∞ if there is no such n.

DEFINITION 2.2.1. A complex $\langle \varepsilon : \mathcal{Y} \to X$ over $X \mid \eta : X \to \mathcal{Y}$ under $X \rangle$ is called a $\langle$ resolution | coresolution $\rangle$ of the module X if the complex $\langle 0 \leftarrow X \leftarrow \mathcal{Y} | 0 \to X \to \mathcal{Y} \rangle$ is exact and admissible (see the beginning of Section 1.1.1).

In the context of Banach structures, we similarly define a *preresolution* of the given module, replacing the condition of the admissibility of $0 \leftarrow X \leftarrow \mathcal{Y}$ by the more liberal condition of the preadmissibility (see idem).

DEFINITION 2.2.2. A resolution $\varepsilon : \mathcal{P} \to X$ of a module X is called a projective resolution if all the modules in $\mathcal{P}$ are projective.

Similarly, according to the properties of the participating modules, we define *injective coresolutions, flat preresolutions, free and cofree* (*co*)*resolutions* etc.

It is clear that every projective module P has a projective resolution of length 0, namely

$$0 \leftarrow P \xleftarrow{\mathbf{1}} P \leftarrow 0$$

(similarly, injective modules have injective resolutions of length 0, and so on). These are certainly the simplest examples of resolutions; the next degree of complexity can be illustrated as follows.

EXAMPLE 2.2.1. Let Ω be a metrizable compact space and Δ be a closed subset of it. Consider the Banach left $\mathrm{C}(\Omega)$-module $\mathrm{C}(\Delta)$ with the pointwise outer multiplication and take the complex

$$0 \longleftarrow \mathrm{C}(\Delta) \xleftarrow{\tau} \mathrm{C}(\Omega) \xleftarrow{i} I \longleftarrow 0,$$

where $I = \{a \in C(\Omega)\colon a|_\Delta = 0\}$, τ sends functions to their restrictions to Δ, and i is the natural embedding. Since I is projective (Theorem 1.1.1) and complementable in $C(\Omega)$, this complex is a projective resolution of $C(\Delta) \in C(\Omega)$-**mod**, and its length is 1. More generally, let A be a Fréchet algebra, let X be a cyclic Fréchet left A-module with a cyclic vector x, and suppose that $I = \{a \in A_+\colon a \cdot x = 0\}$ is projective and complemented in A_+. Then X has a projective resolution (in the category of Fréchet left modules) of length 1, namely, the complex

$$0 \longleftarrow X \xleftarrow{\tau} A_+ \xleftarrow{i} I \longleftarrow 0,$$

where τ sends a to $a \cdot x$ and i is the natural embedding.

Now we show that, despite the possible abundance of projective resolutions for a given module, in essence they are very close to each other.

THEOREM 2.2.1 ("Comparison theorem"). *Let $\varepsilon : \mathcal{P} \to X$ be a projective resolution and $\varepsilon' : \mathcal{Q} \to Y$ be an arbitrary resolution of the relevant modules. Then for any morphism $\varphi : X \to Y$ there exists a morphism $\underline{\varphi}$ between complexes such the diagram*

$$\begin{tikzcd}
\mathcal{P} \arrow[r, "\varepsilon"] \arrow[d, "\underline{\varphi}"'] & X \arrow[d, "\varphi"] \\
\mathcal{Q} \arrow[r, "\varepsilon'"] & Y
\end{tikzcd}$$

is commutative. Any two morphisms with this property are homotopic.

The obvious analogue of this theorem is valid for injective coresolutions (both assertions explain why we are interesting in projective resolutions and injective coresolutions and not, say, in projective coresolutions or injective resolutions).

COROLLARY 2.2.1. *Any two $\langle$ projective resolutions | injective coresolutions $\rangle$ are homotopically equivalent.*

Does every object in a given category of topological modules possess, say, a projective resolution? We shall show that the answer is certainly positive for those categories, which have appropriate notions of a free module and a canonical projection (cf. 0.2 and 1.1.1). In such categories there exist several rather similar methods, which lead to several versions of the so-called standard, or bar-, resolutions.

Let us take, to be definite, a Banach algebra A and an arbitrary Banach left A-module X. Consider the commutative diagram

$$\begin{tikzcd}
0 & X \arrow[l] & A_+ \widehat{\otimes} X \arrow[l, "\pi"'] & & A_+ \widehat{\otimes} K_0 \arrow[ll, "d_0'"'] \arrow[dl, "\pi_0"] & & A_+ \widehat{\otimes} K_1 \arrow[ll, "d_1'"'] \arrow[dl, "\pi_1"] & & \cdots \arrow[ll] \\
& & & K_0 \arrow[ul, "i_0"] & & K_1 \arrow[ul, "i_1"] & & \arrow[ul] &
\end{tikzcd}$$

in which π and π_n, $n = 0, 1, \ldots$, are canonical projections, the i_n are natural embeddings, K_0 is $\operatorname{Ker}\pi$ and the K_n, $n > 0$, are defined inductively as $\operatorname{Ker}\pi_{n-1}$. It is easy to see that the upper row of this diagram is a projective resolution of X in **A-mod**; it is denoted by $0 \longleftarrow X \xleftarrow{\pi} \mathbf{B}'(X)$.

The same resolution, which was just constructed by induction, has also a convenient "simultaneous" description. Consider, for the same X, the free modules

$$\begin{aligned}\mathbf{B}_0(X) &= A_+ \widehat{\otimes} X,\\ \mathbf{B}_1(X) &= A_+ \widehat{\otimes} (A \widehat{\otimes} X), \ldots, \mathbf{B}_n(X) = A_+ \widehat{\otimes} (\underbrace{A \widehat{\otimes} \cdots \widehat{\otimes} A}_{n} \widehat{\otimes} X)\end{aligned}$$

and operators $d_n : \mathbf{B}_{n+1}(X) \to \mathbf{B}_n(X)$, well-defined by

$$\begin{aligned}&d_n(a \otimes a_1 \otimes \cdots \otimes a_{n+1} \otimes x)\\ &\quad = aa_1 \otimes a_2 \otimes \cdots \otimes a_{n+1} \otimes x\\ &\quad + \sum_{k=1}^{n} (-1)^k a \otimes a_1 \otimes \cdots \otimes a_k a_{k+1} \otimes \cdots \otimes a_{n+1} \otimes x\\ &\quad + (-1)^{n+1} a \otimes a_1 \otimes \cdots \otimes a_n \otimes a_{n+1} \cdot x.\end{aligned} \tag{2.1}$$

PROPOSITION 2.2.1. *The sequence*

$$0 \longleftarrow X \xleftarrow{\pi} \mathbf{B}_0(X) \xleftarrow{d_0} \mathbf{B}_1(X) \longleftarrow \cdots \qquad \big(0 \longleftarrow X \xleftarrow{\pi} \mathbf{B}(X)\big)$$

is a projective (*in fact, free*) *resolution of the module X, and it is isomorphic, as a complex of Banach left A-modules, to* $0 \longleftarrow X \xleftarrow{\pi} \mathbf{B}'(X)$.

DEFINITION 2.2.3. The complex $0 \longleftarrow X \xleftarrow{\pi} \mathbf{B}(X)$, as well as its isomorphic copy $0 \longleftarrow X \xleftarrow{\pi} \mathbf{B}'(X)$, is called the normalized standard, or normalized bar-, resolution of the Banach A-module X.

When speaking about the standard resolution of X, we shall often mean the complex $\mathbf{B}(X)$ or $\mathbf{B}'(X)$ itself (with X removed); it will not create a misunderstanding.

Another version of the standard resolution can be provided by the following assertion. Consider the free Banach left modules

$$\beta_n(X) = A_+ \widehat{\otimes} (\underbrace{A_+ \widehat{\otimes} \cdots \widehat{\otimes} A_+}_{n} \widehat{\otimes} X), \qquad n = 0, 1, \ldots.$$

PROPOSITION 2.2.2. *For $X \in$* **A-mod** *there exists a unique free and hence projective resolution of the form*

$$0 \longleftarrow X \xleftarrow{\pi} \beta_0(X) \xleftarrow{\partial_0} \beta_1(X) \xleftarrow{\partial_1} \cdots \qquad \big(0 \longleftarrow X \xleftarrow{\pi} \beta(X)\big)$$

such that the family of operators $\zeta_{-1}: X \to \beta_0(X)$: $x \mapsto e \otimes x$ *and* $\zeta_n: \beta_n(X) \to \beta_{n+1}(X)$: $u \mapsto e \otimes u$, $n = 0, 1, \ldots$, *is its contracting homotopy as of a complex in* **Ban**. *Moreover, the differentials* ∂_n *are defined by the same formulae* (2.1) *as* d_n *in* $\mathbf{B}(X)$.

This resolution is called *non-normalized standard*, or *non-normalized bar-*, (and also *simplicial*) *resolution* of X.

If A and X are unital, then the versions of the bar-resolution can be defined without involving the unitization algebra A_+. The "normalized" version can be defined as the complex

$$0 \longleftarrow X \xleftarrow{\pi} A \widehat{\otimes} X \xleftarrow{d_0''} \cdots \longleftarrow A \widehat{\otimes} (\underbrace{\overline{A} \widehat{\otimes} \cdots \widehat{\otimes} \overline{A}}_{n} \widehat{\otimes} X) \xleftarrow{d_n''} \cdots,$$

where $\overline{A}$ denotes the quotient space $A/\{\lambda e\colon \lambda \in \mathbf{C}\}$, and the "non-normalized" version can be defined as the similar complex with A instead of $\overline{A}$. The differentials are given by the same formulae (2.1); the only modification is that $d_n''(a \otimes \overline{a}_1 \otimes \cdots \overline{a}_{n+1} \otimes x)$ is well-defined as $d_n(a \otimes a_1 \otimes \cdots \otimes a_{n+1} \otimes x)$; here $\overline{a}_1, \ldots, \overline{a}_{n+1} \in \overline{A}$, and a_k is an arbitrary element in the coset $\overline{a}_k$, $k = 1, \ldots, n+1$.

As a matter of fact, all the mentioned constructions are functorial, that is, they can be extended from objects of **A-mod** to morphisms of this category. Thus the *functor* $\langle \mathbf{B} | \beta \rangle$: **A-mod** $\to$ <u>**A-mod**</u> *of* ⟨ *normalized* | *non-normalized* ⟩ *bar-resolution*, as well as its "unital" version arises. Details can be found in [89].

The respective versions of the standard resolution for Banach ⟨ right | two-sided ⟩ modules can be defined in an obvious way by the identification of these modules with Banach left modules over the ⟨ opposite | enveloping ⟩ Banach algebra (see 0.1). Similar constructions manifestly exist for Fréchet, $\widehat{\otimes}$- and $\overline{\otimes}$-modules of arbitrary types. Further, the notion of a cofree module, which exists in the case of Banach structures (see 1.1.2), enables one to construct, on rather parallel lines, several versions of the so-called standard injective (in fact, cofree) coresolution of a given Banach module (cf. [89]). Thus we have

THEOREM 2.2.2. *In the categories of Banach, Fréchet,* $\widehat{\otimes}$*- and* $\overline{\otimes}$*-modules* (of an arbitrary type) *every object has projective resolutions. In the categories of Banach modules* (of an arbitrary type) *every object has, in addition, injective coresolutions.*

As it was mentioned in 1.1.2, the notion of a cofree module, generally speaking, looses its sense in the case of non-Banach topological modules. Therefore questions concerning the existence of injective coresolutions of such modules become to be rather complicated (cf. [209]). In fact, we still do not know whether Fréchet modules have at least one injective coresolution.[10]

2.2.2. *Some special resolutions.* The standard (co)resolutions are good in that they exist "under all regimes". Nevertheless, in many cases, when working with concrete algebras

[10] *Added in proof*: they are not bound to have; see the footnote to p. 171.

and/or modules, some other (co)resolutions turn out to be more useful; these do not resemble the standard resolutions and reflect the peculiarities of the given situation. We have already come across a couple of such examples in the previous subsection; now we indicate some more important special (co)resolutions.

At first let us mention several devices, which enable one to pass from one (co)resolution to another. We begin with the obvious

PROPOSITION 2.2.3. *Let X be a Banach module (of an arbitrary type) over a Banach algebra A. Then the complex $\varepsilon : \mathcal{X} \to X$ over X is a flat preresolution iff the dual complex $\varepsilon^* : X^* \to \mathcal{X}^*$ is an injective coresolution of the module X^* in the respective category.*

The obvious topological version of the algebraic tensor product of complexes over/under modules (see, e.g., [89]) is also useful in building (co)resolutions. If A is, say, a $\widehat{\otimes}$-algebra, and $\langle \varepsilon_1 : \mathcal{P} \to X \mid \varepsilon_2 : \mathcal{Q} \to Y \rangle$ is a projective resolution of a $\langle$ left $\mid$ right $\rangle$ $\widehat{\otimes}$-module $\langle X \mid Y \rangle$ over A, then their (projective) tensor product $\varepsilon_1 \mathbin{\widehat{\otimes}} \varepsilon_2 : \mathcal{P} \mathbin{\widehat{\otimes}} \mathcal{Q} \to X \mathbin{\widehat{\otimes}} Y$ is a projective resolution of $X \mathbin{\widehat{\otimes}} Y$ in the category of $\widehat{\otimes}$-bimodules over A. Moreover, if X_k is a left $\widehat{\otimes}$-module over an $\widehat{\otimes}$-algebra A_k, $k = 1, \ldots, n$, and $\varepsilon_k : \mathcal{P}_k \to X_k$ is its projective resolution, then

$$\varepsilon_1 \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} \varepsilon_n : \mathcal{P}_1 \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} \mathcal{P}_n \to X_1 \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} X_n$$

is the projective resolution of $X_1 \mathbin{\widehat{\otimes}} \ldots \mathbin{\widehat{\otimes}} X_n$ in the category of left $\widehat{\otimes}$-modules over $A_1 \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} A_n$. In the case of Banach structure both facts remain true if we replace projective by flat resolutions.

EXAMPLE 2.2.2. Let Ω_k, $k = 1, \ldots, n$, be a compact metrizable space. Consider the well known "Varopoulos algebra" $A = C(\Omega_1) \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} C(\Omega_n)$ (cf. [214]). Since $\underline{\Omega}(A) = \Omega_1 \times \cdots \times \Omega_n$, every one-dimensional A-module has the form $\mathbf{C}_{t_1} \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} \mathbf{C}_{t_n}$ where $t_k \in \Omega_k$ (cf. Example 0.2.1). As a particular case of Example 2.2.1, $\mathbf{C}_{t_k}$, $k = 1, \ldots, n$, has a projective resolution of length 1. Therefore, taking the (projective) tensor product of respective complexes, we see that all one-dimensional modules over such an A have projective resolutions of length n.

* * *

Now for an arbitrary $\widehat{\otimes}$-, or $\overline{\otimes}$-algebra A, we turn to the extremely important A-bimodule A_+. Let us substitute A_+ for X in the standard resolutions $0 \longleftarrow X \xleftarrow{\pi} \mathbf{B}(X)$ and $0 \longleftarrow X \xleftarrow{\pi} \beta(X)$. It is easy to observe that the resulting complexes $0 \longleftarrow A_+ \xleftarrow{\pi} \mathbf{B}(A_+)$ and $0 \longleftarrow A_+ \xleftarrow{\pi} \beta(A_+)$ happen to be projective – in fact, free – resolutions of A_+ in the category of $\widehat{\otimes}$- (or $\overline{\otimes}$-) *bimodules* over A (and not only in that of left modules). They are so-called *normalized* and, respectively, *non-normalized bimodule bar-resolutions*. Note that their names and notations should not lead to error. They are not the standard resolutions in the formal meaning of the term, that is, not the standard resolutions of the left A^{env}-module, which corresponds to A_+, if we identify A-bimodules with left A^{env}-modules.

The following proposition is somewhat less straightforward. For a Banach algebra A, denote by $(A^n)^*$ the Banach space of continuous n-linear functionals on A; it is a Banach A-bimodule with respect to outer multiplications

$$[a \cdot f](a_1, \dots, a_n) = f(a_1, \dots, a_{n-1}, a_n a)$$

and

$$[f \cdot a](a_1, \dots, a_n) = f(aa_1, a_2, \dots, a_n).$$

Consider also the operators $\zeta : A^* \to (A^2)^*$ and $\zeta^n : (A^{n+2})^* \to (A^{n+3})^*$ defined by

$$\zeta f(a, b) = f(ab)$$

and

$$\zeta^n f(a_1, \dots, a_{n+3}) = \sum_{k=1}^{n+2} (-1)^{k+1} f(a_1, \dots, a_k a_{k+1}, \dots, a_{n+3}).$$

It is obvious that the sequence

$$0 \to A^* \xrightarrow{\zeta} (A^2)^* \xrightarrow{\zeta^0} \cdots \to (A^{n+2})^* \xrightarrow{\zeta^n} (A^{n+3})^* \to \cdots \quad \left(0 \to A^* \xrightarrow{\zeta} \tilde{\beta}^*(A)\right)$$

is a complex in **A-mod-A**, dual to the complex $0 \longleftarrow A \xleftarrow{\pi} \tilde{\beta}(A)$, obtained by substituting A for A_+ in $0 \longleftarrow A_+ \xleftarrow{\pi} \beta(A_+)$,

PROPOSITION 2.2.4. *Let A be a Banach algebra, possessing a b.a.i. Then the complex $0 \leftarrow A \xleftarrow{\pi} \tilde{\beta}(A)$, is a flat preresolution of A in* **A-mod-A**, *and, equivalently,* $0 \to A^* \xrightarrow{\pi^*} \tilde{\beta}^*(A)$, *is an injective coresolution of A in* **A-mod-A**.

In the special case of an operator C*-algebra A, acting on a Hilbert space, let us concentrate on the important Banach A-bimodule A_*^- (see 2.4.2). This bimodule has several similarly constructed injective coresolutions. Three of them, which will happen to be rather useful (see 2.4.5 below) are defined as follows. Note that the complex $0 \longrightarrow A^* \xrightarrow{\pi^*} \tilde{\beta}^*(A)$ has, for such an A, the well-defined subcomplex consisting of bimodules, formed by those polylinear functionals on A, which are normal (see 0.3) on $\langle$ all variables | the first variable | the last variable $\rangle$; we denote it by

$$\left\langle 0 \longrightarrow A_*^- \xrightarrow{\pi_*} \tilde{\beta}_*(A) \,\middle|\, 0 \longrightarrow A_*^- \xrightarrow{\pi_l} \tilde{\beta}_l(A) \,\middle|\, 0 \longrightarrow A_*^- \xrightarrow{\pi_r} \tilde{\beta}_r(A) \right\rangle.$$

PROPOSITION 2.2.5. *For an operator* C*-*algebra A, every one of the just defined three complexes is an injective coresolution of A_*^- in* **A-mod-A**.

* * *

To conclude this subsection, we consider a special resolution of general modules, having length 2. It turns out to be useful in the computation and estimation of homological dimensions of many topological, especially Banach, algebras (see 2.4.5 below). As a rule, this resolution is effective when our algebra is not unital or, which is essentially the same, it has a non-unital two-sided ideal of codimension 1.

To be definite, suppose that A is an $\widehat{\otimes}$-algebra, and X is a left $\widehat{\otimes}$-module over A. (The case of $\overline{\otimes}$-algebras/modules can be treated similarly.) Besides X, we consider the reduced left A-module $X_{\sqcap} = A \widehat{\otimes}_A X$ (see 1.3.3), and we put $P = (A_+ \widehat{\otimes} X) \oplus X_{\sqcap}$ and $Q = (A_+ \widehat{\otimes} X_{\sqcap}) \oplus (A \widehat{\otimes} X)$. Now introduce the morphisms $\langle \partial_{-1} : P \to X \mid \partial_0 : Q \to P \mid \Delta : A \widehat{\otimes} X_{\sqcap} \to Q \rangle$, given "componentwise" with the help of matrices

$$\left\langle (\pi_X, \kappa) \,\middle|\, \begin{bmatrix} \mathbf{1} \otimes \kappa & -i \otimes \mathbf{1} \\ -\pi_{X_{\sqcap}} & \pi' \end{bmatrix} \,\middle|\, (i \otimes \mathbf{1}, \mathbf{1} \otimes \kappa) \right\rangle,$$

where π_X and $\pi_{X_{\sqcap}}$ are the respective canonical projections, κ takes $a \widehat{\otimes}_A x$ to $a \cdot x$, π' takes $a \widehat{\otimes} x$ to $a \widehat{\otimes}_A x$, and i is the natural embedding.

THEOREM 2.2.3 (cf. [84,89]). *The complex*

$$0 \longleftarrow X \xleftarrow{\partial_{-1}} P \xleftarrow{\partial_0} Q \xleftarrow{\Delta} A \widehat{\otimes} X_{\sqcap} \longleftarrow 0 \qquad (0 \leftarrow X \leftarrow \operatorname{Ent} X)$$

is a resolution of the A-module X (that is, it is exact and admissible). This resolution is projective iff the left A-modules A and $X_{\sqcap}$ are both projective; in particular (cf. Subsection 1.3.3), *it is certainly projective, provided A is a biprojective algebra. Finally, in the case of Banach structures this resolution is flat iff the left A-modules A and $X_{\sqcap}$ are flat; in particular* (cf. idem), *it is certainly flat provided A is a biflat algebra.*

The indicated resolution is called the *entwining resolution*. (It works as if it "entwines" the module $A_+ \widehat{\otimes} X$ and reduces it, by "throwing out" $X_{\sqcap}$, to the more manageable module $A \widehat{\otimes} X_{\sqcap}$.)

For the case $X = A_+$, the ingredients of the entwining resolution acquire more transparent form, so we have a resolution

$$0 \leftarrow A_+ \xleftarrow{\partial_{-1}} \big[(A_+ \widehat{\otimes} A_+) \oplus A\big] \xleftarrow{\partial_0} \big[(A_+ \widehat{\otimes} A) \oplus (A \widehat{\otimes} A_+)\big] \xleftarrow{\partial_1} A \widehat{\otimes} A \leftarrow 0$$

which is a complex of A-bimodules. Moreover, the following is evident:

PROPOSITION 2.2.6. *The resolution $0 \longleftarrow A_+ \longleftarrow \operatorname{Ent} A_+$ is projective iff A is a biprojective algebra. In the case of a Banach algebra A this resolution is flat iff A is a biflat algebra.*

COROLLARY 2.2.2. *If A is a $\langle$ biprojective $\widehat{\otimes}$-algebra $\mid$ biflat Banach algebra $\rangle$ then the A-bimodule A_+ has a $\langle$ projective $\mid$ flat $\rangle$ resolution of length 2.*

See [4,148] for some modifications of the entwining resolution, which are useful in the work with some classes of C*-algebras.

2.2.3. *Koszul resolution and Taylor spectrum.* One distinguished type of special resolutions deserves to be discussed at a rather considerable length. It is so-called the Koszul resolution, a term borrowed from pure homology (cf. [134]). Such a resolution is typical for modules over a several important topological algebras, which are, as a rule, non-Banach. Their class includes, in particular, most algebras of test functions and distributions, occurring in distribution theory.

Let E be a (for the moment arbitrary) linear space, and let $\mathcal{S} = (S_1, \dots, S_n)$ be an n-tuple of commuting operators on E. For $m = 0, 1, \dots, n$ we put $E_m = E \otimes (\bigwedge^m \mathbf{C}^n)$; this means that E_m is the direct sum of $\binom{n}{m}$ copies of the space E enumerated by the multi-indices $(i_1, \dots, i_m)$; $1 \leqslant i_1 < \dots < i_m \leqslant n$ and therefore it is generated by elements of the form $x \otimes e_{i_1} \wedge \dots \wedge e_{i_m}$, where $x \in E$, and $e_1, \dots, e_n$ is the natural basis in $\mathbf{C}^n$. We note that E_0 and E_n are naturally identified with E.

For each $m = 0, \dots, n-1$ we introduce the operator $d_m : E_{m+1} \to E_m$ well-defined by

$$d_m(x \otimes e_{i_0} \wedge \dots \wedge e_{i_m}) = \sum_{l=0}^{m} (-1)^l S_{i_l} x \otimes e_{i_0} \wedge \dots \wedge \hat{e}_{i_l} \wedge \dots \wedge e_{i_m}.$$

One can show by an elementary calculation that $d_{m-1} d_m = 0$ for all m.

We let D_m, $m > 0$, denote the subspace $\operatorname{Im} S_1 + \dots + \operatorname{Im} S_n$ in E, and we put $D_0 = \{0\}$. We also introduce the quotient space $E_{-1} = E/D_n$ and the quotient map $\varepsilon : E \to E_{-1}$.

DEFINITION 2.2.4. The complex

$$0 \longleftarrow E \xleftarrow{d_0} E_1 \xleftarrow{d_1} \cdots \xleftarrow{d_{n-1}} E_n \longleftarrow 0$$

(denoted by $\mathcal{K}(E, \mathcal{S})$ or $\mathcal{K}(E, S_1, \dots, S_n)$) is called the Koszul complex of the pair $(E, \mathcal{S})$.The complex $0 \leftarrow E_{-1} \xleftarrow{\varepsilon} \mathcal{K}(E, \mathcal{S})$ (denoted by $\mathcal{K}^+(E, \mathcal{S})$ or $\mathcal{K}^+(E, S_1, \dots, S_n)$) is called the augmented Koszul complex of the pair $(E, \mathcal{S})$.

The fact that the operators $S_1, \dots, S_n$ commute implies that D_m is an invariant subspace for S_{m+1}; therefore the latter operator generates an operator $\overline{S}_{m+1} : E/D_m \to E/D_m$.

Now assume that E is a topological vector space and the operators in $\mathcal{S}$ are continuous; then, obviously, all the operators introduced up to now are also continuous.

PROPOSITION 2.2.7 [208]. *Let E and $\mathcal{S}$ be such that for any $m = 0, \dots, n-1$ the operator $\overline{S}_{m+1}$ has a (continuous) left inverse. Then the augmented Koszul complex $\mathcal{K}^+(E, \mathcal{S})$ splits as a complex of topological vector spaces.*

This proposition is a useful tool to prove that in certain situations the Koszul complex provides projective resolutions for some modules, mostly belonging to the class of reduced modules (see 1.3.3). Note that reduced modules over a unital algebra are obviously just unital modules.

In what follows, A is a $\widehat{\otimes}$- or $\overline{\otimes}$-algebra, in particular, Fréchet algebra. To be definite, we shall use the symbol $\widehat{\otimes}$, although all that is said remains valid if we replace it by $\overline{\otimes}$. Let

$\varphi_m : A \to A,\ m = 1, \ldots, n$, be a system of mutually commuting morphisms of A as of a topological A-bimodule, and let $X = A \widehat{\otimes}_A Y$ be a (reduced) topological left A-module. Put $\psi_m = \varphi_m \widehat{\otimes}_A \mathbf{1}_Y : X \to X$ and $\Sigma_m = \mathbf{1}_A \widehat{\otimes} \psi_m - \varphi_m \widehat{\otimes} \mathbf{1}_X$; it is clear that $\Sigma_m,\ m = 1, \ldots, n$ is an n-tuple of mutually commuting morphisms of topological left A-modules. Consider the complex

$$0 \longleftarrow X \xleftarrow{\pi} \mathcal{K}(A \widehat{\otimes} X, \Sigma_1, \ldots, \Sigma_n), \tag{2.2}$$

where $\pi : A \widehat{\otimes} X \to X$ (like the canonical projection) assigns $a \cdot x$ to $a \otimes x$.

If A is unital, the situation becomes more transparent: in this case φ_m necessarily acts as $a \mapsto aa_m$ for some $a_m = \varphi_m(e) \in A$, and Σ_m acts as $a \otimes x \mapsto a \otimes a_m \cdot x - aa_m \otimes x$.

On the other hand, if $X = A$, and A resembles unital algebras in the sense that $A = A \widehat{\otimes}_A A$, then the complex (2.2) takes the form

$$0 \longleftarrow A \xleftarrow{\pi} \mathcal{K}(A \widehat{\otimes} A, \Sigma_1, \ldots, \Sigma_n) \tag{2.3}$$

with $\Sigma_m = \mathbf{1}_A \otimes \varphi_m - \varphi_m \otimes \mathbf{1}_A$. For a unital A this means that Σ_m acts as $u \mapsto u(e \otimes a_m - a_m \otimes e)$ for some $a_m \in A$. Note that (2.3) is a complex of topological A-bimodules.

In the pure algebraic context, if we take the algebra of complex polynomials in variables $t_1, \ldots, t_n$ as A and a_m as t_m, then the complex (2.2) turns out to be the well-known Koszul resolution of modules over this classical algebra (discovered, in somewhat more general setting, in [134]). The following assertion provides versions of such a resolution for some topological algebras, which are, in a sense, close to the polynomial algebra. As to Taylor's concept of a localization, which explains, at least for unital algebras, the exact meaning of this "closeness", see [208].

The structure of the topological algebras in the cases (i) and (ii) below (these appear to be the most important in the list), was already discussed in the Introduction. As to others see, e.g., [208,150]. here we only mention, that the algebra in the case (v) is the only one in the list which has the convolution, and not the pointwise, multiplication, and that its topology is the strong topology of a dual space.

THEOREM 2.2.4 (cf. [208,156,155]). *In each of the following cases the complex* (2.2) *is a projective resolution of a reduced topological A-module X:*

(i) *A is the Fréchet algebra $\mathcal{O}(\mathcal{U})$ of holomorphic functions on a polydomain $\mathcal{U} = \mathcal{U}_1 \times \cdots \times \mathcal{U}_n \subseteq \mathbf{C}^n$.*
(ii) *A is the Fréchet algebra $C^\infty(\mathcal{U})$ of all infinitely smooth functions on a domain $\mathcal{U} \subseteq \mathbf{R}^n$.*
(iii) *A is the $\widehat{\otimes}$-algebra $\mathcal{D}(\mathcal{U})$ of compactly supported infinitely smooth functions on a domain $\mathcal{U} \subseteq \mathbf{R}^n$.*
(iv) *A is the Fréchet algebra $\mathcal{S}(\mathbf{R}^n)$ of rapidly decreasing infinitely smooth functions on $\mathbf{R}^n$.*
(v) *A is the $\widehat{\otimes}$-algebra $\mathcal{E}^*(\mathbf{R}^n)$ of compactly supported distributions in $\mathbf{R}^n$.*

In the cases (i)–(iv) *φ_m is taken as the morphism of the pointwise multiplication by the m-th coordinate function whereas in the last case φ_m is taken as $\frac{\partial}{\partial x_m}$.*

Recall that, in virtue of the Grothendieck's identification [71], for $A = \langle \mathcal{O}(\mathcal{U}) \mid C^\infty(\mathcal{U}) \rangle$ the module $A \mathbin{\widehat{\otimes}} X$ is topologically isomorphic to the module $\langle \mathcal{O}(\mathcal{U}, X) \mid C^\infty(\mathcal{U}, X) \rangle$ of $\langle$ holomorphic | infinitely smooth $\rangle$ X-valued functions on $\mathcal{U}$, with the outer multiplication $[a \cdot \overline{x}](t) = a(t)\overline{x}(t)$, where a and x are respectively a scalar-valued and a vector-valued function. The similar identification is valid also for $\mathcal{D}(\mathcal{U})$ and $\mathcal{S}(\mathbf{R}^n)$. Therefore the corresponding complex (2.2) acquires a rather simple form. For example, in the "holomorphic" case it can be written as

$$0 \longleftarrow X \xleftarrow{\pi} \mathcal{K}\big(\mathcal{O}(\mathcal{U}, X), S_k\big), \tag{2.4}$$

where $(S_k\overline{u})(z) = [z_k \cdot \overline{u}](z) - z_k\overline{u}(z)$, $\overline{u} \in \mathcal{O}(\mathcal{U}, X)$, and $\pi : \mathcal{O}(\mathcal{U}, X) \to X$ is well-defined by assigning $w \cdot x$ to $w(z)x$, $w \in \mathcal{O}(\mathcal{U})$, $x \in X$. The respective $\mathcal{O}(\mathcal{U})$-bimodule resolution of $\mathcal{O}(\mathcal{U})$ is

$$0 \longleftarrow \mathcal{O}(\mathcal{U}) \xleftarrow{\pi} \mathcal{K}\big(\mathcal{O}(\mathcal{U} \times \mathcal{U}), S_k\big), \tag{2.5}$$

where

$$[S_k w](z_1, \dots, z_n, \zeta_1, \dots, \zeta_n) = (z_k - \zeta_k) w(z_1, \dots, \zeta_n)$$

and π is the morphism of the restriction to the diagonal in $\mathcal{U} \times \mathcal{U}$. The resolutions, corresponding to the algebras in the cases (i)–(iv), acquire a similar form (Taylor [208]).

It $\mathcal{U}$ is a holomorphy domain in $\mathbf{C}^n$ which is, generally speaking, not a polydomain, the complex (2.5) is still exact (idem); however, it is not known whether it is admissible.

It is not by chance, that all algebras participating in the previous theorem, are non-Banach:

PROPOSITION 2.2.8 (idem). *Let A be a unital Banach algebra. Then the complex* (2.3) *is not* (even!) *exact, whatever system φ_m, $m = 1, \dots, n$* (see above) *we would take.*

REMARK. Side by side with the Koszul resolution of the polynomial algebra, another type of a special resolution is well known in pure algebra. These are resolutions of modules over the free algebra of n generators, and they always have the length 1. There exist several examples of topological algebras such that their modules have resolutions, constructed according to the same design. These algebras can be, with some justification, thought as "function algebras of n non-commuting variables". See [210] for the detailes.

* * *

Now we are already able, and therefore are greatly tempted, to formulate the main concept, as well as the main result of Taylor's multi-operator spectral theory; recall that the latter was mentioned in the Preface as one of principal benefactors of topological homology.

Let $\mathcal{T} = (T_1, \dots, T_n)$ be an n-tuple of mutually commuting bounded operators, acting on a Banach space E. Take a vector $\lambda = (\lambda_1, \dots, \lambda_n) \in \mathbf{C}^n$ and consider the Koszul complex $\mathcal{K}(E, \mathcal{T} - \lambda)$, where $\mathcal{T} - \lambda$ stands for the tuple $(T_1 - \lambda_1\mathbf{1}, \dots, T_n - \lambda_n\mathbf{1})$.

DEFINITION 2.2.5. The set of those $\lambda \in \mathbf{C}^n$, for which the complex $\mathcal{K}(E, \mathcal{T} - \lambda)$ is not exact (at least in one term) is called the (joint) Taylor spectrum of the system $\mathcal{T}$.

The Taylor spectrum is always a compact non-empty set, and in the case $n = 1$ it coincides with the classical spectrum of an operator. See [206] about its various (as a rule, very good) geometrical properties.

If $\mathcal{U}$ is a domain in $\mathbf{C}^n$, a *holomorphic calculus of the system $\mathcal{T}$ on $\mathcal{U}$* is, by definition, a unital continuous homomorphism $\Gamma : \mathcal{O}(\mathcal{U}) \to \mathcal{B}(E)$, satisfying the condition $\Gamma(z_k) = T_k$, $k = 1, \dots, n$. Such a definition gives a reasonably exact sense to the concept of "holomorphic function of operator variables". Indeed, if we are given a holomorphic calculus of $\mathcal{T}$ on $\mathcal{U}$, then, for any $w \in \mathcal{O}(\mathcal{U})$, we can talk about the "meaning of this function of $T_1, \dots, T_n$" by putting $w(T_1, \dots, T_n) = \Gamma(w)$.

A well known theorem of Gel'fand (see, e.g., [96]) states that, in the case of a single operator T and a domain $\mathcal{U} \subseteq \mathbf{C}$, the holomorphic calculus of $\mathcal{T} = \{T\}$ on $\mathcal{U}$ exists iff $\mathcal{U}$ contains the spectrum of T, and in this case it is uniquely determined. The search for the "right" multidimensional generalization of this theorem took many years and was not devoid of dramatic turns of events (cf. [126]). Now the following result appears to give the final solution of the problem.

THEOREM 2.2.5 ("Taylor theorem on holomorphic calculus"). *Let E and $\mathcal{T}$ be as in Definition 2.2.5. Then for any domain $\mathcal{U} \subseteq \mathbf{C}^n$ the holomorphic calculus of $\mathcal{T}$ on $\mathcal{U}$ exists if $\mathcal{U}$ contains the Taylor spectrum of $\mathcal{T}$. Moreover, in the case of a domain of holomorphy $\mathcal{U}$, this sufficient condition is also necessary, and the holomorphic calculus is uniquely determined.*

For the moment, let this formulation serve as an appetizer. Some comments and explanations will be appropriate later, after the appearance in 2.3.3 of the spaces Tor.

2.3. *Derived functors*

2.3.1. *General definitions and properties.* Consider a reasonable category of topological modules; let us denote it by **Mod**. In fact, it can be an arbitrary subcategory of the category of all left, right or two-sided modules over a fixed topological algebra; we assume only that every object in **Mod** has at least one projective resolution, belonging, as a complex, to $\underline{\mathbf{Mod}}$.

Let F be an additive covariant functor from **Mod** to some another category of the same kind. For most of the applications, it is sufficient to take, as the range category, just **TVS**; so, for the sake of transparency, let us fix it in this capacity (some modifications, when required, will be explicitly discussed; cf. also [89]).

The ("comparison") Theorem 2.2.1 implies

THEOREM 2.3.1. *Let X and Y be objects in* **Mod** *with projective resolutions, respectively $0 \longleftarrow X \xleftarrow{\varepsilon_k} \mathcal{P}_k$ and $0 \longleftarrow Y \xleftarrow{\xi_k} \mathcal{Q}_k$, $k = 1, 2$, and let $\varphi : X \to Y$ be a morphism. Fur-*

ther, let

$$\begin{tikzcd}
X \arrow[dr, equal] \arrow[dd, "\varphi"'] & & \mathcal{P}_1 \arrow[ll, "\varepsilon_1"'] \arrow[dr] \arrow[dd, "\underline{\varphi_1}"] & \\
& X \arrow[dd, "\varphi"] & & \mathcal{P}_2 \arrow[ll, "\varepsilon_2"'] \arrow[dd, "\underline{\varphi_2}"] \\
Y \arrow[dr, equal] & & \mathcal{Q}_1 \arrow[ll, "\xi_1"'] \arrow[dr] & \\
& Y & & \mathcal{Q}_2 \arrow[ll, "\xi_2"']
\end{tikzcd}$$

be a commutative diagram in $\underline{\mathbf{Mod}}$. *Then, applying* $\underline{F}:\underline{\mathbf{Mod}} \to \underline{\mathbf{TVS}}$ *and then* $H_n:\underline{\mathbf{TVS}} \to \mathbf{TVS}$ *to the right-hand face of this diagram, we get the commutative diagram*

$$\begin{tikzcd}
H_n(\underline{F}(\mathcal{P}_1)) \arrow[r, "\sim"] \arrow[d, "H_n(\underline{F}(\underline{\varphi_1}))"'] & H_n(\underline{F}(\mathcal{P}_2)) \arrow[d, "H_n(\underline{F}(\underline{\varphi_2}))"] \\
H_n(\underline{F}(\mathcal{Q}_1)) \arrow[r, "\sim"] & H_n(\underline{F}(\mathcal{Q}_2))
\end{tikzcd}$$

where the horizontal arrows present topological isomorphisms.

Thus the space $H_n(\underline{F}(\mathcal{P}))$, where $0 \leftarrow X \leftarrow \mathcal{P}$ is a projective resolution of X, does not depend on the particular choice of $\mathcal{P}$ and depends only on X and F. Moreover, if, for every $X \in \mathbf{Mod}$, we fix a copy of this space and denote it by $F_n(X)$, then for every morphism $\varphi: X \to Y$ we obtain a well-defined operator between $F_n(X)$ and $F_n(Y)$, denoted in what follows by $F_n(\varphi)$. The correspondence $X \mapsto F_n(X)$, $\varphi \mapsto F_n(\varphi)$, is obviously a covariant additive functor from **Mod** to **TVS**. We denote it by F_n.

DEFINITION 2.3.1. The functor F_n, $n = 0, 1, 2, \ldots$, is called the n-th projective derived functor of F.

Similarly one defines the *(contravariant) n-th derived functor of a given contravariant additive functor F from* **Mod** *to* **TVS**, denoted by $F^n : \mathbf{Mod} \to \mathbf{TVS}$. Note that the space $F_n(X)$ (or $F^n(X)$) is not bound to be Hausdorff even if F takes values in some subcategory of **TVS**, consisting of Hausdorff spaces. For example, if the range category of F is **Ban**, we can speak about the functors F_n $(F^n) : \mathbf{Mod} \to \langle \mathbf{Ban} \rangle$ (cf. Introduction). The following theorem is a corollary of the "main lemma of homological algebra" in its topological version (Theorem 2.1.2). It describes the most important property of derived functors and at the same time gives an effective method to compute them.

THEOREM 2.3.2. *Let* $0 \leftarrow X \xleftarrow{j} Y \xleftarrow{i} Z \leftarrow 0$ *be an admissible short complex in* **Mod**. *If* $F: \mathbf{Mod} \to \mathbf{TVS}$ *is an additive covariant functor, then there exist linear operators* $\xi_n : F_{n+1}(X) \to F_n(Z)$, $n \geqslant 0$, *such that the sequence*

$$0 \longleftarrow F_0(X) \xleftarrow{F_0(j)} F_0(Y) \xleftarrow{F_0(i)} F_0(Z) \xleftarrow{\xi_0} F_1(X) \xleftarrow{F_1(j)} \cdots$$

is exact. If F *is an additive contravariant functor, then there exist linear operators* $\xi^n : F^n(Z) \to F^{n+1}(X)$, $n \geqslant 0$, *such that the sequence*

$$0 \longrightarrow F^0(X) \xrightarrow{F^0(j)} F^0(Y) \xrightarrow{F^0(i)} F^0(Z) \xrightarrow{\xi^0} F^1(X) \xrightarrow{F^1(j)} \cdots$$

is exact. Moreover, if (in each of these cases) F *takes values in Fréchet, in particular, Banach, spaces, then the operators* ξ_n (ξ^n) *are continuous.*

For a covariant F, let us have a closer look at F_n when $n = 0$. Since for any $X \in \mathbf{Mod}$ and any its projective resolution $0 \leftarrow X \xleftarrow{\varepsilon} P_0 \xleftarrow{d_0} P_1 \leftarrow \cdots$ we have $F_0(X) = F(P_0)/\operatorname{Im} d_0$ and $F(\varepsilon) \circ F(d_0) = 0$, the operator $F(\varepsilon): F(P_0) \to F(X)$ generates a continuous operator $\alpha_X : F_0(X) \to F(X)$. When X runs over **Mod**, these α_X obviously form a *natural transformation of functors*, denoted by $\alpha : F_0 \to F$. In the case of contravariant F, an operator $\alpha^X : F(X) \to F^0(X)$ and a natural transformation $\alpha : F \to F^0$ arise in a similar way. One can easily check

PROPOSITION 2.3.1. *The following properties of* F *are equivalent:*

(i) *for any exact admissible complex*

$$0 \longleftarrow X \longleftarrow X_0 \longleftarrow X_1 \longleftarrow \cdots \qquad (\mathcal{X})$$

in **Mod**, *the complex* $\underline{F}(\mathcal{X})$ *is exact in the terms of* $F(X)$ *and in* $F(X_0)$*;*

(ii) *for any short exact admissible complex* $\underline{Y}$ *in* **Mod** *the complex* $\underline{F}(\underline{Y})$ *is exact at the left hand and middle terms;*

(iii) *for any* X, α_X *(or, according to variance,* α^X*) is a linear isomorphism.*

If, in addition, F *takes values in the class of Fréchet spaces, the word "linear" can be replaced by "topological" (in other words, by the assertion that* α *is a natural equivalence of functors).*

DEFINITION 2.3.2. A functor is called exact on the left if it has the (equivalent) properties indicated in the previous proposition.

EXAMPLE 2.3.1. For any $Y \in \mathbf{Mod}$ the contravariant functor ${}_Ah(?, Y)$ is exact on the left. On the other hand, the functors of the topological tensor product of modules are not, generally speaking, exact on the left; see 2.3.3 below.

We now select a class of "homologically best" functors.

PROPOSITION 2.3.2. *The following properties of F are equivalent:*

(i) *for any exact admissible complex $\mathcal{X}$ in* **Mod** *the complex $\underline{F}(\mathcal{X})$ is exact;*

(ii) *the same, with "short complex" instead of just "complex";*

(iii) *F is exact on the left, and $F_n(X) = 0$ (or, respectively, $F^n(X) = 0$) for all $X \in$* **Mod** *and $n > 0$.*

DEFINITION 2.3.3. A functor is called exact (or, to be precise, projectively exact), if it has the (equivalent) properties, indicated in Proposition 2.3.2.

EXAMPLE 2.3.2. Obviously, the functor $\langle {}_A h(?, J) \mid ? \widehat{\otimes}_A Y \rangle$ is exact iff $\langle J \mid Y \rangle$ is an $\langle$ injective | flat $\rangle$ module.

In certain situations, a derived functor can be computed by means of more general resolutions than projective resolutions.

DEFINITION 2.3.4. An object $Q \in$ **Mod** is called F-acyclic if $F_n(Q) = 0$ when $n > 0$, and $\alpha_Q : F_0(Q) \to F(Q)$ is a linear isomorphism.

Evidently, a projective module is F-acyclic for any functor F.

THEOREM 2.3.3. *Let $0 \leftarrow X \xleftarrow{\varepsilon} \mathcal{Q}$ be an F-acyclic resolution of X in* **Mod**. *Then for every $n \geqslant 0$ there exists a linear isomorphism between the spaces $F_n(X)$ and $H_n(\underline{F}(\mathcal{Q}))$. Moreover, if F takes values in Fréchet spaces, one can replace words "linear isomorphism" by "topological isomorphism"* (or, which is the same, *by "isomorphism in* **TVS**").

Hitherto we discussed projective derived functors. In the situation, when a given category of topological modules is such that its every object has injective coresolutions, one defines by analogy *injective derived functors*; the latter are co- or contravariant depending of the type of the initial functor. All the previous notions and results have obvious versions for injective derived functors. In particular, for any object X the covariant functor ${}_A h(X, ?)$ is injectively exact on the left whereas *this functor is injectively exact iff X is projective.*

2.3.2. *The functor* Ext^n *and some of its applications.* We proceed to a definition of, perhaps, the most important of the derived functors. Now it is appropriate to restrict ourselves to the consideration of the category **A-mod** of Banach left modules over a fixed Banach algebra A. Let X and Y are two of its objects. Denote by $\mathcal{B}^n(A, X, Y)$ the Banach space of continuous polylinear operators from $\underbrace{A \times \cdots \times A}_{n} \times X$ into Y equipped with the polylinear operator norm. We consider the complex

$$0 \to \mathcal{B}(X, Y) \xrightarrow{\delta^0} \mathcal{B}^1(A, X, Y) \to \cdots \to \mathcal{B}^n(A, X, Y) \xrightarrow{\delta^n} \cdots \qquad \big(\tilde{C}(A, X, Y)\big)$$

in which

$$\big[\delta^n f\big](a_1, \ldots, a_{n+1}, x) = a_1 \cdot f(a_2, \ldots, a_{n+1}, x)$$

$$+\sum_{k=1}^{n}(-1)^k f(a_1,\ldots,a_k a_{k+1},\ldots,a_{n+1},x)+(-1)^{n+1} f(a_1,\ldots,a_n,a_{n+1}\cdot x)$$

for $n>0$, and $[\delta^0 f](a,x)=a\cdot f(x)-f(a\cdot x)$. Anticipating on subsequent events, we call it the *standard* Ext*-computing complex.*

In what follows, $\langle F^n | G_n \rangle$ denotes the n-th $\langle$ projective | injective $\rangle$ derived functor of the $\langle$ contra | co $\rangle$ variant functor $\langle F =_A h(\,?\,,Y)\,|\,G={}_A h(X,\,?\,)\rangle$. Using the normalized bar-resolution of X and its respective version among injective coresolutions of Y (see 2.2.1), one can obtain

THEOREM 2.3.4. *For any $n\geqslant 0$, the prenormed spaces $F^n(X)$ and $G_n(Y)$ coincide up to a topological isomorphism, and both can be computed as the n-th cohomology space of the complex $\tilde{C}(A,X,Y)$.*

Because of the indicated coincidence, the following "definition" (in fact, notation) can not lead to a misunderstanding.

DEFINITION 2.3.5. The n-th $\langle$ projective | injective $\rangle$ derived functor from $\langle\ {}_A h(?,Y)\ |\ {}_A h(X,\,?\,)\ \rangle$, considered as a functor between **A-mod** and **TVS** (or, perhaps, between **A-mod** and $\langle$ **Ban** $\rangle$), is denoted by $\langle\ {}_A\operatorname{Ext}^n(\,?\,,Y)\ |\ {}_A\operatorname{Ext}^n(X,\,?\,)\ \rangle$.

Thus ${}_A\operatorname{Ext}^n(X,Y)$ is the complete prenormed space, which is the common value of the first functor, being applied to X, and of the second functor, being applied to Y.

Both functors can be considered as the result of the fixing one of the arguments in the naturally arising bifunctor ${}_A\operatorname{Ext}^n(\,?\,,\,?\,)$. The latter is contravariant on the first and covariant on the second argument.

It is worth noting that some of the introduced spaces can be computed with the help of a wider class of complexes than projective resolutions.

PROPOSITION 2.3.3. *Let $0\leftarrow X\leftarrow \mathcal{X}$ be a flat preresolution of X, and let Y be a dual module. Then ${}_A\operatorname{Ext}(X,Y)$ coincides, up to a topological isomorphism, with the space $H^n(\overline{F}(\mathcal{X}))$ where $F={}_A h(\,?\,,Y)$.*

Now we turn to the respective specialization of Theorem 2.3.2, as well as of its analogue for injective derived functors, in the case of Ext^n. Taking into account that the morphism functors are exact on the left, we obtain the following.

THEOREM 2.3.5. *Let*

$$\big\langle 0\leftarrow X''\leftarrow X\leftarrow X'\leftarrow 0 \qquad (\mathcal{X})\ |\ 0\leftarrow Y''\leftarrow Y\leftarrow Y'\leftarrow 0 \qquad (\mathcal{Y})\big\rangle$$

be a short exact and admissible complex of A-modules. Then for any A-module $\langle Y|X\rangle$, there exists in **TVS** *(in fact, in* $\langle$**Ban**$\rangle$*) an exact sequence*

$$\big\langle 0\to {}_A h(X'',Y)\to {}_A h(X,Y)\to {}_A h(X',Y)\to {}_A\operatorname{Ext}^1(X'',Y)\to\cdots$$

$$\cdots \to {}_A\mathrm{Ext}^n(X', Y) \to {}_A\mathrm{Ext}^{n+1}(X'', Y) \to \cdots \qquad \text{(I)} \mid$$

$$\mid \langle 0 \to {}_Ah(X, Y') \to {}_Ah(X, Y) \to {}_Ah(X, Y'') \xrightarrow{\xi} {}_A\mathrm{Ext}^1(X, Y') \to \cdots$$

$$\cdots \to {}_A\mathrm{Ext}^n(X, Y'') \to {}_A\mathrm{Ext}^{n+1}(X, Y') \to \ldots \qquad \text{(II)} \rangle$$

As one of applications, we obtain a new equivalent definition of the basic classes of "homologically best" modules.

THEOREM 2.3.6. *Let X be a left Banach A-module. Then*

(i) *X is projective iff ${}_A\mathrm{Ext}^1(X, Y) = 0$ for all $Y \in$ **A-mod**, and iff ${}_A\mathrm{Ext}^n(X, Y) = 0$ for all $Y \in$ **A-mod** and $n > 0$,*

(ii) *X is injective iff ${}_A\mathrm{Ext}^1(Y, X) = 0$ for all $Y \in$ **A-mod**, and iff ${}_A\mathrm{Ext}^n(Y, X) = 0$ for all $Y \in$ **A-mod** and $n > 0$,*

(iii) *X is flat iff ${}_A\mathrm{Ext}^1(X, Y) = 0$ for all dual $Y \in$ **A-mod**, and iff ${}_A\mathrm{Ext}^n(X, Y) = 0$ for all dual $Y \in$ **A-mod** and $n > 0$.*

Hitherto we spoke about left modules. Replacing, in the previous discussion, **A-mod** by $\langle$ **mod-A** | **A-mod-A** $\rangle$, one defines functors denoted by $\langle\, \mathrm{Ext}^n_A(\cdot\,,\,\cdot) \mid {}_A\mathrm{Ext}^n_A(\cdot\,,\,\cdot)\rangle$, and obtains the obvious versions of the formulated assertions. We mention a useful connection between some types of these spaces.

PROPOSITION 2.3.4. *For any $X \in \langle$ **A-mod** | **A-mod-A** $\rangle$, $Y \in \langle$ **mod-A** | **A-mod-A** $\rangle$ and $n \geqslant 0$ we have, up to a topological isomorphism,*

$$\big\langle {}_A\mathrm{Ext}^n(X, Y^*) = \mathrm{Ext}^n_A(Y, X^*) \mid {}_A\mathrm{Ext}^n_A(X, Y^*) = {}_A\mathrm{Ext}^n_A(Y, X^*)\big\rangle.$$

The case $n = 1$ explains "why Ext is called Ext". Recall the space $\mathrm{Ext}(\cdot\,,\,\cdot)$ of equivalence classes of module singular extensions, discussed in 1.2.2. Now, in retrospect, we see that it was described just as the quotient space $\mathrm{Ker}\,\delta^1/\mathrm{Im}\,\delta^0$ where δ^1 and δ^0 are the respective operators in the standard Ext-computing complex. Therefore we have

THEOREM 2.3.7. *For all $X, Y \in$ **A-mod**, the space ${}_A\mathrm{Ext}^1(X, Y)$ coincides, up to a topological isomorphism, with ${}_A\mathrm{Ext}(X, Y)$.*

(Again, similar statements are certainly true for right modules and bimodules.)

Thus Propositions 1.2.2 (considered for Banach modules) and 1.2.3 are but particular cases of Theorem 2.3.6.

For $n > 1$, the space ${}_A\mathrm{Ext}^n(X, Y)$ can also be interpreted as the set of equivalence classes of exact admissible complexes modulo a certain (reasonably defined) equivalence relation; now, however, these complexes have the form $0 \leftarrow X \leftarrow Z_1 \leftarrow \cdots \leftarrow Z_n \leftarrow Y \leftarrow 0$. In the pure algebraic context, the respective result (due to Yoneda) is exposed, e.g., in [149].

There is another interpretation of Ext^1, closely connected with the one indicated above. Consider the complex $(\mathcal{Y})$ in Theorem 2.3.5 and take an arbitrary $X \in$ **A-mod** together with a morphism $\varphi : X \to Y''$. Then the second sequence in Theorem 2.3.5 provides an element $\xi(\varphi) \in {}_A\operatorname{Ext}^1(X, Y')$ which is called the *obstruction to lifting the morphism* φ. The reason of such a name is, evidently, that this element is zero iff φ can be "lifted" to some morphism $\psi : X \to Y$, i.e. to such a ψ as makes the diagram

$$\begin{tikzcd} & Y \arrow[d, "\sigma"] \\ X \arrow[ur, "\psi"] \arrow[r, "\varphi"'] & Y'' \end{tikzcd}$$

commutative. Moreover, we have the following

PROPOSITION 2.3.5. *In identifying the spaces* ${}_A\operatorname{Ext}(Y'', Y')$ *and* ${}_A\operatorname{Ext}^1(Y'', Y')$, *the equivalence class of the extension* $\mathcal{Y}$ *becomes the obstruction to lifting the morphism* $\mathbf{1}_{Y''}$ (to a morphism from Y'' to Y).

Another application of Ext^n concerns the question of the stability of module structure under small perturbations. A Banach (say, left) module (X, m) with the outer multiplication $m : A \times X \to X$ is called *stable* if for every left outer multiplication $m' : A \times X \to X$, sufficiently close to m in the sense of the bilinear operator norm $\|m' - m\|$ being small, the new module (X, m') is topologically isomorphic to (X, m).

PROPOSITION 2.3.6. *Let X be such that* ${}_A\operatorname{Ext}^1(X, X) = 0$ *and* ${}_A\operatorname{Ext}^2(X, X)$ *is Hausdorff* (in other words, $\operatorname{Im}\delta^1$ is closed in $\mathcal{B}(A, X)$). *Then X is a stable module.*

We stress, however, that the most important applications of the higher-dimensional functors Ext will be presented in the next section.

Now let us say several words about functors Ext^n in categories of non-Banach modules. It was already mentioned in the previous section that it is still unclear whether all objects of these categories have injective coresolutions.[11] Therefore the question concerning the possibility of the definition of the injective derived functors is still open. Thus in Definition 2.3.5 only the part concerning the projective derived functor ${}_A\operatorname{Ext}^n(\,?\,, Y)$ continues to have sense.

Nevertheless, the covariant functor ${}_A\operatorname{Ext}^n(X, \,?\,)$ for non-Banach modules still can be defined, but in somewhat less straightforward way. Namely, one can assign the obvious precise meaning to the expression ${}_Ah(\mathcal{P}, \,?\,)$, where $\mathcal{P}$ is an arbitrary chosen projective resolution of X, and put ${}_A\operatorname{Ext}^n(X, \,?\,) = H^n({}_Ah(\mathcal{P}, \,?\,))$. It should be noted that Theorem 2.3.5 is valid for non-Banach Ext's as well, although the construction of the second relevant long exact sequence becomes a little bit more complicated matter (see [89], III.4.1).

These inconveniences are, perhaps, counterbalanced by the circumstance that, for some non-Banach modules, the spaces Ext^n can be computed with the help of certain not necessarily admissible complexes whereas some short exact sequences of modules, not being admissible, still generate long exact sequence like those in Theorem 2.3.5. For the most

[11] See footnote to p. 198.

part, such new possibilities arise in the case when some of the underlying spaces of involved modules are nuclear in the sense of Grothendieck. For example, let A be a unital Fréchet algebra, and let $0 \leftarrow A \leftarrow \mathcal{P}$ be an exact (and not, generally speaking, admissible) complex over A, consisting of projective Fréchet bimodules with nuclear underlying spaces. Then, for every Fréchet bimodule X over A which is a DF-space, ${}_A\operatorname{Ext}^n_A(A, X)$ is the n-th cohomology of ${}_Ah_A(\mathcal{P}, X)$. See Taylor [209] for this and several similar results.

2.3.3. *The functor* Tor_n *and the multi-operator holomorphic calculus.* We turn to another class of derived functors, which is of importance second only to Ext^n. Some of its more important applications lie outside the scope of Banach structures. Therefore it is appropriate to discuss it, from the very beginning, in a parallel way for the categories of Banach, Fréchet, $\widehat{\otimes}$- or $\overline{\otimes}$-modules over a fixed algebra A of the respective type. For a while, let us agree to denote by $\langle$ **A-mod** | **mod-A** | **A-mod-A** $\rangle$ the category of $\langle$ left | right | two-sided $\rangle$ A-modules of any of indicated types.

Consider the covariant additive functor $\langle X \widehat{\otimes}_A ?$ | $? \widehat{\otimes}_A Y \rangle$ with domain $\langle$**A-mod** | **mod-A**$\rangle$, an obvious generalization of the "Banach" tensor product functor, introduced in Subsection 0.4. (In the case of $\overline{\otimes}$-modules, the "hat" in the symbol of the tensor product should be replaced everywhere by the "bar".) Depending on the context, we could take the category of Banach spaces, Fréchet spaces or Hausdorff polynormed spaces as the range category of such a functor; however in all cases it will be convenient to choose in this capacity the larger category **TVS**. For $X \in$ **mod-A** and $Y \in$ **A-mod** consider the so-called "*standard* Tor-*computing complex*" $\underset{\sim}{C}(A, X, Y)$

$$0 \leftarrow X \widehat{\otimes} Y \xleftarrow{d_0} X \widehat{\otimes} A \widehat{\otimes} Y \leftarrow \cdots \leftarrow X \widehat{\otimes} \underbrace{A \widehat{\otimes} \cdots \widehat{\otimes} A}_{n} \widehat{\otimes} Y \xleftarrow{d_n} \cdots \qquad (\underset{\sim}{C}(A, X, Y))$$

where

$$\begin{aligned} &d_n(x \widehat{\otimes} a_1 \widehat{\otimes} \cdots \widehat{\otimes} a_{n+1} \widehat{\otimes} y) \\ &\quad = x \cdot a_1 \widehat{\otimes} a_2 \widehat{\otimes} \cdots \widehat{\otimes} a_{n+1} \widehat{\otimes} y \\ &\quad\quad + \sum_{k=1}^{n} (-1)^k x \widehat{\otimes} a_1 \widehat{\otimes} \cdots \widehat{\otimes} a_k a_{k+1} \widehat{\otimes} \cdots \widehat{\otimes} a_{n+1} \widehat{\otimes} y \\ &\quad\quad + (-1)^{n+1} x \widehat{\otimes} a_1 \widehat{\otimes} \cdots \widehat{\otimes} a_n \widehat{\otimes} a_{n+1} \cdot y \end{aligned}$$

for $n > 0$, and $d_0(x \widehat{\otimes} a \widehat{\otimes} y) = x \cdot a \widehat{\otimes} y - x \widehat{\otimes} a \cdot y$.

Denote, for a moment, the functor $\langle X \widehat{\otimes}_A ? \mid ? \widehat{\otimes}_A Y \rangle$ by $\langle F \mid G \rangle$. Computing the projective derived functors $\langle F_n \mid G_n \rangle$, $n \geqslant 0$, with the help of normalized bar-resolution of a given $\langle$ left | right $\rangle$ module, one can obtain the following counterpoint of Theorem 2.3.4.

THEOREM 2.3.8. *For any $n \geqslant 0$, the prenormed spaces $F_n(Y)$ and $G_n(X)$ coincide up to a topological isomorphism, and both can be computed as the n-th homology space of the complex $\underset{\sim}{C}(A, X, Y)$.*

Thus we have the right to give the following "definition".

DEFINITION 2.3.6. The n-th projective derived functor from $\langle X \widehat{\otimes}_A ? \mid ? \widehat{\otimes}_A Y \rangle$ is denoted by $\langle \operatorname{Tor}_n^A(X, ?) \mid \operatorname{Tor}_n^A(?, Y) \rangle$ (it acts from $\langle$ **A-mod** | **mod-A** $\rangle$ to **TVS**).

Thus $\operatorname{Tor}_n^A(X, Y)$ is the complete prenormed space, which is the common value of the first functor, being applied to Y, and of the second functor, being applied to X.

Being combined, both functors lead in an obvious way to the bifunctor $\operatorname{Tor}_n^A(?, ?)$: **A-mod** $\times$ **mod-A** $\to$ **TVS**, which is covariant in both arguments.

If, instead of a one-sided module, an A-bimodule X is given, we can identify it with a right module over A^{env} (cf. 0.1) and thus obtain the functor $\operatorname{Tor}_n^{A^{\mathrm{env}}}(X, ?)$:$\mathbf{A}^{\mathrm{env}}$**-mod** $\to$ **TVS**. Therefore, identifying the category **A-mod-A** with the respective subcategory of $\mathbf{A}^{\mathrm{env}}$**-mod**, we get the "*bimodule* Tor_n *functor*" from **A-mod-A** to **TVS**, denoted by ${}^A\operatorname{Tor}_n^A(X, ?)$. Note that its natural "twin" functor ${}^A\operatorname{Tor}_n^A(?, X)$ (defined by the "symmetric" way) is just the same functor as ${}^A\operatorname{Tor}_n^A(X, ?)$.

The respective analogue of Theorem 2.3.5 reads like this.

THEOREM 2.3.9. *Let*

$$\Big\langle 0 \leftarrow X'' \leftarrow X \leftarrow X' \leftarrow 0 \qquad (\mathcal{X}) \mid 0 \leftarrow Y'' \leftarrow Y \leftarrow Y' \leftarrow 0 \qquad (\mathcal{Y})\Big\rangle$$

be a short exact and admissible complex in $\langle$**mod-A** | **A-mod**$\rangle$. *Then for any* $\langle Y \in$ **A-mod** | $X \in$ **mod-A** $\rangle$ *there exists an exact sequence*

$$\Big\langle 0 \leftarrow \operatorname{Tor}_0^A(X'', Y) \leftarrow \operatorname{Tor}_0^A(X, Y) \leftarrow \operatorname{Tor}_0^A(X', Y) \leftarrow \operatorname{Tor}_1^A(X'', Y) \leftarrow \cdots$$

$$\cdots \leftarrow \operatorname{Tor}_n^A(X', Y) \leftarrow \operatorname{Tor}_{n+1}^A(X'', Y) \leftarrow \cdots \qquad \text{(III)} \mid$$

$$\mid 0 \leftarrow \operatorname{Tor}_0^A(X, Y'') \leftarrow \operatorname{Tor}_0^A(X, Y) \leftarrow \operatorname{Tor}_0^A(X, Y') \leftarrow \operatorname{Tor}_1^A(X, Y'') \leftarrow \cdots$$

$$\cdots \leftarrow \operatorname{Tor}_n^A(X, Y') \leftarrow \operatorname{Tor}_{n+1}^A(X, Y'') \leftarrow \cdots \qquad \text{(IV)} \Big\rangle$$

Here arrows denote linear operators, which are continuous if they connect Tor_n *spaces of the same n. If, in addition,* $\langle \mathcal{X} \mid \mathcal{Y} \rangle$ *consists of Fréchet modules and* $\langle Y \mid X \rangle$ *is also a Fréchet module, then all operators in* $\langle$ (III) | (IV) $\rangle$ *are continuous.*

Note that, unlike the case of Ext^n, we can not replace the 0-th derived functor by the initial functor, that is replace the $\operatorname{Tor}_0^A(\cdot, \cdot)$ spaces by $(\cdot) \widehat{\otimes}_A (\cdot)$. The reason is that the functor of the topological (in contrast to the pure algebraic) tensor product of modules is not, generally speaking, exact to the left. The following simple assertion describes the form of the natural transformation $\alpha : F_0 \to F$ (see 2.3.1) in this particular case; according to the general definition, now we must discuss operators $\langle \alpha_Y | \alpha_X \rangle : \operatorname{Tor}_0^A(X, Y) \to X \widehat{\otimes}_A Y$.

PROPOSITION 2.3.7. *For any pair* $X \in$ **mod-A**, $Y \in$ **A-mod** *the operators* α_Y *and* α_X *coincide. Their kernel is the subspace of points adherent to* $\{0\}$ *in* $\operatorname{Tor}_0^A(X, Y)$, *whereas their image is dense in* $X \widehat{\otimes}_A Y$ *and coincides with the later in the case of Fréchet modules.*

COROLLARY 2.3.1. *If X and Y are Fréchet modules, α_Y $(=\alpha_X)$ is a linear (and hence topological) isomorphism iff* $\operatorname{Tor}_0^A(X,Y)$ *is Hausdorff.*

Combining Theorem 2.3.9 and this Corollary, we get the following characterization of flat modules, parallel to the "projective" (or "injective") part of Theorem 2.3.6.

THEOREM 2.3.10. *Let $\langle X \mid Y \rangle$ be a Fréchet $\langle$ right $\mid$ left $\rangle$ A-module; then the following conditions are equivalent:*

(i) *$\langle X \mid Y \rangle$ is flat,*

(ii) *for all Fréchet $\langle$left$\mid$right$\rangle$ A-module $\langle Y \mid X \rangle$ we have that* $\operatorname{Tor}_1^A(X,Y)=0$, *and* $\operatorname{Tor}_0^A(X,Y)$ *is Hausdorff,*

(iii) *for all Fréchet $\langle$ left $\mid$ right $\rangle$ A-module $\langle Y \mid X \rangle$ and $n>0$ we have that* $\operatorname{Tor}_n^A(X,Y)=0$, *and* $\operatorname{Tor}_0^A(X,Y)$ *is Hausdorff.*

EXAMPLE 2.3.3. Consider the Banach algebra l_p and the Banach l_p-module l_q; $1 \leqslant p, q \leqslant \infty$ with the coordinatewise operations. Then, for the l_p-module $\mathbf{C}$ with the zero outer multiplication, $\operatorname{Tor}_0^{l_p}(l_q,\mathbf{C})$ is Hausdorff just when it vanishes, and it happens iff $q=1$.

As in the case of Ext^n, modules with nuclear underlying spaces provide new opportunities to compute their Tor_n spaces. If, for example, A and $\mathcal{P}$ are the same as at the end of the previous subsection, then, for every Fréchet A-bimodule X, the space ${}^A\operatorname{Tor}_n^A(X,A)$ is the n-th homology of the complex $X \mathbin{\widehat{\otimes}_{A^{\mathrm{env}}}} A$. Again, see [209] for the details of this and similar results.

* * *

One of main applications of the functor Tor to analysis is connected with the Taylor multi-operator spectral theory and its subsequent development (see Preface, where the references were given). It turned out that these functors form a natural setting where the concepts of the joint spectrum and the multi-operator functional calculus, defined in 2.2.3, can be effectively studied. In particular, the key notion of the holomorphic calculus happens to be, after the suitable disrobing, some special Tor space.

Here we indicate some statements which could serve as milestones in Taylor theory. We select those which present an obvious independent interest and at the same time can be formulated without too long preparations.

Taylor theory begins with a given n-tuple $\mathcal{T}=(T_1,\dots,T_n)$ of mutually commuting bounded operators, acting on a Banach space E. The first observation, already done in Example 0.2.5, is that to give such a tuple is equivalent to equip E with the structure of a topological left module, now denoted by X, over the Fréchet algebra $\mathcal{O}(\mathbf{C}^n)$ of entire holomorphic functions on $\mathbf{C}^n$ (in this context T_k, $1 \leqslant k \leqslant n$, is just a notation for the operator $x \mapsto z_k \cdot x$, $x \in X$). In what follows, we use the brief notation $\mathcal{O}$ instead of $\mathcal{O}(\mathbf{C}^n)$.

For any $\lambda=(\lambda_1,\dots,\lambda_n)\in\mathbf{C}^n$ we denote by $\mathbf{C}_\lambda$ the complex plane considered as an $\mathcal{O}$-module with the operation $w\cdot z=w(\lambda)z$; $w\in\mathcal{O}$, $z\in\mathbf{C}$. Take the Koszul complex

$\mathcal{K}(X, \mathcal{T} - \lambda)$ where $\mathcal{T} - \lambda$ means the tuple $(T_1 - \lambda_1 \mathbf{1}, \ldots, T_n - \lambda_n \mathbf{1})$. Using the ("Koszul") resolution, presented in Theorem 2.2.4(i), for the computation of the respective Tor spaces, we easily obtain

PROPOSITION 2.3.8. *For any $m = 0, 1, 2, \ldots$, the space* $\mathrm{Tor}_m^{\mathcal{O}}(\mathbf{C}_\lambda, X)$ *is, up to a topological isomorphism, the m-th homology space* $H_m(\mathcal{K}(X, \mathcal{T} - \lambda))$.

COROLLARY 2.3.2. *The Taylor spectrum of the tuple $\mathcal{T}$ is just* (and thus can be defined as) *the set of those $\lambda \in \mathbf{C}^n$ for which* $\mathrm{Tor}_m^{\mathcal{O}}(\mathbf{C}_\lambda, X) \neq 0$ *for at least one of $m \geqslant 0$* (that is, in fact, for at least one of $m = 0, 1, \ldots, n$).

Now let $\mathcal{U}$ be a domain in $\mathbf{C}^n$. We observe that the notion of a continuous homomorphism $\Gamma : \mathcal{O}(\mathcal{U}) \to \mathcal{B}(X)$ is equivalent to that of a structure of a topological left $\mathcal{O}(\mathcal{U})$-module on X, where $w \cdot x$ is well defined as $[\Gamma(w)](x)$; $w \in \mathcal{O}(\mathcal{U})$, $x \in X$ (cf. the connection between representations and modules indicated in 0.1). Moreover, if we identify every $w \in \mathcal{O}$ with its restriction on $\mathcal{U}$ and thus consider $\mathcal{O}$ as a subalgebra of $\mathcal{O}(\mathcal{U})$, it is evident that a given homomorphism Γ is a holomorphic calculus of $\mathcal{T}$ on $\mathcal{U}$ iff this $\mathcal{O}(\mathcal{U})$-module structure on $\mathcal{U}$ extends the initial $\mathcal{O}$-module structure (in the sense that the former of the relevant outer multiplications $\mathcal{O}(\mathcal{U}) \times X \to X$ and $\mathcal{O} \times X \to X$ extends the latter). Therefore Theorem 2.2.5 ("on holomorphic calculus") can be paraphrased in the following way.

THEOREM 2.3.11. *Let X be a Banach space, and $\mathcal{U}$ be a domain in $\mathbf{C}^n$. Suppose that X has a structure of a topological left $\mathcal{O}$-module. Then this structure can be extended to a structure of a topological left $\mathcal{O}(\mathcal{U})$-module provided $\mathcal{U}$ contains the Taylor spectrum of the tuple $\mathcal{T} = (T_1, \ldots, T_n)$. Moreover, in the case when $\mathcal{U}$ is a domain of holomorphy, this sufficient condition is also necessary, and the extended $\mathcal{O}(\mathcal{U})$-module structure on X is uniquely determined.*

Where to look for the desired $\mathcal{O}(\mathcal{U})$-module structure? Consider, at first for an arbitrary left Fréchet $\mathcal{O}$-module Y, the tensor product space $\mathcal{O}(\mathcal{U}) \mathbin{\widehat{\otimes}_{\mathcal{O}}} Y$ (defined in the same obvious way as the tensor product of Banach modules discussed in 0.4). Observe that it has a structure of a left topological $\mathcal{O}(\mathcal{U})$-module well-defined by $w \cdot (v \otimes_{\mathcal{O}} y) = wv \otimes_{\mathcal{O}} y$, $v, w \in \mathcal{O}(\mathcal{U})$, $y \in Y$, and hence, by restriction, a structure of a left topological $\mathcal{O}$-module as well. Therefore the functor $\mathcal{O}(\mathcal{U}) \widehat{\otimes}_{\mathcal{O}}\,?$ and hence, by virtue of their construction, its derived functors $\mathrm{Tor}_m^{\mathcal{O}}(\mathcal{O}(\mathcal{U}), ?)$, $m = 0, 1, \ldots$, can be considered with values in the category of topological left $\mathcal{O}(\mathcal{U})$-modules and also with the values in that of $\mathcal{O}$-modules (and not just in **TVS**).

In the same way, replacing $\mathcal{O}(\mathcal{U})$ by $\mathcal{O}$, one can obtain functors $\mathrm{Tor}_m^{\mathcal{O}}(\mathcal{O}, ?)$ with values in the category of topological left $\mathcal{O}$-modules. These, however, are very simple: the projectivity of the $\mathcal{O}$-module $\mathcal{O}$ implies that for $m > 0$ the module $\mathrm{Tor}_m^{\mathcal{O}}(\mathcal{O}, Y)$ just vanishes for all Y whereas $\mathrm{Tor}_0^{\mathcal{O}}(\mathcal{O}, Y)$ coincides, up to a topological isomorphism, with Y (cf. "F-acyclic" modules in 2.3.1). Further, by virtue of the functorial properties of Tor (this time with fixed second argument), the embedding of $\mathcal{O}$ into $\mathcal{O}(\mathcal{U})$ generates, for any m, the operator from $\mathrm{Tor}_m^{\mathcal{O}}(\mathcal{O}, Y)$ into $\mathrm{Tor}_m^{\mathcal{O}}(\mathcal{O}(\mathcal{U}), Y)$, which is, as it is easy to see, a morphism of topological $\mathcal{O}$-modules. Thus, in dimension zero we have a morphism between $\mathcal{O}$-modules Y and $\operatorname{Tor}_0^{\mathcal{O}}(\mathcal{O}(\mathcal{U}), Y)$, denoted in what follows by j_Y.

Now we turn to the statement which serves as a principal step in the proof of Theorem 2.3.11 and which is equivalent to that theorem in the case of domains of holomorphy.

THEOREM 2.3.12. *Let X be a Banach space equipped with a structure of a topological $\mathcal{O}$-module, and let $\mathcal{U}$ be a domain of holomorphy. Then the following conditions are equivalent:*

(i) $j_X : X \to \operatorname{Tor}_0^{\mathcal{O}}(\mathcal{O}(\mathcal{U}), X)$ *is a topological isomorphism of $\mathcal{O}$-modules,*

(ii) $\mathcal{U}$ *contains the Taylor spectrum of the tuple $\mathcal{T} = (T_1, \ldots, T_n)$.*

Taking this theorem for granted, one can easily construct the desired $\mathcal{O}(\mathcal{U})$-module structure on X: just put $w \cdot x = j_X^{-1}(w \cdot j_X(x))$, $w \in \mathcal{O}(\mathcal{U})$, $x \in X$, where the dot in brackets corresponds to the already discussed $\mathcal{O}(\mathcal{U})$-module structure on the spaces $\operatorname{Tor}_m^{\mathcal{O}}(\mathcal{O}(\mathcal{U}), Y)$.

In conclusion, we give some comments on the proof of Theorem 2.3.12 itself. It relies on several statements of a combined analytic and algebraic nature. One of these is the known fact concerning the exactness of complexes of analytic Čech cochains in the case of domains of holomorphy (see, e.g., [74]). Equally important to the proof is another statement, which deserves to be specially distinguished. Before its formulation, we note the following. Computing the relevant Tor spaces with the help of the (Koszul) resolution, indicated in Theorem 2.2.4, (where now one should put $\mathcal{U} = \mathbf{C}^n$) we obtain

PROPOSITION 2.3.9. *For any domain $\mathcal{U} \subseteq \mathbf{C}^n$ and $m = 0, 1, 2, \ldots,$ the space $\operatorname{Tor}_m^{\mathcal{O}}(\mathcal{O}(\mathcal{U}), X)$ is the m-th homology space of the Koszul complex $\mathcal{K}(\mathcal{O}(\mathcal{U}, X), \mathcal{S})$, where $\mathcal{S}$ consists of the operators $S_k : \mathcal{O}(\mathcal{U}, X) \to \mathcal{O}(\mathcal{U}, X)$, $k = 1, \ldots, n$, defined by $(S_k\overline{u})(z) = [z_k \cdot \overline{u}](z) - z_k\overline{u}(z)$, $\overline{u} \in \mathcal{O}(\mathcal{U}, X)$* (cf. the complex (2.4) in 2.2.3).

We see that, to speak informally, the complex $\mathcal{K}(\mathcal{O}(\mathcal{U}, X), \mathcal{S})$ is "analytically parametrized" by the complexes of the same form $\mathcal{K}(X, \mathcal{T} - \lambda)$ which participate in the initial definition of the Taylor spectrum. This observation eventually leads to

THEOREM 2.3.13 (Taylor [208]). *Let X be as above, and $\mathcal{U}$ be a domain of holomorphy. Then $\operatorname{Tor}_m^{\mathcal{O}}(\mathcal{O}(\mathcal{U}), X) = 0$ for all $m = 0, 1, 2, \ldots$ iff $\mathcal{U} \cap \operatorname{Sp}\mathcal{T} = \emptyset$. In this case, for any domain $V \subseteq \mathcal{U}$ we have $\operatorname{Tor}_m^{\mathcal{O}}(\mathcal{O}(V), X) = 0$ for all $m = 0, 1, 2, \ldots$.*

Indeed, if there are points $\lambda \in \mathcal{U}$ outside $\operatorname{Sp}\mathcal{T}$, the "fiber" complexes $\mathcal{K}(X, \mathcal{T} - \lambda)$ are exact. Then the techniques of estimating the norms of coefficients in vector-valued power series, originally due to Gleason [58], enables one to prove the exactness of complexes of the form $\mathcal{K}(\mathcal{O}(W, X), \mathcal{S})$ where the W are sufficiently small polydiscs with centers in $\mathcal{U}$. This gives a "local" version of the desired result, and then standard methods of the sheaf theory provide the "global" result concerning our initial domain $\mathcal{U}$.

REMARK. In fact, Theorem 2.3.13 leads to an alternative definition of joint spectrum in terms of sheaves associated with the presheaves $V \mapsto \operatorname{Tor}_m^{\mathcal{O}}(\mathcal{O}(V), X)$. With the help of such an approach, Putinar [174] introduced the notion of the spectrum of the given Fréchet module X over the algebra $\mathcal{O}(\mathcal{M})$ of holomorphic functions on a Stein manifold $\mathcal{M}$ (the initial notion of the Taylor spectrum corresponds to the case of $\mathcal{M} = \mathbf{C}^n$ and the Banach structure on X). In the same paper he gave a substantial generalization of the theorem on the holomorphic calculus. (As to the subsequent ideas and results in this direction, which eventually caused an essential reshaping of the whole theory, see the forthcoming book of Putinar and Eschmeier.)[12]

2.4. *(Co)homology groups and their species*

Cohomology groups are the oldest of all concepts of topological homology, and they are distinguished by their importance. Up to now, a very considerable part of all published papers on the whole area is dedicated to this family of invariants. Being in fact a special case of a derived functor – more precisely, a "bimodule" version of Ext – historically these groups were defined earlier than derived functors, in terms of the so-called standard complex. (Thus, in this respect the development of topological homology just repeated, with a time delay, what happened in pure homological algebra.)

Paying tribute to the history of the topic, we begin with the original definition of these groups, together with some traditional problems in the theory of Banach and topological algebras that lead to their appearance.

2.4.1. *Continuous (co)homology groups; the original definitions.* Let A be a topological algebra, and X be a topological A-bimodule. For $n = 1, 2, \ldots$, we let $C^n(A, X)$ denote the linear space of continuous n-linear operators from

$$\underbrace{A \times \cdots \times A}_{n}$$

into X. ("Continuous" means "separately continuous", with the exception of a given $\widehat{\otimes}$-algebra and $\widehat{\otimes}$-bimodule, when it usually means "jointly continuous". Recall that this rather harmless nuance disappears in the context of Fréchet space where both types of continuity coincide.) If A and X are Banach spaces, then $C^n(A, X)$ is a Banach space with respect to the polylinear operator norm. We also put $C^0(A, X) = X$. The elements in $C^n(A, X)$, $n = 0, 1, \ldots$, are called *n-dimensional cochains.*

Now consider the sequence

$$0 \longrightarrow C^0(A, X) \xrightarrow{\delta^0} C^1(A, X) \xrightarrow{\delta^1} \cdots \xrightarrow{\delta^{n-1}} C^n(A, X) \xrightarrow{\delta^n} \cdots \qquad \big(\widetilde{C}(A, X)\big)$$

where the "coboundary operator" δ^n, $n > 0$, is defined by

$$\begin{aligned}&\delta^n f(a_1, \ldots, a_{n+1})\\ &\quad = a_1 \cdot f(a_2, \ldots, a_{n+1})\end{aligned}$$

[12] It has already appeared; cf. p. 153.

$$+\sum_{k=1}^{n}(-1)^k f(a_1,\ldots,a_{k-1},a_ka_{k+1},\ldots,a_{n+1})$$
$$+(-1)^{n+1}f(a_1,\ldots,a_n)\cdot a_{n+1},$$

$a_1,\ldots,a_{n+1}\in A$, $f\in C^n(A,X)$ and δ^0 is given by $\delta^0x(a)=a\cdot x-x\cdot a$, $a\in A, x\in X$. One can easily verify that for all n we have $\delta^{n+1}\delta^n=0$, and thus $\widetilde{C}(A,X)$ is a complex; it is called *the standard cohomology complex*, or *the Hochschild–Kamowitz complex for A and X*.

The following definition is a topological version of the classical algebraic definition of Hochschild [103].

DEFINITION 2.4.1 (cf. Kamowitz [127], and also Guichardet [72]). The n-th cohomology space of the complex $\widetilde{C}(A,X)$ is called *continuous, or ordinary, n-dimensional cohomology group of a topological algebra A with coefficients in a topological A-bimodule X*. It is denoted by $\mathcal{H}^n(A,X)$.

Note, that the introduced "groups" are, of course, linear spaces. Nevertheless, for historical reasons the term "cohomology groups" is generally accepted.

The spaces $\operatorname{Ker}\delta^n$ and $\operatorname{Im}\delta^{n-1}$ are denoted also by $Z^n(A,X)$ and $B^n(A,X)$, respectively, and their elements are called *n-dimensional cocycles* and *n-dimensional coboundaries*, respectively. Thus

$$\mathcal{H}^n(A,X)=Z^n(A,X)/B^n(A,X).$$

In the case of Banach structure $Z^n(A,X)$ is a closed subspace in $C^n(A,X)$ and $B^n(A,X)$ is, generally speaking, not a closed subspace; thus in this case $\mathcal{H}^n(A,X)$ is considered as a complete prenormed space.

The construction of cohomology groups is essentially functorial. Let us consider the category, whose objects are arbitrary pairs (A,X), where A is a topological algebra, and X is a topological A-bimodule. A morphism $\alpha:(A,X)\longrightarrow(B,Y)$ between two such pairs is, by definition, a pair $\alpha=(\kappa:B\longrightarrow A,\varphi:X_\kappa\longrightarrow Y)$, consisting of a (continuous) homomorphism of algebras and a morphism of topological B-modules; here X_κ is a topological B-module with the underlying space of X and outer multiplication, well defined by $b\cdot x=\kappa(b)\cdot x$ and $x\cdot b=x\cdot\kappa(b)$. Denote this category by **PTAB** and denote its full subcategory, corresponding to the case of Banach structures, by **PBAB**.

Let such an α as above be given. Then every cochain $f\in C^n(A,X)$, $n=0,1,\ldots,$ gives rise to a cochain $g\in C^n(B,Y)$, defined by $g(b_1,\ldots,b_n)=\varphi[f(\kappa(b_1),\ldots,\kappa(b_n)]$ (if $n=0$, we just put $g=\varphi f$). It is easy to observe that, assigning g to f we generate a morphism between the complexes $\widetilde{C}(A,X)$ and $\widetilde{C}(B,Y)$ and hence, for every n, an operator between the respective cohomology spaces $\mathcal{H}^n(A,X)$ and $\mathcal{H}^n(B,Y)$. Thus the association $(A,X)\mapsto\mathcal{H}^n(A,X)$ between objects of **PTAB** and **Lin** can be extended to morphism of these categories, and we get a covariant functor $\mathcal{H}^n:\mathbf{PTAB}\to\mathbf{Lin}$. In the case of Banach structures we get a "more specialized" functor $\mathcal{H}^n:\mathbf{PBAB}\to\langle\mathbf{Ban}\rangle$.

If A is a fixed $\langle$ topological | Banach $\rangle$ algebra, we have a right to treat **A-mod-A**, the category of $\langle$ topological | Banach $\rangle$ A-bimodules, as a subcategory in $\langle$ **PTAB** |

PBAB $\rangle$, consisting of all objects (A, X) and all morphisms $(\mathbf{1}_A, \varphi)$, where A is our fixed algebra. The restriction of $\mathcal{H}^n$ to this subcategory provides a functor, denoted by $\mathcal{H}^n(A, ?): \mathbf{A\text{-}mod\text{-}A} \to \langle \mathbf{Lin} \mid \langle \mathbf{Ban} \rangle \rangle$

REMARK. In fact, the latter functor is but a special case of the covariant Ext functor, introduced in 2.3.2; cf. Theorem 2.4.19 below.

Some other substantial functors, this time with "running" A, arise if we restrict $\mathcal{H}^n$ on some other suitable categories in **PTAB** (or **PBAB**). One of them, the functor of simplicial cohomology, will be considered later, in 2.6.1.

REMARK. If a subalgebra S of a topological algebra A is given, a reasonable "relative" version of the cohomology groups $\mathcal{H}^n(A, X)$ arises (cf. [118] and also [204]). A cochain $f \in C^n(A, X)$, $n > 0$, is called *S-relative*, if it has the properties

$$f(ba_1, a_2, \ldots, a_n) = b \cdot f(a_1, \ldots, a_n),$$

$$f(a_1, \ldots, a_k b, a_{k+1}, \ldots, a_n) = f(a_1, \ldots, a_k, ba_{k+1}, \ldots, a_n)$$

and

$$f(a_1, \ldots, a_n b) = f(a_1, \ldots, a_n) \cdot b$$

for all $a_1, \ldots, a_n \in A$, $b \in S$. It is easy to check that such cochains form a subcomplex in $\widetilde{C}(A, X)$, denoted by $\widetilde{C}(A, , S; X)$ (the space in dimension 0 is taken as $Z^0(S, X)$). The space $H^n(\widetilde{C}(A, , S; X))$ is called the *S-relative n-dimensional cohomology group of A with coefficients in X*; it is denoted by $\mathcal{H}^n(A, S; X)$ (or $\mathcal{H}^n(A, X/S)$). Sometimes it is possible to establish relations between "relative" and "absolute" cohomology. In particular, if, in the context of Banach structures, S is amenable and X is dual, then the spaces $\mathcal{H}^n(A, S; X)$ and $\mathcal{H}^n(A, X)$ are isomorphic (L. Kadison [118]). Results of this kind provide means of computation of the absolute cohomology with the help of the relative version; see [118,204] for the details.

For Banach algebras and bimodules, the vanishing of $\mathcal{H}^n(A, X)$ for $\langle$ all $\mid$ all dual $\rangle$ X and some $n > 0$ implies the vanishing of $\mathcal{H}^k(A, X)$ for $\langle$ all $\mid$ all dual $\rangle$ X and all $k > n$. The part of this assertion, concerning all bimodules, remains to be true in more general context of Fréchet, $\widehat{\otimes}$- and $\overline{\otimes}$-algebras and bimodules. These phenomena will be explained later, from the overview of homological dimension (see 2.5.1); as to the original proof, based on so-called reduction, see [103,109].

The second traditional invariant of topological algebras, the homology groups, has sense only for those cases, where there is a proper notion of the topological tensor product. To be definite, we consider the case when the tensor product "$\widehat{\otimes}$" exists; the "$\overline{\otimes}$-case" can be treated similarly, with obvious modifications.

So let A be a $\widehat{\otimes}$-algebra, and X be a $\widehat{\otimes}$-bimodule over A. For $n = 1, 2, \ldots$ we put

$$C_n(A, X) = \underbrace{A \widehat{\otimes} \cdots \widehat{\otimes} A}_{n} \widehat{\otimes} X$$

and, in addition, we put $C_0(A, X) = X$; elements of $C_n(A, X)$, $n = 0, 1, \ldots$, are called n-dimensional chains. Further, consider so-called *standard homological complex*

$$0 \longleftarrow C_0(A,X) \xleftarrow{d_0} C_1(A,X) \xleftarrow{d_1} \cdots \xleftarrow{d_{n-1}} C_n(A,X) \xleftarrow{d_n} \cdots \qquad (\underset{\sim}{C}(A,X))$$

with "boundary operators" well-defined by

$$\begin{aligned} & d_n(a_1 \otimes \cdots \otimes a_{n+1} \otimes x) \\ & \quad = a_2 \otimes \cdots a_{n+1} \otimes x \cdot a_1 \\ & \quad + \sum_{k=1}^{n} (-1)^k a_1 \otimes \cdots \otimes a_k a_{k+1} \otimes \cdots \otimes a_{n+1} \otimes x \\ & \quad + (-1)^{n+1} a_1 \otimes \ldots \otimes a_n \otimes a_{n+1} \cdot x. \end{aligned}$$

(It is easy to check that it is indeed a complex.)

DEFINITION 2.4.2 (cf. Guichardet [72]). The n-th homology space of $\underset{\sim}{C}(A, X)$ is called *continuous, or ordinary homology group of a $\widehat{\otimes}$-algebra A with coefficients in a $\widehat{\otimes}$-bimodule X over A*. It is denoted by $\mathcal{H}_n(A, X)$.

Note that in the described general situation this "group" is a polynormed (= locally convex) space, and a complete prenormed space in the context of Banach structures.

REMARK. Again, as for cohomology groups, the construction of the space $\mathcal{H}_n(A, X)$ leads to a covariant functor $\mathcal{H}_n$, this time from the category of all pairs ($\widehat{\otimes}$-algebra A, $\widehat{\otimes}$-bimodule over A) to **TVS**. The restriction of $\mathcal{H}_n$ to the subcategory of the latter category, corresponding to a fixed A, provides a functor $\mathcal{H}_n(A, ?)$: **A-mod-A** $\to$ **TVS**. This functor turns out to be a specialization of a certain Tor; cf. Theorem 2.4.19 below.

In the case of Banach structures, it is obvious that the complex in **Ban**, dual to $\underset{\sim}{C}(A, X)$, is isometrically isomorphic to $\widetilde{C}(A, X^*)$. Therefore Theorem 2.1.1 implies the following connection between cohomology and homology groups.

PROPOSITION 2.4.1. *Let A be a Banach algebra, X be a Banach A-bimodule, and let $n > 0$ be an integer. The following conditions are equivalent:*
(i) $\mathcal{H}_k(A, X) = 0$ *for $k \geqslant n$, and the space $\mathcal{H}_{n-1}(A, X)$ is Hausdorff,*
(ii) $\mathcal{H}^k(A, X^*) = 0$ *when $k \geqslant n$.*

REMARK. Similarly to the case of abstract algebras over a field, a Banach algebra A can have an additional structure of a so-called Banach algebra over a given commutative Banach algebra Z. Apart from "usual" algebras, corresponding to the case $Z = \mathbf{C}$, the case when Z is the centre of A (or of A_+) is also of considerable interest. Cohomology groups of such a "Banach Z-algebra" with the properly defined coefficients ("$A - Z$-bimodules") were introduced by Phillips and Raeburn [166] on the model of Definition 2.4.1 above, and their expression in terms of a suitable version of Ext was done by Lykova [146] on the model of Theorem 2.4.19 below. For the details of the definition of these groups, their applications, computations and further related results, see [166,167,146,147].

Now we concentrate on the cohomology groups in small dimensions and proceed to the description of their standard and time-honoured interpretations.

First of all, the case of $n = 0$ is quite transparent. The "group" $\mathcal{H}^0(A, X)$ coincides with $Z^0(A, X)$ and it is the so-called *centre* of the bimodule X, consisting of such x that $a \cdot x = x \cdot a$ for all $a \in A$. If, in particular, A is a Banach algebra and $X = A^*$, $\mathcal{H}^0(A, X) = \{f \in A^*\colon\ f(ab) = f(ba) \text{ for all } a, b \in A\}$; this space is also called the *space of bounded traces on* A and is denoted by A^{tr}. As another example, if A is an operator algebra on H, and $X = \mathcal{B}(H)$, then $\mathcal{H}^0(A, X)$ is just the commutant of A.

REMARK. By comparison, the 0-dimensional homology group, being $X/\operatorname{Im} d_0$ is somewhat more complicated object, because it is not so simple matter to describe $\operatorname{Im} d_0$: it is, generally speaking, larger than the linear span of "commutators" $a \cdot x - x \cdot a$, $a \in A, x \in X$, but smaller than its closure.

As to $\mathcal{H}^1(A, X)$, we already came across this space in 1.2.1 and know what it does mean: it is the space of continuous derivations of A with values in X modulo the subspace of inner derivations (cf. Definition 1.2.3).

We turn to the dimension 2; it deserves an expanded discussion.

2.4.2. *Two-dimensional cohomology and singular extensions of algebras.* We discussed in Subsection 1.2.2 extensions of modules. Now, in somewhat similar spirit we speak about the more complicated and delicate subject of extensions of algebras.

DEFINITION 2.4.3. Let A and I be topological algebras. An extension of A by I is a topological algebra $\underline{\mathbf{A}}$ together with a topologically injective homomorphism $i : I \to \underline{\mathbf{A}}$, the image of which is a two-sided ideal in $\underline{\mathbf{A}}$, and a topologically surjective (= open) homomorphism $\sigma : \underline{\mathbf{A}} \to A$, of which this two-sided ideal is the kernel.

Thus $\underline{\mathbf{A}}$ is a "large" algebra in which there is a certain two-sided ideal that can be identified with I and A can be identified with the quotient algebra modulo this two-sided ideal. Equivalently, we can say that there is a triple $(\underline{\mathbf{A}}, \sigma, i)$ for which the sequence

$$0 \longleftarrow A \xleftarrow{\sigma} \underline{\mathbf{A}} \xleftarrow{i} I \longleftarrow 0 \qquad (\mathcal{E})$$

is exact and, in addition, $\langle i \mid j \rangle$ is topologically $\langle$ injective | surjective $\rangle$.

By virtue of Banach theorem, in the case of Fréchet, in particular, Banach algebras, for such an exact sequence the indicated additional conditions are automatically satisfied.

Suppose we know everything about A and I; what can be said about $\underline{\mathbf{A}}$? Historically this question in the theory of "algebras of analysis" arose first of all for extensions of semisimple algebras by radical algebras [42,51,7]; such extensions are called radical extensions.

Therefore for radical extensions the posed question leads to an attempt to reduce the study of arbitrary algebras to the study of semisimple and of radical algebras.

An ideal form of answer to the question posed would be a description of all extensions up to the following equivalence relation.

DEFINITION 2.4.4. Two extensions $(\underline{\mathbf{A}}, \sigma, i)$ and $(\underline{\mathbf{A}}', \sigma', i')$ of A by I are equivalent if there is a commutative diagram

$$
\begin{tikzcd}
& \underline{\mathbf{A}} \arrow[dd, "\kappa"] \arrow[dr, "\sigma"] & \\
I \arrow[ur, "i"] \arrow[dr, "i'"'] & & A \\
& \underline{\mathbf{A}}' \arrow[ur, "\sigma'"'] &
\end{tikzcd}
$$

in which κ is a topological isomorphism.

(In fact, for Fréchet algebras any continuous homomorphism κ making the diagram commutative is a topological isomorphism.)

REMARK. In the present exposition we do not touch the theory of extensions of C*-algebras (that is the case when $(\mathcal{E})$ consists of such algebras), highly developed and rich in applications (see, e.g., [183,40,9]). We only mention that, investigating such extensions, one usually takes a specific equivalence relation which is less "rigorous" than given in the previous definition.

The question of describing all extensions with given "edges" (A and I) turns out, with the exception of a few special cases, to be very difficult. At the same time classical Wedderburn's theorem suggests a more tangible question, which we naturally deal with in the first instance.

DEFINITION 2.4.5. The extension $\mathcal{E}$ is said to be split if there is a (necessarily closed) subalgebra B of $\underline{\mathbf{A}}$ such that $\sigma|_B : B \to A$ is a topological isomorphism, and $\underline{\mathbf{A}}$, as a topological vector space, decomposes into a topological direct sum of B and $\operatorname{Im} i$.

Obviously, $(\underline{\mathbf{A}}, \sigma, i)$ splits iff σ has a right inverse continuous homomorphism $\rho : A \to \underline{\mathbf{A}}$ whose image as a subspace of $\underline{\mathbf{A}}$ is a topological direct summand of $\operatorname{Im} i$. Such a ρ is called a *splitting homomorphism* for the given extension. In the case of Fréchet algebras, the very existence of a closed subalgebra B of A such that $\underline{\mathbf{A}} = B \oplus \operatorname{Im} i$ (in **Lin**) is, by virtue of Banach theorem, sufficient for splitting.

REMARK. A weaker version of the concept of splitting is algebraic splitting, that is the existence of a, generally speaking, non-closed subalgebra of $\underline{\mathbf{A}}$ with $\underline{\mathbf{A}} = B \oplus \operatorname{Im} i$ (in **Lin**) or, equivalently, the existence of a, generally speaking, discontinuous algebra homomorphism that is a right inverse to σ. This notion and its interplay with "topological" splitting was explored in a pioneering article of Bade and Curtis [7]; as to later work, see [77,5] and also [89]. Note that in most of cited papers the authors use another terminology and, instead of $\langle$splitting | algebraic splitting$\rangle$ speak about $\langle$strong Wedderburn | Wedderburn$\rangle$ decomposition.

Now we pass to the class of extensions that can be studied effectively with the help of cohomology groups.

DEFINITION 2.4.6. An extension $\mathcal{E}$ is said to be singular if I has the zero multiplication and σ has a right inverse continuous operator $\rho : A \to \underline{\mathbf{A}}$ (called the splitting operator for the given extension).

The existence of a splitting operator is obviously equivalent to $\operatorname{Im} i$ being complemented in $\underline{\mathbf{A}}$, or, in other words, $\mathcal{E}$ splitting as a complex in the (additive) category **TVS**. Regarding the condition $I^2 = 0$, this easily implies that I has the structure of a polynormed A-bimodule with operations $a \cdot x = i^{-1}(i(x)b)$, where $x \in I$, $a \in A$, and b is any element of $\underline{\mathbf{A}}$ with $\sigma(b) = a$. If we let X denote I in its new capacity as a bimodule, we shall call the original $\mathcal{E}$ a singular extension of the algebra A by the A-bimodule X.

It should be mentioned that in some cases the condition, imposed on σ, is automatically satisfied. This concerns annihilator extensions, that is those with $i(a)b = bi(a) = 0$ for all $a \in I, b \in \underline{\mathbf{A}}$.

PROPOSITION 2.4.2 (Karyaev and Yakovlev [130]). *Let A be a Banach algebra such that the operator $\sigma : A \widehat{\otimes}_A A \to A$: $a \widehat{\otimes}_A b \mapsto ab$ is topologically injective, and* $\dim A/\overline{A^2} < \infty$. *Then any annihilator extension of A* (in the category of Banach algebras) *is singular.*

The indicated condition on A is satisfied by any Banach algebra with a bounded approximate identity (see Proposition 0.4.2; other examples are l_1, $\mathcal{N}(H)$ and, say, all maximal ideals in the disk-algebra. Nevertheless, in general the condition on σ cannot be abandoned. As it was shown by Yakovlev [220], there exists a construction (in fact, a functor) that takes each pair, consisting of a Banach space E and a non-complemented subspace E_0 of it, to an annihilator extension of a Banach algebra $(\underline{\mathbf{A}}, \sigma, i)$ such that $\operatorname{Im} i$ is not complementable (as a subspace) radical of $\mathcal{A}$.

The set of equivalence classes of singular extensions of A by X is denoted by $\operatorname{Ex}(A, X)$. It has a natural distinguished point θ, namely, the class consisting of all splittable extensions.

In what follows, when we talk about a singular extension $(\underline{\mathbf{A}}, \sigma, i)$, we identify X with $\operatorname{Im} i$. If ρ is a splitting operator, the element $\rho(a)\rho(b) - \rho(ab)$, denoted by $f_\rho(a, b)$, obviously lies in X for all $a, b \in A$. In this way we obtain a continuous (separately or, according to the context, jointly) bilinear operator $f_\rho : A \times A \to X$, that is, an element $f_\rho \in C^2(A, X)$. Now one can observe the following

PROPOSITION 2.4.3. *For any splitting operator ρ, $f_\rho \in Z^2(A, X)$. If ρ and ρ' are two such operators, then $f_\rho - f_{\rho'} \in B^2(A, X)$.*

That leads to the following topological version of a well-known result of Hochschild [103].

THEOREM 2.4.1 (cf. [127]). *Let A be a topological algebra, and X be a topological A-bimodule. Then there exists a bijection between* $\mathrm{Ex}(A, X)$ *and* $\mathcal{H}^2(A, X)$ *taking the equivalence class of an extension with the splitting operator ρ to the coset $f_\rho + B^2(A, X) \in \mathcal{H}^2(A, X)$. Moreover, this bijection takes the class θ to $0 \in \mathcal{H}^2(A, X)$.*

COROLLARY 2.4.1. *All the singular extensions of A by the A-bimodule X split iff $\mathcal{H}^2(A, X) = 0$.*

Thus we see that the 2-dimensional cohomology group is "responsible" for the splitting of singular extensions. To this in its turn, one can reduce the question of splittability of some more general extensions. We call an extension *nilpotent*, if, for some n, $a_1 \dots a_n = 0$ for any $a_1, \dots, a_n \in I$.

PROPOSITION 2.4.4 (idem).
(i) *Suppose $\mathcal{H}^2(A, X) = 0$ for every A-bimodule X. Then every nilpotent extension $\underline{\mathbf{A}}$ of A by I, such that I and all powers of I^2 are complemented in $\underline{\mathbf{A}}$, splits.*
(ii) *Suppose $\mathcal{H}^2(A, X) = 0$ for every finite-dimensional X. Then every extension of A by finite-dimensional I splits.*

Note that the algebra $A \oplus X$, considered much earlier in Proposition 1.2.1, is actually a splittable singular extension of A by X. This proposition implies that the discussion of splittable extensions involves 1-dimensional cohomology:

PROPOSITION 2.4.5. *In a splittable singular extension $\underline{\mathbf{A}}$ of A by X the splitting homomorphism* (or, equivalently, in the context of Fréchet algebras, the closed subalgebra B in $\underline{\mathbf{A}}$ with $\underline{\mathbf{A}} = B \oplus X$) *is unique iff $\mathcal{H}^1(A, X) = 0$ and $\mathcal{H}^0(A, X) = X$.*

* * *

Now we cite a number of known facts about the 2-dimensional cohomology groups of certain algebras. In what follows, "*annihilator bimodule*" means a bimodule with zero outer multiplications.

THEOREM 2.4.2 [89]. *Let A be a Banach algebra, and X be a Banach annihilator A-module. Then $\mathcal{H}^2(A, X) = 0$ iff the operator $A \mathbin{\widehat{\otimes}}_A A \to \overline{A^2}$ is a topological isomorphism, and every continuous operator from $\overline{A^2}$ to X can be extended to a continuous operator from A to X.*

(Note that the latter condition is certainly satisfied if $\overline{A^2}$ is complementable in A or $\dim X < \infty$.)

COROLLARY 2.4.2. *If A has a left or right bounded approximate identity, then $\mathcal{H}^2(A, X) = 0$ for every annihilator X.*

This, in its turn, implies

THEOREM 2.4.3 [78,106]. *Let A be a Banach algebra such that every two-sided ideal in A of finite codimension has a bounded approximate identity* (e.g., A is a C*-algebra). *Then every extension of A by a finite-dimensional I splits.*

The first example of a non-trivial 2-dimensional cohomology group with annihilator coefficients, namely $\mathcal{H}^2(l_2, \mathbf{C}_\infty)$, was presented by Feldman [51]. (It was done in the language of extensions; later it was observed [6] that the constructed extension was algebraically splittable.) The analysis of this extension leads to

EXAMPLE 2.4.1 [89]. The space $\mathcal{H}^2(l_p, X)$, $1 \leqslant p \leqslant \infty$, where $X \neq 0$ is an annihilator l_p-bimodule, vanishes iff $p = 1$ or ∞ (and for other p and $X = \mathbf{C}_\infty$

$$\mathcal{H}^2(l_p, \mathbf{C}_\infty) = \begin{cases} l_{p/(p-2)}/l_{p/(p-1)}, & \text{if } 2 < p < \infty, \\ l_\infty/l_{p/(p-1)}, & \text{if } 1 < p < 2). \end{cases}$$

EXAMPLE 2.4.2. Let $\mathcal{A}(\mathbf{D})$ be disk algebra (see Introduction), and $A = \{w \in \mathcal{A}(\mathbf{D})\colon w(0) = w'(0) = 0\}$. Then for the annihilator bimodule $\mathbf{C}_\infty$, $\mathcal{H}^2(A, \mathbf{C}_\infty) \simeq \mathbf{C}^2$; in particularly, the coset of the cocycle

$$f(w_1, w_2) = \frac{\partial^3 w_1}{\partial z^3}(0)\frac{\partial^3 w_2}{\partial z^3}(0)$$

is not trivial [89]. At the same time $\mathcal{H}^n(\mathcal{A}(\mathbf{D}), \mathbf{C}) = 0$ for any $\mathcal{A}(\mathbf{D})$-bimodule structure in $\mathbf{C}$ and $n \geqslant 2$ [109].

From annihilator coefficients we pass to so-called commutative coefficients, that is modules X with $a \cdot x = x \cdot a$ for all $a \in A$ and $x \in X$. The following result was established by Kamowitz for $A = C(\Omega)$ (being, incidentally, the first explicit computation of Banach cohomology groups). Its present general form is due to Johnson [109].

THEOREM 2.4.4. *Let A be a commutative amenable Banach algebra and X be a commutative Banach A-bimodule. Then $\mathcal{H}^1(A, X) = \mathcal{H}^2(A, X) = 0$.*

(The part concerning the 1-dimensional cohomology was already mentioned when we discussed weakly amenable algebras; see 1.3.3.)

It is not known whether the equality $\mathcal{H}^n(A, X) = 0$ where A and X are as before, is valid for all, or at least for some, $n \geqslant 3$. The problem is open even for $A = C(\Omega)$, and, being posed for this particular algebra, is apparently the oldest among all open problems in topological homology (cf. [109,35,89]).

In the same basic paper [103] where he discovered the cohomology groups of (associative) algebras, Hochschild introduced the class of so-called *completely segregated* algebras A – those, which have zero 2-dimensional cohomology groups with all possible coefficients. This class, in particular, contains quite a few commutative semisimple algebras over $\mathbf{C}$, such as the polynomial algebra in one variable, the group algebra of $\mathbf{Z}$ or, say, the algebra c_{00} of finitary sequences. The following assertion shows that the situation with the 2-dimensional cohomology in the context of Banach algebras is strikingly different from what could be expected from the experience in pure algebra.

THEOREM 2.4.5 [85]. *Let A be a commutative Banach algebra with infinite spectrum (= maximal ideals space). Then there exists a Banach A-bimodule X such that* $\mathcal{H}^2(A, X) \neq 0$ (or, equivalently, there exists a singular extension of A which does not split).

COROLLARY 2.4.3. *Every infinite-dimensional function (= commutative and semisimple) Banach algebra has a non-splittable radical extension.*

Thus, in the class of commutative semisimple Banach algebras – contrary to the respective class of abstract algebras – the condition $\mathcal{H}^2(A, X) = 0$ for all X just means that our algebra is $\mathbf{C}^n$ for some n (and hence, it is equivalent to the condition $\mathcal{H}^1(A, X) = 0$ for all X, that is to the contractibility of A).

The theorem presented follows from a somewhat stronger statement, the so-called "global dimension theorem". This statement, together with the construction of bimodules X, possessing the desired property, will be discussed in 2.5.2.

If A, in addition to the assumptions of Theorem 2.4.5, is semisimple and amenable, the space $\mathcal{H}^2(A, X)$ does not vanish already for $X = A \widehat{\otimes} A$ [108].

2.4.3. *The spaces $\mathcal{H}^n(A, A)$ and stability problems.* There is another area of applications of cohomology groups: questions concerning perturbations (or deformations).

Let A be a Banach algebra with multiplication $m : A \times A \to A$. We let (A, m') denote the Banach algebra with the same underlying space but with multiplication m'. A property of the algebra A is called *stable under small perturbations*, or just *stable* if there exists $\varepsilon > 0$ such that all algebras (A, m') with $\|m' - m\| < \varepsilon$ have the given property. The algebra itself is *stable* (or *rigid*) if there exists $\varepsilon > 0$ such that every (A, m') with $\|m' - m\| < \varepsilon$ is topologically isometric to A.

Here is a sufficient condition of stability, expressed in terms of cohomology.

THEOREM 2.4.6 (Johnson [112], see also Raeburn and Taylor [176]). *Let A be a Banach algebra such that $\mathcal{H}^2(A, A) = 0$ and $\mathcal{H}^3(A, A)$ is Hausdorff* (i.e. $B^3(A, A)$ is closed in $Z^3(A, A)$). *Then A is stable.*

We should like to clarify, in informal way, how such cohomological properties work to establish the stability.

Consider the sequence of (non-linear) maps

$$\begin{aligned}
&\{\text{the set of invertibles in } \mathcal{B}(A)\} \xrightarrow{\alpha} C^2(A, A) \xrightarrow{\beta} C^3(A, A),\\
&\quad\text{where } \alpha(\varphi) : (a, b) \mapsto \varphi^{-1}\big(m(\varphi(a), \varphi(b))\big),\\
&\quad\text{and } \beta(\psi) : (a, b, c) \mapsto \psi\big(\psi(a, b), c\big) - \psi\big(a, \psi(b, c)\big).
\end{aligned} \tag{2.6}$$

Obviously $\beta^{-1}(0)$ is the set of all (continuous) multiplications in A and $\operatorname{Im}\alpha$ is the set of those multiplications m' for which (A, m') is topologically isomorphic to A. Therefore the stability of the given multiplication $m = \alpha(1)$ is expressed by the fact that, if the bi-operator ψ is sufficiently near m, then $\beta(\psi) = 0$ implies $\psi = \operatorname{Im}\alpha$; that is to say, the sequence (2.6) has a "local exactness" property. Since both maps α and β are smooth it is natural to be interested in the "infinitesimal picture of what is going on", that is, to consider the complex

$$\mathcal{B}(A) = C^1(A, A) \xrightarrow{\alpha'(1)} C^2(A, A) \xrightarrow{\beta'(m)} C^3(A, A), \tag{2.7}$$

where a dash means the Fréchet derivative; further, simple calculations show that $\alpha'(1) = \delta^1$, and $\beta'(m) = \delta^2$. It turns out that, if $B^2(A, A) = Z^2(A, A)$ (the "infinitesimal" complex (2.7) is exact) and at the same time $B^3(A, A)$ is closed, then this ensures that (2.6) is "locally exact", that is, the multiplication m is stable. (To transfer "from the infinitesimal to the local" like this is in itself nontrivial; it is shown to be possible in [176] as a corollary of a new version of the implicit function theorem.)

Thus a Banach algebra is certainly stable, if $\mathcal{H}^n(A, A) = 0$ for all $n \geqslant 2$. At the moment, it is known that many, although by no means all, of the "popular" examples of algebras possess latter property. In particular, the class of such algebras includes:
- $C(\Omega)$, where Ω is a compact metrizable set [112],
- $\mathcal{B}(E)$, where E is an arbitrary Banach space [125] (this will be commented on later, in 2.4.5),
- $\mathcal{K}(H)$, where H is a Hilbert space [108],
- $\mathcal{N}(E)$, where E is a Banach space with the approximation property ([85] and [191] with [195] combined; cf. Theorem 2.4.21 below),
- $L^1(G)$, where G is an amenable compact group [112].

Note that for the first two algebras from this list we have $\mathcal{H}^1(A, A) = 0$ as well (see Theorems 2.4.4 and 1.2.9) whereas for the last three ones we have, generally speaking, $\mathcal{H}^1(A, A) \neq 0$ (see 1.2.1 and also [216]). Finally, $\mathcal{H}^2(\mathcal{N}(E), \mathcal{N}(E)) = 0$ for all E whereas $\mathcal{H}^3(\mathcal{N}(E), \mathcal{N}(E)) = 0$ iff E has the approximation property [195].

(The important question of the vanishing $\mathcal{H}^n(A, A)$ for von Neumann algebras and other operator algebras on Hilbert spaces, as well as some applications to stability type problems, will be discussed in two next subsections.)

As the opposite end, the algebras l_p, $1 < p < \infty$, with the coordinatewise multiplication, and also $l_1(F_2)$, where F_2 is the free group with two generators, are not stable [112]. Other examples of stable and not stable algebras see [112,113], and also the book [105].

REMARK. The conditions of Theorem 2.4.6 are not necessary for the stability. As an example, the disk-algebra $\mathcal{A}(\mathbf{D})$ is stable (Rochberg [182]); at the same time $\mathcal{H}^2(\mathcal{A}(\mathbf{D}), \mathcal{A}(\mathbf{D})) \neq 0$ [109].

One can speak not only about the stability of Banach algebras, but also about the stability of some of their maps. A unital continuous homomorphism $\kappa : A \to B$ between two unital Banach algebras is called *stable* if there exists $\varepsilon > 0$ such that, for any homomorphism $\kappa' : A \to B$ with $\|\kappa' - \kappa\| < \varepsilon$, there exists an invertible element $b \in B$ with $\kappa'(a) = b^{-1}\kappa(a)b$, $a \in A$.

THEOREM 2.4.7 (Johnson [112], see also Raeburn and Taylor [176]). *Let κ be as above and B be considered as a Banach A-bimodule with the operations $a \cdot b = \kappa(a)b$ and $b \cdot a = b\kappa(a)$. Suppose that $\mathcal{H}^1(A, B) = 0$, and $\mathcal{H}^2(A, B)$ is Hausdorff. Then κ is stable.*

REMARK. Theorem 2.4.23 below will show that Proposition 2.3.6 is a particular case of this theorem, corresponding to $B = \mathcal{B}(X)$.

Note also that amenability is always a stable property of Banach algebras, and the same is true for commutativity provided that, for the given algebra, we have $\mathcal{H}_2(A, A) = 0$ [112].

* * *

In the study of derivations, two-dimensional cohomology groups can also be useful. Let I be a closed two-sided ideal in topological algebra A and let D and $\widehat{D}$ be derivations of A and A/I respectively. D is called a *lifting* of $\widehat{D}$ if the diagram

$$\begin{array}{ccc} A & \xrightarrow{D} & A \\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\tau} \\ A/I & \xrightarrow{\widehat{D}} & A/I \end{array}$$

in which τ is the natural projection, is commutative.

PROPOSITION 2.4.6 (cf. Ringrose [180]).

(i) *Let I be complemented, as a subspace, in A and $\mathcal{H}^2(A, I) = 0$. Then every continuous derivation of A/I admits a continuous lifting.*

(ii) *Let $\mathcal{H}^2(A, A) = 0$ and $\overline{I^2} = I$. If every continuous derivation of A/I has a lifting, then $\mathcal{H}^2(A, I) = 0$.*

2.4.4. *Cohomology of operator algebras: interplay of different versions, and computation results.* The continuous, or ordinary, cohomology is the earliest and best known of all topological modifications of Hochschild cohomology. However, there are versions which meet other topological conditions on cohains; most of these make sense for operator algebras and reflect their specific features.

Let A be an operator algebra (so far arbitrary), acting on a Hilbert space H, and let $X = (X_*)^*$ be a dual Banach A-bimodule. Recall the notions of a normal operator from A to X, and of the left normal, right normal and normal parts of X, introduced in 0.3. A cochain $f \in C^n(A, X)$; $n > 0$ is called *normal* if this polylinear operator is normal in each argument.

Denote by $C^n_{\mathrm{w}}(A, X_{\mathrm{w}})$; $n > 0$ the (norm closed) subspace in $C^n(A, X)$ consisting of normal cochains with values in X_{w}, the normal part of X, and put $C^0_{\mathrm{w}}(A, X_{\mathrm{w}}) = X_{\mathrm{w}}$.

Since the coboundary operators δ^n in $\widetilde{C}(A, X)$ preserve the property of a cochain to be normal, this complex has a subcomplex

$$0 \longrightarrow C_{\mathrm{w}}^0(A, X_{\mathrm{w}}) \longrightarrow \cdots \longrightarrow C_{\mathrm{w}}^n(A, X_{\mathrm{w}}) \longrightarrow \cdots \qquad \big(\widetilde{C}_{\mathrm{w}}(A, X_{\mathrm{w}})\big)$$

which is called the *standard normal cohomology complex* (for A and X).

DEFINITION 2.4.7. The n-th cohomology space of the complex $\widetilde{C}_{\mathrm{w}}(A, X_{\mathrm{w}})$ is called normal n-dimensional cohomology group of the operator algebra A with coefficients in a (dual) A-bimodule X.

This cohomology, for the most important case of normal X, was introduced by Kadison and Ringrose [121]. As to other interesting cases, it was noticed in [101] that for a von Neumann algebra A and $X = A^*$ (and thus $X_{\mathrm{w}} = A_*$) we obtain the special spaces "H_σ^n", considered by Christensen and Sinclair [24].

At the beginning, the normal cohomology was invented as a useful tool for computing the ordinary cohomology. The matter is that sometimes it is easier to deal with the normal, rather than the ordinary, cohomology, and then to pass from the former to the latter. This strategy is relied on the following result, which was established by Johnson, Kadison and Ringrose [116] for normal bimodules with some mild additional assumptions.

THEOREM 2.4.8 (cf. [116,91]). *Let A be an operator* C*-*algebra, and X be a dual Banach A-bimodule. Then, up to a topological isomorphism,*

$$\mathcal{H}_{\mathrm{w}}^n(A, X_{\mathrm{w}}) = \mathcal{H}^n(A, X_l) = \mathcal{H}^n(A, X_r)$$

for any $n \geqslant 0$. In particular, if X is normal, then normal and ordinary cohomology groups coincide in all dimensions.

The sketch of the proof will be given later; see 2.4.5 below. As to the original argument in [116], it essentially uses the following "averaging" principle, which has a considerable independent interest.

PROPOSITION 2.4.7. *Let A be an operator* C*-*algebra, and X be a normal A^--, and hence A-bimodule. Let, further, B is a closed amenable subalgebra of A. Then, for any $n \geqslant 1$, every coset in $\mathcal{H}^n(A, X) = Z^n(A, X)/B^n(A, X)$ contains a cochain $g : A \times \cdots \times A \to X$ which vanishes whenever any of arguments lies in B.*

REMARK. This result can be interpreted as the coincidence of $\mathcal{H}^n(A, X)$ with the relative cohomology $\mathcal{H}^n(A, B; X)$ mentioned in 2.4.1. See [204] for related results and details.

The established relation between normal and ordinary cohomology, together with the observation that $\mathcal{H}_{\mathrm{w}}^n(A, X) = \mathcal{H}_{\mathrm{w}}^n(A^-, X)$ for any operator C*-algebra A and a normal A^--bimodule X, enabled the authors of [116] to give the first proof of the vanishing of both varieties of cohomology for hyperfinite von Neumann algebras with normal coefficients in positive dimensions (cf. Theorem 2.4.26 below). That meant, in particular, that to the list of algebras with $\mathcal{H}^n(A, A) = 0$ for $n \geqslant 2$, presented in the previous subsection, one should add all hyperfinite von Neumann algebras. And this, in its turn, leads to some stability type properties for closed subalgebras in $\mathcal{B}(H)$ in a framework of an approach, initiated by Kadison and Kastler [120] and Christensen [17].

Recall that the distance between two subspaces F and G of a Banach space E is the number $\mathrm{d}(F, G) = \sup\{\mathrm{d}(a, S_B), \mathrm{d}(b, S_A);\ a \in S_A,\ b \in S_B\}$, where $\langle \mathrm{d}(a, S_B) \mid \mathrm{d}(b, S_A) \rangle$ denotes the distance between $\langle a \mid b \rangle$ and the unit ball $\langle S_B \mid S_A \rangle$ in $\langle B \mid A \rangle$.

THEOREM 2.4.9 (Raeburn and Taylor [176]). *Let A be a hyperfinite von Neumann algebra on H. Then there exists $\delta > 0$ such that if B is another von Neumann algebra on H with $\mathrm{d}(A, B) < \delta$, then it is (necessary isometrically) *-isomorphic to A.*

* * *

The year 1995 was the silver anniversary of the normal version of cohomology groups, but now we turn to a much more recent modification of these groups, the so-called completely bounded cohomology groups.

Let E be a linear space. We denote by $M_n(E)$ the space of all $n \times n$ matrices with entries in E, and for $u \in M_n(E)$, $v \in M_n(E)$ we denote by $u \oplus v$ the element in M_{m+n} with $\langle u \mid v \rangle$ at the $\langle$ upper left $\mid$ lower right $\rangle$ corner.

DEFINITION 2.4.8 (Ruan [185]). The space E, equipped with the family $\|\cdot\|_n$; $n = 1, 2, \ldots$, where $\|\cdot\|_n$ is a norm on $M_n(E)$, is called an (abstract) operator space if it satisfies $\|u \oplus v\|_{m+n} = \max\{\|u\|_m, \|v\|_n\}$ and $\|\alpha v \beta\|_m \leqslant \|\alpha\| \|v\|_n \|\beta\|$ for all $u \in M_m(E)$, $v \in M_n(E)$ and $\alpha, \beta \in M_n(\mathbf{C})$.

(Here $\alpha v \beta$ is defined by the obvious way via the pattern of the matrix multiplication, and the norm in $M_n(\mathbf{C})$ is that in $\mathcal{B}(\mathbf{C}^n)$, where $\mathbf{C}^n$ has the usual Hilbert space structure.)

The main example is $\mathcal{B}(H)$, equipped with the operator norm in $M_n(\mathcal{B}(H))$, naturally identified with $\mathcal{B}(H \oplus \cdots \oplus H)$.

If E is an operator space, then every subspace F of it is also an operator space with respect to norms in $M_n(F)$ as in the subspace in $M_n(E)$. Thus every subspace in $\mathcal{B}(H)$ is an operator space; such operator spaces are called *concrete*.

DEFINITION 2.4.9. Let E and F be operator spaces. A linear operator $\varphi: E \to F$ is called completely bounded if every operator $\varphi_n : M_n(E) \to M_n(F)$: $[x_{ij}] \mapsto [\varphi(x_{ij})]$ is bounded, and $\sup\{\|\varphi_n\|: n = 1, 2, \ldots\} < \infty$. The latter number is called the cb-norm of φ and is denoted by $\|\varphi\|_{\mathrm{cb}}$. An operator φ is called a complete $\langle$ contraction $\mid$ isometry $\rangle$ if each φ_n is a $\langle$ contraction $\mid$ isometry $\rangle$.

The name "operator space" is justified by the following principal result.

THEOREM 2.4.10 (Ruan; idem). *Every abstract operator space is completely isometric to some concrete operator space.*

The crucial concept is

DEFINITION 2.4.10 (Christensen and Sinclair [23]). Let E_k, $1 \leqslant k \leqslant m$, and F be operator spaces. A polylinear operator $f : E_1 \times \cdots \times E_m \to F$ is called completely bounded if, for any $n \geqslant 1$, the polylinear operator $f_n : M_n(E_1) \times \cdots \times M_n(E_m) \to M_n(F)$, taking a tuple $([x^1_{ij}], \ldots, [x^m_{ij}])$ to $[y_{ij}]$ with

$$y_{ij} = \sum_{k_1,\ldots,k_{m-1}=1}^{n} f(x_{ik_1}, x_{k_1k_2}, \ldots, x_{k_{m-1}j})$$

is bounded (relative to norms, included in the operator space structure), and the respective polylinear operator norms $\|f_n\|$ have a finite upper bound. The latter is called cb-norm of f and is denoted by $\|f\|_{\mathrm{cb}}$. (Thus, f_n is constructed by the imitation of matrix multiplication).

Now let A be an operator norm closed subalgebra of $\mathcal{B}(H)$, $A_\times = \mathrm{l.h.}(A, \mathbf{1}_H)$ and X be an operator space equipped with the structure of a unital $A_\times$-bimodule.

DEFINITION 2.4.11. The bimodule X is called a completely bounded A-bimodule if the trilinear operator $A_\times \times X \times A_\times \to X$: $(a, x, b) \mapsto a \cdot x \cdot b$ is completely bounded with respect to $\langle$ given | natural $\rangle$ operator space structure on $\langle X \mid A_\times \rangle$.

Finally, if A is as above and X is a completely bounded A-bimodule, then a completely bounded n-linear operator from $A \times \cdots \times A$ to X is called an n-dimensional completely bounded cochain. The spaces $C^n_{\mathrm{cb}}(A, X)$, consisting of such cochains, together with the corresponding coboundary operators, form a subcomplex, denoted by $\widetilde{C}_{\mathrm{cb}}(A, X)$, of the standard complex $\widetilde{C}(A, X)$.

DEFINITION 2.4.12 (Christensen, Effros and Sinclair [20]). The n-th cohomology of $\widetilde{C}_{\mathrm{cb}}(A, X)$ is called the n-dimensional completely bounded cohomology group of A with coefficients in the completely bounded A-bimodule X. It is denoted by $\mathcal{H}^n_{\mathrm{cb}}(A, X)$.

If, in addition, a completely bounded A-bimodule X is also a dual Banach A-bimodule, we can speak of those completely bounded cochains which are also normal and have values in X_{w}. Thus a subcomplex $\widetilde{C}_{\mathrm{cbw}}(A, X_{\mathrm{w}})$ of $\widetilde{C}_{\mathrm{cb}}(A, X)$ (and of $\widetilde{C}_{\mathrm{w}}(A, X_{\mathrm{w}})$ as well) appears, consisting of such cochains. Its cohomology, the "normal completely bounded" version of the continuous cohomology, is denoted by $\mathcal{H}^n_{\mathrm{cbw}}(A, X)$.

Completely bounded cohomology, together with its normal version, helped to achieve considerable advances in the two following problems about ordinary cohomology that, in their full generality, are still open.

PROBLEM 1 (raised by Kadison/Ringrose, cf. [180]). Let A be a von Neumann algebra. Is it true that $\mathcal{H}^n(A, A) = 0$ for all $n > 0$?

(Recall Theorem 1.2.7 which gives the positive answer for $n = 1$.)

PROBLEM 2 (raised by Christensen, Effros and Sinclair [20]). Let A be an operator C*-algebra on H. Is it true that $\mathcal{H}^n(A, \mathcal{B}(H)) = 0$ for all $n > 0$?

(Recall that the answer is not known even for $n = 1$; cf. the problem of Christensen, discussed in 1.2.1.)

Note that $\mathcal{H}^n(A, \mathcal{B}(H)) = 0$ implies $\mathcal{H}^n(A, B) = 0$ for any hyperfinite von Neumann algebra B, which contains A, and the same is true for other versions of cohomology groups. This follows from the existence of a B-bimodule projection, so-called conditional expectation, from $\mathcal{B}(H)$ onto B [212], and functorial properties of the cohomology functor $\mathcal{H}^n(A, ?)$.

As it was discovered, positive answers are ensured if we replace ordinary cohomology by its completely bounded version:

THEOREM 2.4.11 (Christensen and Sinclair [26]). *Let A be a von Neumann algebra. Then $\mathcal{H}^n_{\mathrm{cb}}(A, A) = \mathcal{H}^n_{\mathrm{cbw}}(A, A) = 0$ for all $n > 0$.*

THEOREM 2.4.12 (Christensen, Effros and Sinclair [20]). *Let A be an operator* C*-*algebra on H. Then, for all $n > 0$, $\mathcal{H}^n_{\mathrm{cb}}(A, B) = \mathcal{H}^n_{\mathrm{cbw}}(A, B) = 0$ for $B = \mathcal{B}(H)$ and hence for every hyperfinite von Neumann algebra B, containing A.*

Fortunately, "in most cases" all versions of cohomology groups do coincide. In what follows, a von Neumann algebra is said to be *good in its* II_1 *summand* if its canonical central direct summand of type II_1 is *-isomorphic to its von Neumann tensor product with the (unique) hyperfinite type II_1 factor (for example, it can have no summand of type II_1 or have, in the presence of such a summand, this hyperfinite factor).

THEOREM 2.4.13 (Christensen, Effros and Sinclair (idem)). *Let A be an operator* C*-*algebra such that A^- is good in its type* II_1 *summand, and let X be a normal completely bounded A^--bimodule* (and hence an A-bimodule). *Then, up to a topological isomorphism,*

$$\mathcal{H}^n(A, X) = \mathcal{H}^n_{\mathrm{cb}}(A, X) = \mathcal{H}^n_{\mathrm{w}}(A, X) = \mathcal{H}^n_{\mathrm{cbw}}(A, X).$$

As to the proof, the normal versions of the cohomology play the central role. The principal step is to establish the equality $\mathcal{H}^n_{\mathrm{w}}(A, X) = \mathcal{H}^n_{\mathrm{cbw}}(A, X)$; the rest follows from the theorem of Johnson, Kadison and Ringrose (see Theorem 2.4.8) and its "completely bounded" version (which can be proved by an argument closely resembling that in [116]).

When combined, the cited results immediately give

THEOREM 2.4.14.

(i) [26] *Let A be a von Neumann algebra, which is good in its type* II_1 *summand. Then $\mathcal{H}^n(A, A) = 0$ for all $n > 0$.*

(ii) [20] *Let A be an operator* C*-*algebra such that A^- is good in its type* II_1 *summand. Then, for all $n > 0$, $\mathcal{H}^n(A, B) = 0$ for $B = \mathcal{B}(H)$ and hence for every hyperfinite von Neumann algebra B, containing A.*

(Another condition, which gives a similar result, will be indicated in Theorem 2.4.23 below.)

Thus, the only possible obstruction to the vanishing of the discussed cohomology lies in the existence of "bad" type II_1 von Neumann algebras. However, something positive can be said even about these, if we content ourselves with the low dimensions. Recall that a maximal commutative selfadjoint subalgebra B of a von Neumann algebra A is called a *Cartan subalgebra* if the set of unitary elements u of A with the property $u^*Bu = B$ generates A as a von Neumann algebra.

THEOREM 2.4.15 (Christensen, Pop, Sinclair and Smith [22]). *Let A be a finite von Neumann algebra with a Cartan subalgebra. Then $\mathcal{H}^2(A, A) = 0$. If, an addition, A has a separable predual, then $\mathcal{H}^3(A, A) = 0$.*

Again the result is achieved by reducing to the completely bounded cohomology.

REMARK. Most of known examples of finite von Neumann algebras have Cartan subalgebras (cf. [204]; in particular, it is the case for the first example of a non-hyperfinite von Neumann algebra A with $\mathcal{H}^2(A, A) = 0$, due to Johnson [110]). However, quite recently Voiculescu has discovered that the group von Neumann algebra of F_2, the free group on two generators, has no Cartan subalgebra.

* * *

We turn to the cohomology of non-selfadjoint operator algebras and begin with the most explored of these, nest algebras. Here, instead of the open problems 1 and 2 (see above), we have a positive result, generalising the "one-dimensional" Theorem 1.2.6

THEOREM 2.4.16 (Nielsen [153]; Lance [140]). *Let A be a nest algebra, and X be an arbitrary ultraweakly closed sub-A-bimodule of $\mathcal{B}(H)$, containing A. Then $\mathcal{H}^n(A, X) = 0$ for all $n > 0$.*

This theorem was generalised to a certain class of CSL-algebras in [54].

(For another class of coefficients, providing the assertion of the cited theorem, see the next subsection.)

The argument in [140] is based on the following observation, which has an independent interest.

PROPOSITION 2.4.8 (compare Theorem 2.4.8). *Let A be an ultraweakly closed operator algebra on H such that the ultraweak closure of $A \cap \mathcal{K}(H)$ contains the identity operator. Then $\mathcal{H}^n(A, X) = \mathcal{H}^n_{\mathrm{w}}(A, X)$ for any normal A-bimodule and $n \geqslant 0$.*

We have already seen (Example 1.2.1) that for general CSL-algebras, the assertion of Theorem 2.4.14 is not valid even for the one-dimensional cohomology of algebras of finite (linear) dimension. (And therefore the cohomology groups, in principle, provide substantial invariants for algebras of this class.) We proceed to a general important construction, giving examples of non-trivial cohomology.

DEFINITION 2.4.13 (Gilfeather and Smith [57]). Let $\langle A \mid B \rangle$ be a norm closed unital operator algebra acting on a Hilbert space $\langle H \mid K \rangle$. Then the join of A and B is the operator algebra on $H \oplus K$ consisting of operators represented by block-matrices

$$\begin{bmatrix} b & 0 \\ c & a \end{bmatrix}$$

where $\langle a \mid b \mid c \rangle$ runs over $\langle A \mid B \mid \mathcal{B}(K, H) \rangle$. It is denoted by $A\#B$. The join $\langle A\#\mathbf{C} \mid A\#\mathbf{C}^2 \rangle$ is called the $\langle$ cone $\mid$ suspension $\rangle$ of A and is denoted by $\langle C(A) \mid S(A) \rangle$.

Note that, if A and B are CSL-algebras, then so is their join.

Here is, apparently, the main result concerning cohomology groups of joins. In what follows, for any unital operator algebra A, acting on H, we denote the space $\mathcal{H}^n(A, \mathcal{B}(H))$; where $n > 0$, by $\widetilde{\mathcal{H}}^n(A)$, and the space $\mathcal{H}^0(A, \mathcal{B}(H))/l.h.\{\mathbf{1}\}$ by $\widetilde{\mathcal{H}}^0(A)$ (we recall that the latter cohomology group is the commutant of A and thus contains the identity operator).

THEOREM 2.4.17 (idem). *Let A and B be as above and, in addition, either H or K is finite-dimensional. Then, up to a linear isomorphism,*

$$\widetilde{\mathcal{H}}^n(A\#B) = \bigoplus_{k=0}^{n-1} \big(\widetilde{\mathcal{H}}^k(A) \otimes \widetilde{\mathcal{H}}^{n-k-1}(B)\big)$$

for any $n \geqslant 1$. Moreover, for any A and B we have $\widetilde{\mathcal{H}}^0(A\#B) = 0$.

COROLLARY 2.4.4 [56]. *Let A be any norm closed operator algebra. Then $\widetilde{\mathcal{H}}^n(S(A)) = \widetilde{\mathcal{H}}^{n-1}(A)$ for $n \geqslant 1$ and $\widetilde{\mathcal{H}}^0(S(A)) = 0$. Besides, $\widetilde{\mathcal{H}}^n(C(A)) = 0$ for all $n \geqslant 0$.*

(As we shall see, the part, concerning cones, is also a corollary of Theorem 2.4.23 below.)

The cited result explains the terminology: operator suspensions and cones behave exactly as their prototypes in topology.

EXAMPLE 2.4.3. Take A as $S^n(\mathbf{C}^{k+1})$, where S^n is the suspension repeated n times. Then $\widetilde{\mathcal{H}}^n(A) = \mathbf{C}^k$, and $\widetilde{\mathcal{H}}^m(A) = 0$ for any $m \neq n$.

This example generalises and, to some extent, explains Example 1.2.1 (at least for $B = \mathcal{B}(H)$): one must only observe that the algebra, considered there, is just $S(\mathbf{C}^2)$.

Using such suspensions, one can construct CSL-algebras with arbitrary prescribed cohomology:

THEOREM 2.4.18 (Gilfeather and Smith [55]). *Let n_k, $k = 0, 1, 2, \ldots$, be a sequence of non-negative integers or ∞. Then there exists a* CSL-*algebra with* $\dim \widetilde{\mathcal{H}}^k(A) = n_k$ *for all k.*

The main technical tool, which helped the authors mentioned to construct the desired algebras from the "bricks" like $S^n(\mathbf{C}^{k+1})$ is

PROPOSITION 2.4.9 (idem). *Let $\mathcal{M}$ be the smallest class of operator algebras, containing commutative unital* C*-*algebras and closed under the operation of taking suspensions. Further, let A_k, $k = 1, 2, \ldots$, be in $\mathcal{M}$, and let A be the l_∞-direct sum of these algebras. Then*

$$\widetilde{\mathcal{H}}^n(A) = \bigoplus_{k=1}^{\infty} \widetilde{\mathcal{H}}^n(A_k) \quad \textit{for any } n \geqslant 1.$$

REMARK. As is mentioned in the cited papers, the ideas and results, connected with operator-theoretic suspensions and joins, were much influenced by the algebraic exploration of Kraus and Shack (alas, still not published). They discovered a relationship between the (Hochschild) cohomology of certain finite-dimensional operator algebras and the (topological) cohomology of certain simplicial complexes, naturally related to these algebras.

* * *

We spoke about cohomology with coefficients in $\mathcal{B}(H)$; much less is known, if we replace the latter bimodule by $\mathcal{K}(H)$. In particular, it is an open problem, whether $\mathcal{H}^n(A, \mathcal{K}(H)) = 0$ for an arbitrary von Neumann algebra, acting on H, and $n > 1$ (in other words, whether the "one-dimensional" result of Popa, cited in Subsection 1.2.1, can be extended to higher dimensions).

The strongest result in this direction is, as far as we know, due to Radulescu [175], who proved that $\mathcal{H}^2_{\mathrm{w}}(A, \mathcal{K}(H)) = 0$ for a finite countably decomposable von Neumann algebra A on H not containing a certain "bad" type II_1 factor as a direct summand. (The subscript w indicates that we mean the cohomology of the subcomplex in $\widetilde{C}(A, \mathcal{B}(H))$, consisting of normal cochains, taking values in $\mathcal{K}(H)$.)

2.4.5. *(Co)homology as derived functors: principal facts and applications.* Many of the results already presented on the computation of cohomology groups were obtained by methods working with cocycles in a suitable version of the standard complex. Now we turn to results, which were mostly obtained in the framework of another approach. This approach has its origin in the work of the founding fathers of homological algebra – Cartan, Eilenberg, MacLane [16,149], and also Hochschild [104]. It is based on the treating cohomology, as well as homology, groups of algebraic systems as particular cases of the unifying notion of a derived functor. Such an approach, properly carried over to "algebras of analysis", saves one from being tied to standard complexes. Instead, one can use suitable chosen (co)resolutions and the machinery of long exact sequences.

The key fact is

THEOREM 2.4.19. *Let A be a Banach algebra, and X be a Banach A-bimodule. Then, up to a topological isomorphism,*

$$\mathcal{H}^n(A,X) = {}_A\operatorname{Ext}^n_A(A_+,X) \quad \textit{and} \quad \mathcal{H}_n(A,X) = {}^A\operatorname{Tor}^A_n(X,A_+)$$

for all $n \geqslant 0$. If, in addition, A and X are unital, then we have also

$$\mathcal{H}^n(A,X) = {}_A\operatorname{Ext}^n_A(A,X) \quad \textit{and} \quad \mathcal{H}_n(A,X) = {}^A\operatorname{Tor}^A_n(X,A).$$

A similar assertion ia valid for $\widehat{\otimes}$- and $\overline{\otimes}$- (in particular, Fréchet) algebras and respective classes of coefficients; the only difference is that the sign of equality should be understood in these cases as linear isomorphism.

To prove the first two identities, one must just compute the mentioned $\langle \mathrm{Ext}|\mathrm{Tor}\rangle$ by means of the normalised bimodule bar-resolution $0 \leftarrow A_+ \xleftarrow{\pi} \mathbf{B}(A_+)$, introduced in 2.2.1: the emerging complex $\langle {}_Ah_A(\mathbf{B}(A_+),X) \mid X \widehat{\otimes}_A \mathbf{B}(A_+)\rangle$ turns out to be isomorpic to the standard complex $\widetilde{C}(A,X) \mid \underset{\sim}{C}(A,X)\rangle$. The remaining identities can be established by the similar use of the unital version of $\mathbf{B}(A_+)$, also mentioned in 2.2.1.

As the first application, this theorem provides a characterization of contractible and (Johnson) amenable algebras, extending that in Theorem 1.3.1 and 1.3.4. Indeed, combining Theorem 2.4.19 with just mentioned results and with Theorem 2.3.6 and 2.3.10, we have

THEOREM 2.4.20. *Let A be a Banach algebra. Then*

(i) *A is contractible iff $\mathcal{H}^n(A,X) = 0$ for all Banach A-bimodules X and any $n > 0$,*
(ii) *A is amenable iff $\mathcal{H}^n(A,X) = 0$ for all dual Banach A-bimodules X and any $n > 0$,*
(iii) *A is amenable iff $\mathcal{H}_n(A,X) = 0$ for all Banach A-bimodules X and any $n > 0$, and the space $\mathcal{H}_0(A,X)$ is Hausdorff.*

The assertions (i) and (iii) remain true in the wider context of $\widehat{\otimes}$- and $\overline{\otimes}$-algebras, and respective classes of coefficients.

The second major application is based on employing, for the computation of derived functors, the entwining bimodule resolution of A_+ (see again 2.2.1). Since this resolution has the length 2, Theorem 2.4.19, combined with Propositions 2.2.6 and 2.3.3, immediately implies

THEOREM 2.4.21 cf. [85].

(i) *If A is a biprojective $\widehat{\otimes}$-algebra, then $\mathcal{H}^n(A,X) = 0$ for all Banach A-bimodules X and $n \geqslant 3$.*
(ii) *If A is a biflat Banach algebra, then $\mathcal{H}^n(A,X) = 0$ for all dual Banach A-bimodules X and $n \geqslant 3$. Moreover, $\mathcal{H}_n(A,X) = 0$ for all Banach A-bimodules X and $n \geqslant 3$, and the space $\mathcal{H}_2(A,X)$ is Hausdorff.*

We see, in particular, that such algebras as l_1, $L^1(G)$ for a compact G, $\mathcal{N}(E)$ for E with the approximation property, and $E \mathbin{\widehat{\otimes}} E^*$ for any E (see 1.3.3) have the vanishing cohomology groups with arbitrary coefficients, beginning with the dimension 3.

In some important cases, the ⟨ cohomology | homology ⟩ groups coincide with certain one-sided, and not only with two-sided, as above, ⟨ Ext | Tor ⟩ spaces:

THEOREM 2.4.22. *Let A be a Banach algebra. Then for any* ⟨ $X \in$ **A-mod** | **mod-A** ⟩ *and $Y \in$ **A-mod**, we have, up to a topological isomorphism,*

$$\big(\mathcal{H}^n\big(A, \mathcal{B}(X, Y)\big) = {}_A\operatorname{Ext}^n(X, Y) \,|\, \mathcal{H}_n(A, X \widehat{\otimes} Y) = \operatorname{Tor}_n^A(X, Y)\big)$$

for all $n \geqslant 0$.

The part, concerning homology and Tor, is valid for arbitrary $\widehat{\otimes}$-algebras and $\widehat{\otimes}$-modules (and also for $\overline{\otimes}$-algebras and $\overline{\otimes}$-modules, if we replace the symbol "$\widehat{\otimes}$" by $\overline{\otimes}$).

The indicated identities become evident if we notice that the respective standard ⟨ cohomology | homology ⟩ complex is isomorphic to the respective standard ⟨ Ext | Tor ⟩ computing complex.

As an immediate consequence, we get a "high-dimensional" extension of the criterion, presented in the Corollary 1.2.2:

COROLLARY 2.4.5. *Let A be a Banach algebra. A Banach left A-module ⟨$X|Y|X$⟩ is ⟨ projective | injective | flat ⟩ iff $\mathcal{H}^n(A, \mathcal{B}(X, Y)) = 0$ for ⟨ any | any | any dual ⟩ Banach left A-module ⟨$Y \mid X \mid Y$⟩ and any $n > 0$. Furthermore, a Banach ⟨ right | left ⟩ A-module ⟨$X \mid Y$⟩ is flat iff the space $\mathcal{H}_0(A, X \widehat{\otimes} Y)$ is Hausdorff, and $\mathcal{H}_n(A, X \widehat{\otimes} Y) = 0$ for any Banach ⟨ left | right ⟩ A-module ⟨$Y \mid X$⟩ and any $n > 0$.*

A particular case of this corollary deserves to be distinguished:

THEOREM 2.4.23. *Let A be a spatially ⟨ projective | flat ⟩ operator algebra on a ⟨ Banach | reflexive Banach ⟩ space E. Then $\mathcal{H}^n(A, \mathcal{B}(E)) = 0$ for any $n > 0$ and, more generally, $\mathcal{H}^n(A, \mathcal{B}(E, X)) = 0$ for ⟨ all | all dual ⟩ left Banach A-modules X and any $n > 0$.*

Combining this theorem with the information about spatially projective and spatially flat Banach algebras presented in 1.1, we obtain examples of algebras with vanishing "spatial cohomology groups" $\mathcal{H}^n(A, \mathcal{B}(E))$ and their generalizations $\mathcal{H}^n(A, \mathcal{B}(E, X))$ (for all or at least for all dual $X \in$ **A-mod**). The class of such operator algebras includes:

- algebras, containing a column of rank-one operators (compare the "one-dimensional" case, presented as Theorem 1.3.7; among them we have $\mathcal{B}(E)$, $\mathcal{K}(E)$, $\mathcal{N}(E)$ and, obviously, cones of arbitrary norm closed operator algebras on a Hilbert space (cf. Corollary 2.4.4);
- Connes (= hyperfinite von Neumann) algebras;
- nest algebras (as well as their generalizations in [54]).

Here is one more

EXAMPLE 2.4.4. Let A be an operator C*-algebra on H such that A_+ has an algebraically cyclic vector. That means (see Proposision 0.2.1) that the A-module H is isomorphic to some A_+/I where I is a closed left ideal in A_+. Since left ideals in C*-algebras have right bounded approximate identities, A_+/I and hence H are flat (Theorem 1.1.5(ii)). Thus $\mathcal{H}^n(A, \mathcal{B}(H, X)) = 0$ for all dual $X \in$ **A-mod**.

* * *

The case of dual bimodules provides additional possibilities for computing cohomology groups. Combining Theorem 2.4.19 with Proposition 2.3.4, we obtain

THEOREM 2.4.24. *Let A be a Banach algebra, and $X = (X_*)^*$ be a dual Banach A-bimodule. Then, up to a topological isomorphism, $\mathcal{H}^n(A, X) = {}_A\mathrm{Ext}^n_A(X_*, A^*_+)$. Moreover, in the case of unital A and X, $\mathcal{H}^n(A, X) = {}_A\mathrm{Ext}^n_A(X_*, A^*)$, $n \geqslant 0$.*

What is, perhaps, more interesting is that the normal cohomology groups of operator C*-algebras, although defined in terms of non-normed topology, can also be expressed in the language of the standard ("Banach") Ext spaces, but with somewhat different variables:

THEOREM 2.4.25 (cf. [91]). *Let A be an operator C*-algebra, and $X = (X_*)^*$ be a dual Banach A-bimodule. Then, up to a topological isomorphism,*

$$\mathcal{H}^n_{\mathrm{w}}(A, X_{\mathrm{w}}) = {}_A\mathrm{Ext}^n_A(X_*, A^-_*), \quad n \geqslant 0.$$

We recall that X_{w} is the normal part of X whereas A^-_* is the predual bimodule to A^-, the ultraweak closure of A (see 0.3).

The proof is based on the computation of the indicated Ext by means of the injective coresolution $0 \longrightarrow A^-_* \xrightarrow{\pi_*} \tilde{\beta}_*(A)$ (see Proposition 2.2.5). The emerging complex ${}_Ah_A(X_*, \tilde{\beta}_*(A))$ turns out to be isomorphic to $\tilde{C}_{\mathrm{w}}(A, X_{\mathrm{w}})$, and the theorem follows. (Similarly, Theorem 2.4.24 could also be obtained by means of the coresolution $0 \longrightarrow A^* \xrightarrow{\pi^*} \tilde{\beta}^*(A)$; mentioned in Proposition 2.2.4.)

But what if we compute the same Ext with the help of other coresolutions mentioned in the same Proposition 2.2.4? Then we get complexes ${}_Ah_A(X_*, \tilde{\beta}_k(A))$, $k = 1, 2$, which, as one can show, are isomorphic respectively to the complexes $\tilde{C}(A, X_l)$ and $\tilde{C}(A, X_r)$. Thus Theorem 2.4.25 implies Theorem 2.4.8, formulated earlier.

Theorem 2.4.25 also plays the principal role in the proving Theorem 1.3.9 (about the equivalence of two approaches to the notion of a Connes amenable algebra). Since the injectivity of the second argument implies the vanishing of respective Ext^n for all $n > 0$, the latter theorem can now be presented in the following extended form:

THEOREM 2.4.26. *Let A be an operator C*-algebra. Then the following properties are equivalent (and hence, any of them can be taken as a definition of a Connes amenable algebra):*

(i) $\mathcal{H}^1(A, X) = 0$ *for all normal A-bimodules X,*
(ia) $\mathcal{H}^1_{\mathrm{w}}(A, X) = 0$ *for all normal A-bimodules X,*
(ib) $\mathcal{H}^1_{\mathrm{w}}(A, X_{\mathrm{w}}) = 0$ *for all dual A-bimodules X,*
(ii) $\mathcal{H}^n(A, X) = 0$ *for all normal A-bimodules X and $n > 0$,*

(iia) $\mathcal{H}^n_{\mathrm{w}}(A, X) = 0$ *for all normal* A*-bimodules* X *and* $n > 0$,
(iib) $\mathcal{H}^n_{\mathrm{w}}(A, X_{\mathrm{w}}) = 0$ *for all dual* A*-bimodules* X *and* $n > 0$,
(iii) *the Banach* A*-dimodule* A^-_* *is injective.*

By virtue of Theorem 1.3.10, the indicated properties exactly mean, in intrinsic of operator-algebraic terms, that A^- is hyperfinite (or, equivalently, "Connes injective"). We mention that the implication "hyperfiniteness $\Rightarrow$ (ii), (iia)" was originally proved in [116], and "(i) $\Rightarrow$ Connes injectivity" in [29].

As to completely bounded cohomology groups, they still wait to be interpreted in terms of a suitable derived functor.[13] It seems probable that the recent modification of these groups, mentioned by Effros and Ruan [46] as "matricially bounded cohomology" will happen to be fairly tractable, and that the necessary ingredients (free modules, bar-resolution etc.) will be defined with the help of so-called "operator projective tensor product". This very promising type of the tensor product of operator spaces was recently introduced by Effros and Ruan (idem) and, independently, Blecher and Paulsen [10].

2.5. *Homological dimensions*

2.5.1. *General definitions and properties.* To speak informally, the homological dimension of a module is an integer (or ∞) which measures how far this module is "homologically bad". There are several versions of this notion, depending on which modules we want to consider as best. The most explored characteristic arises if we start from the class of projective modules.

DEFINITION 2.5.1. Let X be a $\langle$ left | right | two-sided $\rangle$ $\widehat{\otimes}$-module over an $\widehat{\otimes}$-algebra A; then the length of its shortest projective resolution is called its projective homological dimension and is denoted by $\langle {}_A\mathrm{p.dh}\, X \mid \mathrm{p.dh}_A\, X \mid {}_A\mathrm{p.dh}_A\, X \rangle$.

When there is no danger of confusion, we say just "*homological dimension*" and omit "p." in the respective notation.

In more traditional language, the homological dimension of X is obviously the least n for which X can be represented as $P_0/(P_1/\ldots/(P_{n-1}/P_n)\ldots)$ where all modules are projective and every P_k is complemented as a subspace of P_{k-1}. In particular, it is equal to 0 *just when* X *is projective.*

In fact, projective homological dimension can be similarly defined in any category of topological modules such that its objects have projective resolutions. Thus, we could replace the symbol "$\widehat{\otimes}$" by "$\overline{\otimes}$" or to restrict ourselves to Banach or Fréchet modules.

In the next and subsequent propositions of this subsection we consider, to be definite, the case of left modules; these propositions have obvious analogues for other types of modules.

[13] These lines were written in Spring of 1995 in Moscow. As it turned out, by that time the following discovery was already made on the opposite side of the globe. Vern Paulsen has succeeded to express the spaces $\mathcal{H}^n_{\mathrm{cb}}(\cdot, \cdot)$ in terms of a proper "completely bounded" version of the spaces $\mathrm{Ext}^n(\cdot, \cdot)$. The latter, however, were defined not by means of derived functors, but as the spaces of equivalence classes of n-multiple extensions in the spirit of Yoneda (cf. 2.3.2 above). The achieved overview enabled Paulsen to give a new elegant proof of Theorems 2.4.11 and 2.4.12. See these striking results in his forthcoming paper "Relative Yoneda cohomology for operator spaces".

By virtue of Theorem 2.3.1, the given definition implies the following

PROPOSITION 2.5.1. *Let* ${}_A\operatorname{dh} X < k$. *Then for any additive functor* F *from the category of left* $\widehat{\otimes}$*-modules over* A *to* **TVS** *we have* $F_k(X) = 0$ (*or, respectively,* $F^k(X) = 0$).

For one of derived functors, the indicated property provides alternative approaches to the notion of the homological dimension:

PROPOSITION 2.5.2. *The following properties of* X *are equivalent:*

(i) ${}_A\operatorname{dh} X \leqslant n$;
(ii) ${}_A\operatorname{Ext}^{n+1}(X, Y) = 0$ *for any relevant* Y;
(iii) *if the complex of the form*

$$0 \leftarrow X \leftarrow P_0 \leftarrow \cdots \leftarrow P_{n-1} \leftarrow Q \leftarrow 0 \qquad (\mathcal{P})$$

is a resolution of X *in which* $P_0, \ldots, P_{n-1}$ *are projective, then* Q *is also projective.*

(The proof is an easy consequence of Theorems 2.3.5 and 2.3.6.)

Note that we can now remove an apparent ambiguity in the definition of ${}_A\operatorname{dh} X$ for a Banach module X. In fact, if the latter has a certain projective resolution, consisting of $\widehat{\otimes}$-modules, then it has a projective resolution of the same length consisting only of Banach modules. Indeed, we can take any of existing resolutions, consisting of Banach modules (say, the bar-resolution), and apply the equivalence of (i) and (iii) above. (The same is obviously true for Fréchet, instead of Banach, modules as well.)

Further, a similar argument shows that if X is a unital module over a unital algebra, then ${}_A\operatorname{dh} X$ can be defined as the length of the shortest projective resolution of X, consisting of unital modules.

Proposition 2.5.2, together with the Theorem 2.3.5, implies:

PROPOSITION 2.5.3. *Let* $0 \leftarrow X'' \leftarrow X \leftarrow X' \leftarrow 0$ *be an admissible exact complex of* A*-modules. Then*

(i) ${}_A\operatorname{dh} X \leqslant \sup\{{}_A\operatorname{dh} X'', {}_A\operatorname{dh} X'\}$. *Moreover, equality certainly holds if the complex splits* (*that is,* $X = X' \oplus X''$);
(ii) ${}_A\operatorname{dh} X'' \leqslant \sup\{{}_A\operatorname{dh} X, {}_A\operatorname{dh} X' + 1\}$. *Moreover, if* X *is projective and the complex does not split, then* ${}_A\operatorname{dh} X'' = {}_A\operatorname{dh} X' + 1$;
(iii) ${}_A\operatorname{dh} X' \leqslant \sup\{{}_A\operatorname{dh} X, {}_A\operatorname{dh} X'' - 1\}$.

Now we mention several examples of computing homological dimension.

EXAMPLE 2.5.1. Let Ω, Δ, $A = C(\Omega)$ and $X = C(\Delta)$ be as in Example 2.2.1. Then the form of the resolution, presented in that example, implies that

$${}_A\operatorname{dh} X = \begin{cases} 0 & \text{if } \Omega \setminus \Delta \text{ is closed,} \\ 1 & \text{otherwise.} \end{cases}$$

In particular, for the point A-module $\mathbf{C}_s$, $s \in \Omega$ we have ${}_A\operatorname{dh} \mathbf{C}_s = 0$ if s is isolated in Ω, and ${}_A\operatorname{dh} \mathbf{C}_s = 1$ otherwise.

EXAMPLE 2.5.2. Let A be $\mathbf{C}$ with zero multiplication. Then, considering the resolution

$$0 \leftarrow A \xleftarrow{\pi} A_+ \xleftarrow{d} A_+ \xleftarrow{d} A_+ \leftarrow \cdots$$

of the A-module A, where π and d assign the usual product λz to $\lambda e + z$, we see that ${}_A\operatorname{dh} A = \infty$.

EXAMPLE 2.5.3. Let $A = \mathcal{O}(\mathbf{C}^n)$ and $\mathbf{C}_0$ be $\mathbf{C}$ with the outer multiplication $w \cdot z = w(0)z$. Then, as a particular case of Theorem 2.2.4 (i), the latter module has the Koszul resolution $0 \leftarrow \mathbf{C}_0 \xleftarrow{\pi} \mathcal{K}(A, \Sigma_1, \dots, \Sigma_n)$ where $\Sigma_m : A \to A$ takes w to $z_m w$. By means of this we see that ${}_A\operatorname{Ext}^n(\mathbf{C}_0, \mathbf{C}_0) = \mathbf{C} \neq 0$ and hence ${}_A\operatorname{dh} \mathbf{C}_0 = n$.

More sophisticated results on computing the homological dimension of concrete modules will be presented later; see Theorems 2.5.1, 2.5.10–2.5.13 and Proposition 2.5.11.

* * *

To consider the "injective" counterpoint of Definition 2.5.1, one naturally must require that our modules have injective coresolutions. Therefore we restrict ourselves to Banach structures, where this is certainly the case (see 2.2.1). Within this framework, the length of the shortest injective coresolution of a given A-module X is called its injective homological dimension and it is denoted, depending on a type of our module, by ${}_A\operatorname{i.dh} X$, $\operatorname{i.dh}_A X$ or ${}_A\operatorname{i.dh}_A X$.

Propositions 2.5.1–2.5.3 have obvious "injective" analogues; in particular, ${}_A\operatorname{i.dh} X \leqslant n$ iff ${}_A\operatorname{Ext}^{n+1}(Y, X) = 0$ for all relevant (now Banach) Y. Note that the latter equality makes sense outside the "Banach" context (see 2.3.2) and thus provides a way to define injective homological dimension, say for Fréchet modules etc.[14]

Finally, the *flat* or, as is more frequently said, *weak homological dimension* of a $\widehat{\otimes}$-module X over an $\widehat{\otimes}$-algebra A can be initially defined as the length of the shortest flat resolution of this module; it is denoted respectively by ${}_A\operatorname{w.dh} X$, $\operatorname{w.dh}_A X$ or ${}_A\operatorname{w.dh}_A X$. Alternative definitions are provided in the next proposition.

PROPOSITION 2.5.4. *Let X be a left Banach module over a Banach algebra A. Then the following properties are equivalent:*

(i) ${}_A\operatorname{w.dh} X \leqslant n$.
(ii) $\operatorname{Tor}^A_{n+1}(Y, X) = 0$, *and* $\operatorname{Tor}^A_n(Y, X)$ *is Hausdorff for any (right)* Y.
(iii) *if the complex of the form* $(\mathcal{P})$ *as in Proposition* 2.5.2(iii) *is a resolution of X in which $P_0, \dots, P_{n-1}$ are flat, then Q is also flat.*
(iv) $\operatorname{i.dh}_A X^* \leqslant n$.
(v) ${}_A\operatorname{Ext}^{n+1}(X, Y) = 0$ *for any dual* Y.
(vi) *the same as* (iii), *but with "preresolution" instead of "resolution".*

(The part of this proposition concerning first three properties is valid also in the context of Fréchet modules.)

[14] These dimensions are computed for many (non-Banach) Fréchet algebras by Pirkovskii.

Thus, in the Banach context ${}_A\,\mathrm{w.dh}\,X$ is just $\mathrm{i.dh}_A\,X^*$, and it is the length of the shortest flat preresolution of X. As in the case of projective dimension, the weak dimension of a Banach or Fréchet module can be defined in terms of flat resolutions consisting of modules of the respective category. Similarly, in the unital case, we can define both injective and weak dimensions in terms of respective (co)resolutions, consisting of unital modules.

Needless to say, $\langle {}_A\,\mathrm{i.dh}\,X \mid {}_A\,\mathrm{w.dh}\,X \rangle = 0$ means just that X is $\langle$ injective | flat $\rangle$.

* * *

Now, starting from modules, we introduce a few numerical characteristics of the algebras themselves.

DEFINITION 2.5.2. Let A be an $\widehat{\otimes}$-algebra. The number (or ∞) $\sup_A \mathrm{dh}\,X$, where X runs over all left $\widehat{\otimes}$-modules over A, is called the *left global homological dimension of* A. It is denoted by $\mathrm{dgl}\,A$, or, more frequently, just by $\mathrm{dg}\,A$.

In a parallel way, one can define the right and the two-sided global homological dimension of A; it is easy to see that the first one is just $\mathrm{dg}\,A^{\mathrm{op}}$ whereas the second is $\mathrm{dg}\,A^{\mathrm{env}}$. Below, we shall simply say "global dimension of A" meaning $\mathrm{dg}\,A$.

The "flat" version of $\mathrm{dg}\,A$, called the *weak global dimension* of A and denoted by $\mathrm{w.dg}\,A$, is defined as $\sup_A \mathrm{w.dh}\,X$,with X running over the same class of modules.

It would seem that our terminology is not precise and could lead to a confusion because we did not say "*projective* left global dimension" and "*left* weak global dimension". As a matter of fact, the words in italics were omitted deliberately. The reason is that, introducing the "injective" version of dg and the "right" version of w.dg, we would not get new characteristics:

PROPOSITION 2.5.5. *Let A be a Banach algebra. Then*

(i) $\mathrm{dg}\,A = \sup\{{}_A\,\mathrm{i.dh}\,Y\colon\ Y \in \mathbf{A\text{-}mod}\} = \min\{n\colon\ {}_A\,\mathrm{Ext}^{n+1}(X, Y) = 0$ *for all* $X, Y \in \mathbf{A\text{-}mod}\}$*;*

(ii) $\mathrm{w.dg}\,A = \sup\{\mathrm{w.dh}_A\,Y\colon\ Y \in \mathbf{mod\text{-}A}\} = \min\{n\colon\ \mathrm{Tor}^A_{n+1}(Y, X) = 0,$ *and* $\mathrm{Tor}^A_n(Y, X)$ *is Hausdorff for all* $X \in \mathbf{A\text{-}mod}$, $Y \in \mathbf{mod\text{-}A}\}$.

The assertion (ii) is also valid in larger classes of Fréchet algebras.

In particular, we see that $\mathrm{dg}\,A = 0$ means that all left A-modules are projective and, which is equivalent, that all of them are injective. At the same time $\mathrm{w.dg}\,A = 0$ means that all left and, which is equivalent, all right A-modules are flat. We do not know at the moment, whether $\mathrm{dgl}\,A$ is always equal to "$\mathrm{dgr}\,A$" (cf. below).

Now we turn to another group of numerical invariants of algebras, which are most important from the point of view of (co)homology groups.

DEFINITION 2.5.3. Let A be a $\widehat{\otimes}$-algebra. The least n for which

$$\left\langle \mathcal{H}^{n+1}(A, X) = 0 \,\middle|\, \mathcal{H}_{n+1}(A, X) = 0 \text{ and } \mathcal{H}_n(A, X) \text{ is Hausdorff} \right\rangle$$

for all $\widehat{\otimes}$-bimodules X over A, or ∞ if there is no such n, is called the $\langle$ *homological* | *weak homological* $\rangle$ *bidimension of* A, and it is denoted by $\langle\, \mathrm{db}\, A \mid \mathrm{w.db}\, A\, \rangle$.

Combining Propositions 2.5.2 and 2.5.4 with the expression of (co)homology in terms of derived functors (see 2.4.5), we immediately get

PROPOSITION 2.5.6. *For any Banach algebra* A, $\langle\, \mathrm{db}\, A \mid \mathrm{w.db}\, A\, \rangle$ *coincides with* $\langle\, {}_A \mathrm{dh}_A\, A_+ \mid {}_A \mathrm{w.dh}_A\, A_+\, \rangle$ *and, moreover* $\mathrm{w.db}\, A$ *coincides also with the least n for which* $\mathcal{H}^{n+1}(A, X) = 0$ *for all dual Banach A-modules X.*

Evidently, $\langle\, \mathrm{db}\, A \mid \mathrm{w.db}\, A\, \rangle = 0$ means just that A is $\langle$ contractible | (Johnson) amenable $\rangle$. In this connection, algebras with $\mathrm{w.db}\, A \leqslant n - 1$ are sometimes referred as n-amenable algebras (cf. [163]). Besides, some authors, saying "the cohomological dimension of A", mean $\mathrm{db}\, A$.

If A is an operator C*-algebra, one can also define its *normal*, or *ultraweak homological bidimension* "n.db A" as the least n for which $\mathcal{H}^{n+1}(A, X) = 0$ for all normal A-bimodules X. It is true, however, that our knowledge about this quantity almost confines to the banality that it is equal to zero iff A is Connes amenable.

It is easy to see that all of the introduced dimensions do not change if we pass from an algebra to its unitization. Thus, $\mathrm{dg}\, A = \mathrm{dg}\, A_+$, $\mathrm{w.db}\, A = \mathrm{w.db}\, A_+$ etc.

Certainly we always have $\mathrm{w.dg}\, A \leqslant \mathrm{dg}\, A$ and $\mathrm{w.db}\, A \leqslant \mathrm{db}\, A$, and rather often strict inequalities occur. One of the more interesting things is:

PROPOSITION 2.5.7. *If A is an* $\widehat{\otimes}$*-algebra, then*

(i) $\mathrm{dg}\, A \leqslant \mathrm{db}\, A$.

If, moreover, it is a Fréchet algebra, then

(ii) $\mathrm{w.dg}\, A \leqslant \mathrm{w.db}\, A$.

In the Banach case,the indicated inequalities immediately follow from the formulae, presented in Theorem 2.4.22, and the same is true for (ii) in the general case. As to the proof of (i) in the general case, see, e.g., [96].

As a matter of fact, for all $\widehat{\otimes}$-algebras for which both $\mathrm{dg}\, A$ and $\mathrm{db}\, A$ have been computed, they coincide and the same is true for the pair of weak dimensions (see below). It is an open question, whether the equalities $\mathrm{dg}\, A = \mathrm{db}\, A$ and $\mathrm{w.dg}\, A = \mathrm{w.db}\, A$ are always valid in the class of $\widehat{\otimes}$-, at least Banach algebras. (We do not even know, whether $\langle\, \mathrm{dg}\, A = 0 \mid \mathrm{w.dg}\, A = 0\, \rangle$ implies $\langle\, \mathrm{db}\, A = 0 \mid \mathrm{w.db}\, A = 0\, \rangle$. This was discussed, in equivalent terms, in 1.3.) Since we always have $\mathrm{db}\, A = \mathrm{db}\, A^{\mathrm{op}}$, every hypothetical algebra A with $\mathrm{dg}\, A < \mathrm{dg}\, A^{\mathrm{op}}$ would have $\mathrm{dg}\, A < \mathrm{db}\, A$. In pure algebra such things are possible (Kaplansky [129]).

Finally note that Theorem 2.4.21 can be reformulated as follows.

PROPOSITION 2.5.8.

(i) *If A is a biprojective Banach algebra, then* $\mathrm{db}\, A \leqslant 2$ *and hence* $\mathrm{dg}\, A \leqslant 2$.

(ii) *If A is a biflat Banach algebra, then* $\mathrm{w.db}\, A \leqslant 2$, *and hence* $\mathrm{w.dg}\, A \leqslant 2$.

In fact, the first assertion is valid for all biprojective $\widehat{\otimes}$-algebras.

2.5.2. *Peculiarities of Banach structures: forbidden values and additivity formulae.* We proceed to describe a phenomenon which is specific for homological dimensions of Banach algebras. It has no analogue for abstract algebras and at the same time for non-normed topological algebras, even for metrizable Arens–Michael algebras, which appear to be so close to Banach algebras.

We begin with a certain concrete computation, serving as a core for the subsequent results of a general character. Consider c_b as a Banach module over either of c_0 and l_1, with coordinatewise outer multiplication.

THEOREM 2.5.1. *For both $A = c_0$ or l_1, we have ${}_A \operatorname{dh} c_b = 2$.*

Why it is so? The explanation is that the entwining resolution of c_b, being now projective, provides ${}_A \operatorname{dh} c_b \leqslant 2$. The main part of the argument is to establish the opposite inequality or, equivalently, to show that the morphism Δ in the mentioned resolution has no left inverse (continuous) morphism. It is rather easy for $A = l_1$ ("the crude case") because of the well-known structure of tensor products by l_1. The proof for $A = c_0$ ("the delicate case") is somewhat more complicated. Here, playing on the absence of a Banach complement of c_0 in c_b, one gets that the hypothetical left inverse of Δ must take some bounded sets to sets in $c_0 \widehat{\otimes} c_0$, containing the so-called n-triangular elements for any $n = 1, 2, \dots$. The latter are defined in such a way that their norms are equal to those of the elements $u_n \in \mathbf{C}^n \widehat{\otimes} \mathbf{C}^n$ depicted by triangular matrices with 1 above the diagonal. Since it is known that $\|u_n\| \to \infty$ as $n \to \infty$ [84,108]; our hypothetical morphism can not be continuous.

Note that the formulated equalities already contradict the purely algebraic experience. Indeed, a well known theorem of Auslander (see, e.g., [16]) states that, in the framework of abstract algebra, $\operatorname{dg} A$ can be computed as $\sup_A \operatorname{dh} X$, where X runs over only cyclic modules. However, for the Banach algebras $A = c_0$ and l_1 the latter quantity is certainly 1 and thus must be strictly less than $\operatorname{dg} A$.

Theorem 2.5.1 has a "non-commutative" version. In fact, using similar ideas, one can prove that for $A = \mathcal{K}(H)$ or $\mathcal{N}(H)$, we have ${}_A \operatorname{dh} \mathcal{B}(H) = 2$.

Now we turn to more general results, demonstrating the peculiar behaviour of "Banach" homological dimensions. Here is the oldest of them, often referred as "the global dimension theorem".

THEOREM 2.5.2 [85]. *Let A be a commutative Banach algebra with an infinite Gel'fand spectrum, in particular, an infinite-dimensional function algebra. Then $\operatorname{dg} A \geqslant 2$, and* (hence) $\operatorname{db} A \geqslant 2$.

The part, concerning db, was already presented as Theorem 2.4.5, and thus it implies Corollary 2.4.3 (concerning the existence of non-splitting singular extensions for all infinite-dimensional Banach function algebras). Recall (see 2.4.2) that the indicated estimates are something definitely exotic from the point of view of pure homological algebra, and in that realm the class of completely segregated algebras (which now could be defined as those with $\operatorname{db} A \leqslant 1$) contains quite a few infinite-dimensional commutative semisimple algebras. Note that for all three concrete pure algebras mentioned in 2.4.2 (the polynomial algebra, the group algebra of $\mathbf{Z}$ and c_{00}) we have $\operatorname{dg} A = \operatorname{db} A = 1$.

The proof of the global dimension theorem (see [86] and [169] for a detailed exposition) absorbs various bits of information about Banach structures. It relies on some specific features of Banach geometry (the possible absence of Banach complements and norm estimates of some elements in Varopoulos algebras) which were already mentioned in connection with Theorem 2.5.1. Apart from this, the existence of the Shilov boundary as the subset, to speak informally, of the "non-analycity" of Gel'fand transforms also plays an important role (suggested by Theorem 1.1.2(i) which is essentially used in the proof).

Where to find X with ${}_A\operatorname{dh} X \geqslant 2$? The following statement provides, in fact, a detailed formulation of Theorem 2.5.2.

THEOREM 2.5.3. *Let A be as above, Ω_+ be the spectrum of A_+ and $\partial\Omega_+$ be the Shilov boundary of Ω_+. Then at least one of the following two possibilities occurs: either* (i) *there is $s \in \partial\Omega_+$ such that $\Omega_+ \setminus \{s\}$ is not paracompact, or* (ii) *there is a converging sequence $\mathbf{s} = \{s_n\}$ of pairwise different points of $\partial\Omega_+$. If* (i) *holds, then ${}_A\operatorname{dh} \mathbf{C}_s \geqslant 2$* (cf. Example 0.2.1). *If* (ii) *holds, then $\mathbf{s}$ has a subsequence $\mathbf{t} = \{t_n\}$ such that for the respective A-module $c_{\mathrm{b}} = c_{\mathrm{b}}(\mathbf{t})$* (cf. Example 0.2.2) *we have ${}_A\operatorname{dh} c_{\mathrm{b}} \geqslant 2$.*

The assertion in case (i) is an obvious corollary of Theorem 1.1.1. In case (ii), taking a sufficiently sparse subsequence $\mathbf{t}$ of $\mathbf{s}$ and using standard homological techniques, one can manage to show that ${}_A\operatorname{dh} c_{\mathrm{b}}(\mathbf{t}) \geqslant 2$ provided ${}_B\operatorname{dh} c_{\mathrm{b}}(\mathbf{t}) \geqslant 2$ where B is the quotient algebra A_+/I with $I = \{a\colon a(t_n) = 0, \text{ for all } n = 1, 2, \ldots\}$. Finally, continuing the lacunization, one ensures that B behaves, to speak informally, "either as c_0 or as l_1" and reduces the desired estimate to the consideration of the two model examples, presented in Theorem 2.5.1.

It turned out that phenomena, similar to the global dimension theorem, frequently happen in other "natural" classes of Banach algebras. In particular, the estimate $\operatorname{dg} A \geqslant 2$ is valid for all infinite-dimensional CCR-algebras (Lykova [148]). The next result actually claims that the same estimate occurs for "the majority" of separable C*-algebras.

Let I be a two-sided ideal in an algebra A. Recall that it is called *complementable as a subalgebra* if there exists another two-sided ideal J in A such that $A = I \oplus J$.

THEOREM 2.5.4 (Aristov [4]). *Let A be a separable* C*-*algebra without an identity, or, more generally, such that there exists in A_+ a closed two-sided ideal of finite codimension, which is not complementable as a subalgebra. Then* $\operatorname{dg} A \geqslant 2$, *and hence* $\operatorname{db} A \geqslant 2$. *As a corollary, both estimates are valid for all separable infinite-dimensional tame* (= *postliminal, or* GCR-) *algebras and for all group* C*-*algebras of amenable infinite group with a countable base.*

In process of the proof, the concrete A-module X with ${}_A\operatorname{dh} X = 2$ is displayed. Namely, one can always extend I to a larger two-sided ideal I_1 which is again noncomplementable as a subalgebra and also such that $A_+/I \simeq M_n$, the algebra of all $n \times n$ matrices, for some n. Then $X = \mathcal{M}(I_1)$, the multiplier algebra of I_1, has the desired property.

However, it is an open question, whether the global dimension theorem is valid for all infinite-dimensional C*-algebras, even if they are separable. As to the latter, the following conditional result of Aristov (idem) is worth of notice: if we have $\operatorname{dg} A \geqslant 2$ for all simple unital separable infinite-dimensional C*-algebras, then the same is true for all separable infinite-dimensional C*-algebras.

Note that the respective parts of the cited theorems could be reformulated as follows: the number 1 is a forbidden value for dg, as well as for db, in the class of function Banach algebras and in the class of separable tame C*-algebras. Later we shall see that the same is true for group algebras of amenable groups (Theorem 2.5.13) and for $\mathcal{N}(E)$, where E is an arbitrary Banach space (Theorem 2.5.12). All these results, combined with other available information about various homological dimensions (see the remaining subsections), provoke to state the following aggregative "problem of forbidden values".

PROBLEM.
- (i) *Is it true that* $\operatorname{dg} A$ *and/or* $\operatorname{db} A$ *are never equal to* 1 *in the class of semisimple Banach algebras? Or at least* C*-*algebras? Or all commutative Banach algebras?*
- (ii) *The same, with* "$\mathrm{w.dg}\, A$ *and/or* $\mathrm{w.db}\, A$" *instead* "$\operatorname{dg} A$ *and/or* $\operatorname{db} A$".[15]
- (iii) *Is it true that* $\mathrm{n.db} A$ *is never equal to* 1 *in the class of operator* C*-*algebras? Or at least von Neumann algebras?*

Such a problem, modulo some nuances in formulation, was referred as "one of the most intriguing unanswered questions in functional analysis" by E. Effros and A. Kishimoto [45] (cf. also [163]) – and, in our humble opinion, quite deservedly. As a part of it, let us distinguish a "triple" question, which appears to be most urgent:
- Do there exist non-contractible (= infinite-dimensional) C*-algebras A with $\mathcal{H}^2(A, X) = 0$ for all $X \in$ **A-mod-A**?
- Do there exist non-Johnson-amenable (= non-nuclear) C*-algebras A with $\mathcal{H}^2(A, X) = 0$ for all *dual* $X \in$ **A-mod-A**?
- Do there exist non-Connes-amenable (= non-hyperfinite) von Neumann algebras A with $\mathcal{H}^2(A, X) = 0$ for all *normal* $X \in$ **A-mod-A**?

As to the third question, a very recent result of Christensen and Sinclair [25] gives the impression that, most probably, it will be answered in negative. They have assigned to every von Neumann algebra A an A-bimodule X and, moreover, a cocycle $f \in Z^2(A, X)$ such that f is a coboundary iff A is hyperfinite. However, this bimodule is not dual (although, in a sense, it is "rather near" to dual bimodules).

Here is a partial result on behalf of the conjecture, concerning the commutative case.

THEOREM 2.5.5 (Selivanov [197]). *Let A be a commutative Banach algebra without a bounded approximate identity. If, in addition, A has the bounded approximation property,*[16] *then* $\operatorname{dg} A \geqslant 2$ *and hence* $\operatorname{db} A \geqslant 2$.

Now we want to emphasize that, stating questions about forbidden values, we deliberately considered specific classes of Banach algebras, and not all Banach algebras. The thing is that Banach, and even finite-dimensional algebras A with $\operatorname{dg} A = \operatorname{db} A = \mathrm{w.dg}\, A =$

[15] Selivanov has shown that $\mathrm{w.db}\, \mathcal{K}(\mathcal{N}(H)) = 1$ (cf. p. 191).

[16] It is a strengthened version of the "usual" approximation property; see, e.g., [143].

w.db $A = 1$ exist; the simplest example is the algebra consisting of 2×2 matrixes of the form

$$\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}, \quad a, b \in \mathbf{C}$$

([16]; see also [189]). Also it is known that for the algebra of upper triangular $n \times n$ matrixes all the four dimensions are again equal to 1 (two different proofs are given by Smith [164] and Selivanov). At the moment, however, all Banach algebras occurring in known examples of that kind, are non-semisimple and at the same time non-commutative.

REMARK. We do not consider in our exposition the so-called strict homological theory of Banach algebras and modules, based on the injective ("$\widehat{\widehat{\otimes}}$-") instead of projective ("$\widehat{\otimes}$-") tensor product. The behavior of homological characteristics in this theory essentially differs from that in the "standard" Banach homology and more resembles to the known facts of abstract homological algebra. In particular, the problem of forbidden values does not arise. For example, both "strict" dimensions, global and bi-, of $C(\Omega)$ in the case of infinite metrizable space Ω are equal to 1 (Kurmakaeva [139]).

* * *

The homological dimensions of Banach algebras have one more peculiarity: they behave "too regularly" under the operation of projective tensor product of algebras.

A recent[17] result of Selivanov claims the following: if A_k, $1 \leqslant k \leqslant n$, are Banach algebras with b.a.i., then the "*additivity formula*"

$$\text{w.db}(A_1 \widehat{\otimes} \cdots \widehat{\otimes} A_n) = \text{w.db}\, A_1 + \cdots + \text{w.db}\, A_n \tag{2.8}$$

is always valid.

We do not know at the moment, whether the additivity formulae like (2.8) hold for the "strong" dimensions dg and db. In this setting, of course, we must assume that the algebras A_k are unital. (This condition, as well as that of the presence of b.a.i. in the additivity formulae for weak dimensions, is unavoidable: for example, $\langle$ dg and db | w.dg and w.db $\rangle$ are equal to 2 for $\langle c_0 \mid l_1 \rangle$ as well as for $\langle c_0 \widehat{\otimes} c_0 \mid l_1 \widehat{\otimes} l_1 \simeq l_1 \rangle$). Such a formula is proved, for both db [63] and dg [137] in cases where all A_k are unitizations of biprojective Banach function algebras.

In the pure algebraic context, the dimensions dg and db are equal to 1 for the algebra of all eventually constant sequences (that is for $(c_{00})_+$), and the same equality holds for the tensor powers of this algebra. It is easy to see, that the already mentioned specific features of Banach structures, like the abundance of uncomplementable subspaces, prevent the construction of similar examples in Banach homology.

[17] It was recent in 1995.

2.5.3. *Dimensions of biprojective and biflat algebras.* For several traditional classes of biprojective algebras (cf. 1.3.3), their "strong" dimensions can be exactly computed. Combining the mutually opposite estimates, given in Proposition 2.5.8(i) and in the previous subsection, we get:

THEOREM 2.5.6. *For all of the following algebras A we have* $\operatorname{dg} A = \operatorname{db} A = 2$*:*

(i) *A is a biprojective commutative Banach algebra with infinite spectrum* (say, $A = C_0(\Omega)$ with Ω discrete or $A = l_1$).
(ii) *A is a C^*-algebra with discrete spectrum and finite-dimensional irreducible representations* (for example, $A = C^*(G)$ for a compact group G).
(iii) *$A = L^1(G)$ for a compact group G.*
(iv) [195] *$A = \mathcal{N}(H)$ for a Hilbert space H.*

For some of these algebras, it is possible to characterize A-modules of the biggest possible homological dimension in terms of Banach space geometry.

PROPOSITION 2.5.9 (Krichevets [136]). *Let A be a biprojective function Banach algebra with a b.a.i., X a Banach A-module. Then ${}_A\operatorname{dh} X = 2$ iff the subspace $\overline{A \cdot X}$ in X has no Banach complement.*

(In some properly modified form,the result is valid also without the assumption of the existence of b.a.i. [196].)

Now, knowing that the ⟨ "strong" | weak ⟩ dimensions of ⟨ biprojective | biflat ⟩ Banach algebras can assume, generally speaking, only three possible values 0, 1 or 2, it is natural to ask, under which conditions every of these possibilities really happens. (In fact, the case of 0, that is of the ⟨ contractibility | amenability ⟩ was already considered in 1.3, but we include it in the present discussion for the sake of completeness.) At the moment, the picture is perfectly transparent for the weak dimensions:

THEOREM 2.5.7 (Selivanov [200]). *Let A be a biflat Banach algebra. Then*

$$\operatorname{w.db} A = \begin{cases} 0 & \text{iff } A \text{ has a b.a.i.,} \\ 1 & \text{iff } A \text{ has a left or right, but no two-sided b.a.i.,} \\ 2 & \text{iff } A \text{ has neither a left nor a right b.a.i.} \end{cases}$$

Moreover, in the latter case $\mathcal{H}^2(A, X) \neq 0$ for the concrete dual A-bimodule $X = (A^* \widehat{\otimes} A^*)^*$.

(An argument, similar to that of [200], shows that the same result is valid for w.dg instead of w.db.)

Note that the algebra $\mathcal{N}(E)$, in the case of an arbitrary infinite-dimensional E, has no one-sided b.a.i. [195]. Hence, if E has, in addition, the approximation property, then we have $\operatorname{w.dg} \mathcal{N}(E) = \operatorname{w.db} \mathcal{N}(E) = 2$, and thus the weak dimensions of this algebra are equal to the "strong" dimensions (cf. Theorem 2.5.6(iv)).

The available respective result for "strong" dimensions is not so complete:

PROPOSITION 2.5.10 (idem). *Let A be a biprojective Banach algebra. Then*

$$\operatorname{db} A = \begin{cases} 0 & \textit{iff } A \textit{ has an identity,} \\ 1 & \textit{if } A \textit{ has a left or right, but no two-sided, identity,} \\ 2 & \textit{if } A \textit{ has neither a left nor a right identity and does not have a} \\ & \textit{b.a.i.} \end{cases}$$

Moreover, in the latter case $\mathcal{H}^2(A, X) \neq 0$ for the concrete A-bimodule $X = A \widehat{\otimes} A$.

Most probably, the assumption, concerning the absence of a b.a.i., is redundant, and hence the result can be completed by replacing both "if" by "iff". But this is still not proved in the general case, and besides, it is unknown whether Proposition 2.5.10 is valid for dg instead of db.

2.5.4. *Dimensions and function algebras.* From now on we shall concentrate on the behaviour of homological dimensions in some traditional classes of "algebras of analysis". Naturally we begin with Banach function algebras, the object of the classical Gel'fand theory.

The results already presented and the first of all the global dimension theorems attract interest to the following rather old question: *what is the set of values taken by global dimension, and what is the set of values, taken by the bidimension, for the whole class of Banach function algebras?*

Combining Theorems 2.5.2 and 2.5.6(i) with the additivity formulae for dg and db mentioned at the very end of Subsection 2.5.1 (see also [189]), we get that both sets contain all even non-negative numbers, and do not contain 1. Further, both sets contain ∞: for example, $\operatorname{dg} A = \operatorname{db} A = \infty$ for the algebra $A = C_n[a, b]$, $n = 1, 2, \ldots$, of n-smooth functions on the segment (Kleshchov [133]), and, say, for $A = l_2$ [196]. Finally, Selivanov [194] proved that for any unital Banach algebra B, and the sequence algebra c (see Introduction) $\operatorname{dg} c \widehat{\otimes} B = \operatorname{dg} B + 2$, and the same is valid for db. This obviously implies that both our sets contain, with every n, also $n + 2$.

Summing up, we see that each of these sets must necessary have one of the following descriptions: (1) all even numbers and ∞, or (2) all even numbers, all odd numbers save some initial segment $1, \ldots, 2n + 1$, $n \geqslant 0$, and ∞. However, this is all we know at the moment. Note that if a Banach function algebra A with $\langle \operatorname{dg} A = 3 \mid \operatorname{db} A = 3 \rangle$ would be discovered that would mean that there is only one "forbidden" value for the $\langle$ global dimension | bidimension $\rangle$ of Banach function algebras, and this would be 1.

* * *

The problem of describing the sets of values, assumed by dg and db, is still open if we restrict the class of considered algebras and take only $C(\Omega)$ for Ω compact. The respective question, being intimately connected with questions of general topology and set theory, is of an independent interest.

Note that the function algebras of even homological dimensions, provided by the additivity formulae, are not uniform and therefore do not belong to the class $C(\Omega)$. Nevertheless, "even-dimensional" algebras, belonging to the latter class, were eventually discovered:

THEOREM 2.5.8 (Krichevets [138]). *Let K be the one-point compactification of a discrete topological space of a cardinality exceeding at least a countable set of different cardinalities, let K^n be its n-th Cartesian power, $n = 1, 2, \ldots$ Then* $\operatorname{dg} C(K^n) = \operatorname{db} C(K^n) = 2n$.

It was also proved (idem) that, for $A = C(K^n)$, the quantities ${}_A \operatorname{dh} \mathbf{C}_s$, where s runs over K^n (= the spectrum of A), assume all possible values in $\{0, 1, \ldots, n\}$. Besides, ${}_A \operatorname{dh} \mathbf{C}_s = n$ iff $s = (\omega, \ldots, \omega)$ where ω is a point at infinity in K.

Thus, for both dimensions, in the class $C(\Omega)$, again all even numbers are permitted, and 1 is forbidden. The only extra information we have is that ∞ is also a permitted value. The relevant example is due to Moran [152] who took the segment $[0, \alpha]$ of the transfinite line with the order topology and proved that $\operatorname{dg} C[0, \alpha] = \operatorname{db} C[0, \alpha] = \infty$ provided the cardinality of α is sufficiently big (e.g., the same as in the previous theorem).

REMARK. It is probable that the general relation $\operatorname{dg} C(\Omega \times K) = \operatorname{dg} C(\Omega) + 2$, and a similar one for db, hold where Ω is an arbitrary compact space, and K is that from Theorem 2.5.8. If so, we have for these dimensions of algebras $C(\Omega)$ the same two possibilities which were indicated above for general Banach function algebras.

* * *

From algebras, important in general topology, we pass to those in complex analysis. Here, as well as above, the behaviour of the homological dimensions of "point modules" $\mathbf{C}_s$, $s \in \underline{\Omega}(A)$, is rather significant. First, we indicate the situation in the case of the so-called polydomain algebras. For the sake of transparency,we restrict ourselves to the consideration of the polydisc algebra $\mathcal{A}(\overline{\mathbf{D}}^n)$ (cf. Introduction). Recall that the spectrum of this algebra is just $\overline{\mathbf{D}}^n$.

PROPOSITION 2.5.11 [80]. *Suppose $s = (\lambda_1, \ldots, \lambda_n) \in \overline{\mathbf{D}}^n$ is such that m of its coordinates belong to the open disc $\mathbf{D}$ (and hence the remaining coordinates belong to the unit circle $\mathbf{T}$). Then, for $A = \mathcal{A}(\overline{\mathbf{D}}^n)$, we have*

$$
{}_A \operatorname{dh} \mathbf{C}_s = \begin{cases} m + 1 & \text{if } m < n, \\ n & \text{if } m = n. \end{cases}
$$

Thus the dimensions of point modules reflect the natural classification of points of the topological boundary of the polydisc.

As to the proof, if $m = n$, then the augmented Koszul complex

$$
0 \leftarrow \mathbf{C}_s \leftarrow \mathcal{K}(A, S_1, \ldots, S_n),
$$

where $S_k : w \mapsto (z_k - \lambda_k) w$ (z_k is the k-th coordinate function) provides the shortest projective resolution of $\mathbf{C}_s$. If $m < n$, one displays some "aggregate" resolution of $\mathbf{C}_s$, tensoring a suitable augmenting Koszul complex of length m by a complex of length 1, representing, in a sense, "boundary" coordinates of s.

REMARK. If we consider, instead of the polydisc algebra, the similarly defined "ball algebra" $\mathcal{A}(\overline{S}^n)$, where $\overline{S}^n$ is the closed unit ball in $\mathbf{C}^n$, Koszul complexes stop to work, and the situation becomes rather obscure. We do not know whether, for such A, ${}_A\operatorname{dh}\mathbf{C}_s = n$ if s is a point of the open ball. (As to the points of the unit sphere, we have ${}_A\operatorname{dh}\mathbf{C}_s = 1$.)

More interesting is that the equality ${}_A\operatorname{dh}\mathbf{C}_s = n$ implies (under reasonable conditions), that A "locally" consists of holomorphic functions of n variables, and so the part of Proposition 2.5.11 concerning inner points of $\overline{\mathbf{D}}^n$ reflected the general situation. Recall that, for a commutative Banach algebra A with spectrum Ω, a subset Δ of Ω is called an *analytic n-disc in Δ*, if there exists a homeomorphism between Δ and the open unit n-disc $\mathbf{D}^n \subset \mathbf{C}^n$ such that for any $a \in A$ the function $a(t)$, $t \in \Omega$, becomes analytic on Δ after identifying the latter with $\mathbf{D}^n$. The following theorem can be considered as a proper "n-dimensional" generalization of the "one-dimensional" Theorem 1.1.2(ii).

THEOREM 2.5.9 (Pugach [172]). *Let A be an arbitrary commutative Banach algebra with spectrum Ω, and s be a point of Ω such that ${}_A\operatorname{dh}\mathbf{C}_s = n > 0$ or, equivalently, ${}_A\operatorname{dh} I_s = n - 1 \geqslant 0$, where I_s is the maximal ideal corresponding to s. Then*
(i) *(linear)* $\dim I/\overline{I}^2 \leqslant n$, *and actually it can be any integer among* $0, 1, \ldots, n$,
(ii) *if, in addition,* $\dim I/\overline{I}^2 = n$ ("non-degenerate case"), *then there exists a neighbourhood of s in Ω which is an analytic n-disc.*

Note that the second, evidently most substantial, part of the theorem uses results of Read [177] about the connection of the behaviour of powers of maximal ideals in Banach algebras with the analytic structure on their spectra.

The situation when we have ${}_A\operatorname{w.dh}\mathbf{C}_s = n$ instead of ${}_A\operatorname{dh}\mathbf{C}_s = n$, and again $\dim I/\overline{I}^2 = n$, was clarified quite recently. In this case (cf. Theorem 1.1.7), the point s still belongs to an analytic n disc but the latter, generally speaking, does not form its neighbourhood in the Gel'fand topology. (See the forthcoming paper of Pugach and/or White.)

* * *

Hitherto we spoke almost exclusively about Banach function algebras. It was already mentioned that one of main specific features of their homological theory, the existence of forbidden values, does not occur if we proceed to more general algebras. The remaining results of this subsection show that one can find function algebras of any prescribed dimension even among Fréchet and simultaneously Arens–Michael algebras.

One of the oldest and comparatively simple results of that kind is as follows.

PROPOSITION 2.5.12. *Suppose $A = \mathcal{O}(\mathcal{U})$ where $\mathcal{U}$ is a polydomain in $\mathbf{C}^n$. Then* $\operatorname{dg} A = \operatorname{db} A = n$.

Indeed, recall that the augmented Koszul complex (2.5) in 2.2.3 (suggested by Taylor [208]), is a projective A-bimodule resolution of A. Since it has the length n, $\operatorname{db} A \leqslant n$.

Further, using the resolution indicated in Theorem 2.2.4(i), to compute Ext spaces, we easily see that ${}_A\operatorname{Ext}^n(\mathbf{C}_\lambda, \mathbf{C}_\lambda) = \mathbf{C} \neq 0$ where $\lambda \in \mathcal{U}$, and $\mathbf{C}_\lambda$ is $\mathbf{C}$ with the outer multiplication $w \cdot z = w(\lambda)z$. This implies $\operatorname{dg} \geqslant n$.

In fact, taking any of the algebras A, involved in Theorem 2.2.4 and using that theorem, one can prove the same equality $\operatorname{dg} A = \operatorname{db} A = n$ [156]. In particular, it is true for $A = \mathrm{C}^\infty(\mathcal{U})$, $\mathcal{U} \subseteq \mathbf{R}^n$. The next result, however, concerns much wider class of algebras, which, generally speaking, have no Koszul resolution.

THEOREM 2.5.10 (Ogneva [154]). *Let $\mathcal{M}$ be any infinitely smooth manifold of topological dimension n, $A = \mathrm{C}^\infty(\mathcal{M})$. Then $\operatorname{dg} A = \operatorname{db} A = n$ and besides, ${}_A\operatorname{dh} \mathbf{C}_\lambda = n$ for any A-module $\mathbf{C}_\lambda$, $\lambda \in \mathcal{M}$. (The latter is $\mathbf{C}$ with the outer multiplication $a \cdot z = a(\lambda)z$.)*

Roughly speaking, the idea of the proof is as follows. Despite the fact that A itself is not bound to have the Koszul resolution, such resolutions exist, by virtue of Proposition 1.1.6, for A-bimodules $\mathrm{C}^\infty(\mathcal{U})$ where $\mathcal{U}$ is is a local chart in $\mathcal{M}$. This enables one to get, for these $\mathcal{U}$, the equality ${}_A\operatorname{dh}_A \mathrm{C}^\infty(\mathcal{U}) = n$, the "local" version of the desired result. Then, starting from a family of charts $\mathcal{U}$, forming an atlas of $\mathcal{M}$, and using the technics of long sequences for Ext, one can "glue" the respective "local" equalities to the "global" equality ${}_A\operatorname{dh}_A \mathrm{C}^\infty(\mathcal{M}) = n$.

REMARK. Under some topological assumptions on $\mathcal{M}$, Connes [31] explicitly displayed a projective bimodule resolution of $\mathrm{C}^\infty(\mathcal{M})$ of length n. Its bimodules consist of smooth sections of certain vector bundles over $\mathcal{M} \times \mathcal{M}$.

REMARK. It is worth mentioning that the picture drastically changes if we pass from $\mathrm{C}^\infty(\mathcal{M})$ to the larger Fréchet algebra $\mathrm{C}^n(\mathcal{M})$ of n-smooth functions on $\mathcal{M}$. (It is in fact Banach algebra if $\mathcal{M}$ is compact.) Whatever $n = 1, 2, \ldots$ would be, we always have $\operatorname{dg} \mathrm{C}^n(\mathcal{M}) = \operatorname{db} \mathrm{C}^n(\mathcal{M}) = \infty$, and, in addition, ${}_A\operatorname{dh} \mathbf{C}_\lambda = \infty$ $(A = \mathrm{C}^n(\mathcal{M}))$ for all $\lambda \in \mathcal{M}$. (This generalizes what was said about $\mathrm{C}^n[a, b]$ in 2.5.2.)

The theorem of Ogneva suggests the question whether Proposition 2.5.12 can also be generalized, this time to an arbitrary Stein manifold $\mathcal{S}$. However, the dimensions of $\mathcal{O}(\mathcal{S})$ are still waiting to be computed; the objective difficulty is that Proposition 1.1.6 has no analogue for subdomains of complex manifolds (cf. 1.1.1). At the same time, Putinar [174] has announced that (in equivalent terms) $\operatorname{w.dg} \mathcal{O}(\mathcal{S})$ coincides with the complex dimension of $\mathcal{S}$ (proof not given).

2.5.5. *Spatial dimensions, dimensions of group algebras and miscellany.* We turn from function algebras to some other popular classes of algebras. Practically all what we know about the dimensions of C^*-algebras was already discussed in 2.5.2 and 2.5.3. Now we mention several results on operator algebras outside the class of C^*-algebras.

Terminology like "spatially projective" or "spatially flat" (see 1.1), suggests:

DEFINITION 2.5.4. Let A be an operator algebra on a Banach space E. The $\langle$ (projective) $\mid$ injective $\mid$ weak $\rangle$ homological dimension of the natural Banach left A-module E

is called the spatial $\langle$ (projective) | injective | weak $\rangle$ homological dimension of A, and it is denoted by $\langle$ sp.d A | i.sp.d A | w.sp.d A $\rangle$.

Needless to say, Proposition 1.1.4 and Theorems 1.1.4 and 1.1.7 can be reformulated giving conditions of vanishing the respective spatial dimensions. Here we pass to a result concerning, generally speaking, positive spatial dimensions. Its main meaning is that the impact of taking suspensions, and also some other "good" joins, on the projective spatial dimension resembles the impact of the constructions of the same names on the spatial cohomology groups (see Corollary 2.4.4 and Theorem 2.4.17).

THEOREM 2.5.11 (Golovin [61]). *Let $A, B_1, \dots, B_m$ be unital norm closed operator algebras acting on* (generally speaking, different) *Hilbert spaces and let $B = B_1 \oplus \cdots \oplus B_m$ be the spatial direct sum of these algebras. Further, suppose that every B_j, $1 \leqslant j \leqslant m$, contains a column of rank-one operators* (see 1.1.1). *Then* (i) $\mathrm{sp.d}(A\#B) = \mathrm{sp.d}\,A + 1$ *if $m > 1$, and* (ii) $\mathrm{sp.d}(A\#B) = 0$ *if $m = 1$.*

In particular, taking the suspension (i.e., transfer from A to $A\#\mathbf{C}^2$) *always increases the spatial homological dimension by* 1.

Note that case (ii) generalizes the already mentioned fact (see 1.1.1) that the taking a cone (i.e., passing from A to $A\#\mathbf{C}$) always gives a spatially projective algebra.

EXAMPLE 2.5.4. If, for any H, $A = \mathcal{B}(H)$, then $\mathrm{sp.d}\,S^n(\mathcal{B}(H)) = n$, where S^n is the n-fold suspension. In particular, $\mathrm{sp.d}\,S^n(\mathbf{C}) = n$. Besides,

$$\mathrm{sp.d}\,\underbrace{S(S(\dots}_{\infty}\mathcal{B}(H)\dots)) = \infty$$

Now we can finish the discussion of various dimensions of the algebras $\mathcal{N}(E)$ and related modules. For the sake of completeness, the following final result includes the information already given earlier (in Theorem 2.5.6(iv)).

THEOREM 2.5.12 (Selivanov [195]). *Let E be an arbitrary Banach space. Then, for $A = \mathcal{N}(E)$, we have*

$$\begin{aligned} \mathrm{dg}\,A &= \mathrm{w.dg}\,A = \mathrm{db}\,A = \mathrm{w.db}\,A \\ &= \begin{cases} 0 & \textit{iff } E \textit{ is finite-dimensional,} \\ 2 & \textit{iff } E \textit{ is infinite-dimensional and has the approximation property,} \\ \infty & \textit{iff } E \textit{ does not have the approximation property.} \end{cases} \end{aligned}$$

It is also known, that for the algebra of Hilbert–Schmidt operators on a Hilbert space, as well as for its "commutative prototype" l_2, all dimensions, indicated in the previous theorem, are equal to ∞ (cf. [196]).

Finally, it is worth noting that any infinite-dimensional von Neumann algebra, even an AW*-algebra, contains non-projective ideals (Lykova [145]).

* * *

Passing to L^1-algebras of locally compact groups, recall that we already know that in the case of a compact infinite group we have $\operatorname{dg} L^1(G) = \operatorname{db} L^1(G) = 2$ (Theorem 2.5.6). The following theorem tells what usually happens outside of that class of groups.

THEOREM 2.5.13 (Sheinberg). *Let G be a locally compact group which contains a closed amenable non-compact subgroup. Then, for $A = L^1(G)$, we have*

$$_A\operatorname{dh} I_{\mathrm{aug}} = {_A\operatorname{dh}} \mathbf{C}_{\mathrm{aug}} = \operatorname{dg} A = \operatorname{db} A = \infty$$

(for the definition of I_{aug} and $\mathbf{C}_{\mathrm{aug}}$, see Example 0.2.4).

This theorem was proved in [202] under the assumption that G itself is amenable and non-compact. However, the argument still applies in the more general case.

To give the general idea of the proof, we shall restrict ourselves to the model case of the Wiener algebra $W = L^1(\mathbf{Z})$. One can observe that the W-module $\mathbf{C} = \mathbf{C}_{\mathrm{aug}}$ has a resolution, consisting of modules, which are simultaneously projective and injective. (Indeed, we can take the standard resolution

$$0 \leftarrow \mathbf{C} \leftarrow L^1(\mathbf{Z}) \widehat{\otimes} \mathbf{C} \leftarrow \cdots \leftarrow L^1(\mathbf{Z}) \widehat{\otimes} \cdots \widehat{\otimes} L^1(\mathbf{Z}) \widehat{\otimes} \mathbf{C} \leftarrow \cdots$$

and see that its modules, being dual to $c_0(\mathbf{Z} \times \cdots \times \mathbf{Z})$, are injective by virtue of the amenability of our algebra; cf. 1.3.2.) Therefore an easy argument, concerning kernels = images of morphisms in such a resolution, gives that the inequality $_W\operatorname{dh} \mathbf{C}_{\mathrm{aug}} \leqslant n$, for a finite $n > 0$, would imply $_W\operatorname{dh} \mathbf{C}_{\mathrm{aug}} \leqslant n - 1$ and eventually $_W\operatorname{dh} \mathbf{C}_{\mathrm{aug}} = 0$. However, the latter equality is impossible because $\mathbf{Z}$ is not compact (see Proposition 1.1.10(i)).

It is an open question, whether the assertion of the Sheinberg's theorem is valid for an arbitrary non-compact group G. It is true that the great majority of known groups satisfy the condition of the theorem. However, this condition is not ubiquitous. For example, there are discrete groups of Ol'shanskii [157] which are not amenable, but all their proper subgroups are finite. Similar exotic groups certainly require some specific treatment.

* * *

We have a rather scant knowledge about the behaviour of dimensions in the class of radical Banach algebras. For all algebras of that class with explicitly computed dimensions we have invariably $_A\operatorname{dh} \mathbf{C}_\infty = \operatorname{dg} A = \operatorname{db} A = \infty$. This is the case for all nilpotent Banach algebras [196] and for weighted convolution algebras $l_1(\omega)$ with a very rapidly decreasing wheight $\omega = \omega(n)$, e.g., for $\omega(n) = \exp(-n^{1+\varepsilon})$, $\varepsilon > 0$ (Gumerov [73]). For somewhat larger classes of radical algebras it is established that $3 \leqslant \operatorname{dg} A$, $\operatorname{db} A$: this is the case for all uniformly radical commutative Banach algebras, that is, A with $\lim_{n\to\infty} \sqrt[n]{r_n} = 0$, where $r_n = \sup\{\|r^n\|: r \in A, \|r\| \leqslant 1\}$ [196,197]. (In particular, this applies to the algebra $C_*[0, 1]$ with underlying space $C[0, 1]$, but with convolution multiplication.)

Finally, many radical algebras satisfy the conditions of Theorem 2.5.5 and hence have $2 \leqslant \operatorname{dg} A, \operatorname{db} A$; the same estimate is valid for all radical commutative Banach algebras A such that $\overline{A}^2 \nsubseteq A$ [86,73]. Both classes obviously include $l_1(\omega)$ with an arbitrary radical weight ω. As to radical algebras, possessing a bounded approximate identity, such an estimate is given for the Volterra algebra $L^1[0, 1]$ [197] and for $L^1(\mathbf{R}^+, \omega)$ with a measurable radical weight $\omega = \omega(t)$ (this is a very recent result of Ghahramani and Selivanov).[18]

* * *

Despite rather a lot of results presented in this section, we feel obliged to end it with some kind of complaint. The challenging thing is that the dimensions of quite a few well-known concrete algebras are still not exactly computed. Indeed, the "strong" dimensions (dg and db) of such time-honoured algebras as $\mathbf{C}[0, 1]$, c_b, $\mathcal{K}(H)$ are unknown, and the same is true of the strong, as well as weak, dimensions of $\mathcal{A}(\mathbf{D}^n)$, $n = 1, 2, \ldots$, $\mathcal{B}(H)$, $L^1[0, 1]$.

Where is the difficulty? To get the estimate, say, $\langle \operatorname{db} A \leqslant n \mid \operatorname{db} A > n \rangle$ we must show that in some projective resolution

$$0 \longleftarrow A_+ \text{ (or } A) \longleftarrow P_0 \longleftarrow \cdots \xleftarrow{d_{n-2}} P_{n-1} \longleftarrow \cdots \tag{2.9}$$

the A-bimodule $\operatorname{Ker} d_{n-2}$ $\langle$ is | is not $\rangle$ projective. In "lucky" cases like c_0, $L^1(\mathbf{T})$, $\mathcal{N}(H)$, $C(K^n)$, W etc. one has managed to find sufficiently convenient resolutions to work with. However, in the cases like $A = C[0, 1]$ or $\mathcal{K}(H)$ the only known projective resolution of A is the bar-resolution. Its bimodules are projective tensor powers of A_+ or A (see 2.2.1), and they can be fairly complicated things even if A itself appears to be rather simple. As an example, for $A = C[0, 1]$ the bimodule P_{n-1} in (2.1) happens to be the Varopoulos algebra (cf. Example 2.2.2) consisting of functions on the $(n + 1)$ dimensional cube $[0, 1]^{n+1}$. There is no transparent description of these functions (somewhere between 1-smooth and all continuous functions). The sub-bimodule $\operatorname{Ker} d_{n-2}$ of P_{n-1} is still less understandable. It is true that one can deduce from the theorem of global dimension in some roundabout way that $\operatorname{Ker} d_{n-2}$ is not projective if $n = 1$. However, the endeavour to solve the question of its projectivity for larger n – at least before the appearance of essentially new ideas – encounters considerable difficulties.

2.6. *Cyclic and simplicial (co)homology (topological versions)*

A new important algebraic type of (co)homology groups of associative algebras was discovered in early 80-ies by Connes [31] and Tzygan [213]. In the present subsection we discuss the principal topological version of this notion. We want to emphasize that the rich and powerful pure algebraic theory of cyclic (co)homology, including deep connections

[18] Now it is published.

with K-theory, remains outside the scope of our exposition; see [144] and, as an excellent introduction to the subject, the last chapter of [184]. Apart from this, we almost do not touch the pure algebraic cyclic (co)homology of those algebras which in fact have a natural Banach or Fréchet structure; such groups were studied by Wodzicki [217–219] who has demonstrated that these things can be indeed very interesting and instructive.

2.6.1. *Simplicial (co)homology and simplicially amenable algebras.* To prepare the stage for the topological version of the cyclic (co)homology, we begin with a special case of the ordinary (co)homology (see Definition 2.4.1); it has a considerable independent interest.

Let A be an arbitrary (so far) topological algebra. A *simplicial n-cochain*, $n = 0, 1, \dots$ is, by definition, a continuous $(n + 1)$-linear functional on A (the meaning of the word "continuous" is the same as in 2.4.1). With n fixed, such cochains form a linear space, denoted, in this context, by $C^n(A)$. In turn, these spaces form a complex

$$0 \longrightarrow C^0(A) \xrightarrow{\delta^0} \cdots \longrightarrow C^n(A) \xrightarrow{\delta^n} C^{n+1} \longrightarrow \cdots \qquad \big(\widetilde{C}(A)\big)$$

with the *"coboundary operators"* δ^n given by

$$\begin{aligned}\delta^n f(a_0, \dots, a_{n+1}) &= \sum_{k=0}^{n} (-1)^k f(a_0, \dots, a_k a_{k+1}, \dots, a_{n+1})\\ &\quad + (-1)^{n+1} f(a_{n+1}a_0, a_1, \dots, a_n).\end{aligned}$$

This complex is called the *simplicial cochain complex for A*.

DEFINITION 2.6.1. The n-th cohomology of the complex $\widetilde{C}(A)$ is called the (continuous) n-dimensional simplicial cohomology group of the topological algebra A. It is denoted by $\mathcal{H}^n(A)$.

Note that $\mathcal{H}^0(A)$ is evidently nothing else but the subspace A^{tr} of A^*, consisting of (continuous) traces (see 2.4.1).

If A is a Banach algebra, then $C^n(A)$ is a Banach space of bounded $(n + 1)$-linear functionals on A with respect to the polylinear operator norm, δ^n is bounded and $\mathcal{H}^n(A)$, $n = 0, 1, \dots$, has the structure of a complete prenormed space. Recall that bounded $(n+1)$-linear functionals on A can be identified with bounded n-linear operators from A to its dual Banach A-bimodule A^*. It follows from this that the space $C^n(A)$ is isometrically isomorphic to the "ordinary" cochain space $C^n(A, A^*)$ and, moreover, the complex $\widetilde{C}(A)$ is isomorphic to the standard cohomology complex $\widetilde{C}(A, A^*)$. Hence $\mathcal{H}^n(A)$ is isometrically isomorphic to $\mathcal{H}^n(A, A^*)$, $n = 0, 1, \dots$, and thus the simplicial cohomology of Banach algebras is but a particular case of their ordinary cohomology $\mathcal{H}^n(A, X)$, with $X = A^*$.

To define simplicial homology, we must assume that the topological algebra A has well-defined topological tensor powers, say, it is an $\widehat{\otimes}$-algebra, in particular, a Fréchet or Banach algebra (or an $\overline{\otimes}$-algebra). Then *simplicial n-chains* are, by definition, elements of the space

$$\underbrace{A \widehat{\otimes} \cdots \widehat{\otimes} A}_{n+1},$$

the latter now denoted by $C_n(A)$. (As always, here and in what follows we could replace "$\widehat{\otimes}$" by $\overline{\otimes}$.) These spaces form a complex

$$0 \longleftarrow C_0(A) \xleftarrow{d_0} \cdots \longleftarrow C_n(A) \xleftarrow{d_n} C_{n+1}(A) \longleftarrow \cdots \qquad (\underset{\sim}{C}(A))$$

with the "boundary operators" well-defined by

$$\begin{aligned} d_n(a_0 \otimes \cdots \otimes a_{n+1}) &= \sum_{k=0}^{n} (-1)^k a_0 \otimes \cdots \otimes a_k a_{k+1} \otimes \cdots \otimes a_{n+1} \\ &\quad + (-1)^{n+1} a_{n+1} a_0 \otimes a_1 \otimes \cdots \otimes a_n. \end{aligned}$$

It is so-called *simplicial chain complex for A*.

DEFINITION 2.6.2. The n-th homology of $\underset{\sim}{C}(A)$ is called the (continuous) n-dimensional simplicial homology group of the ($\widehat{\otimes}$)-algebra A. It is denoted by $\mathcal{H}_n(A)$. Note that $\mathcal{H}_n(A)$ has a structure of a polynormed space, which is a complete $\langle$ prenormed | premetric $\rangle$ space in the case when A is a $\langle$ Banach | Fréchet $\rangle$ algebra.

It is obvious that $\underset{\sim}{C}(A)$ coincides with the standard homology complex $\underset{\sim}{C}(A, A)$, and thus $\mathcal{H}_n(A)$ is just another way of writing $\mathcal{H}_n(A, A)$, the ordinary homology group of A with coefficients in itself.

REMARK. In many papers, simplicial (co)homology, being considered as a special case of Hochschild (co)homology (or of its respective topological version) is just called a Hochschild (co)homology; cf. [184,217].

Let $\kappa : A \to B$ be a continuous homomorphism of Banach algebras. Then the family of operators $\kappa^n : C^n(B) \to C^n(A)$, $n = 0, 1, \ldots$, given by

$$\kappa^n f(a_0, \ldots, a_{n+1}) = f\big(\kappa(a_0), \ldots, \kappa(a_{n+1})\big),$$

form a morphism between the complexes $\widetilde{C}(B)$ and $\widetilde{C}(A)$ in $\underline{\mathbf{Ban}}$ and hence generate, for every $n = 0, 1, \ldots$, a continuous operator $\mathcal{H}^n(\kappa)$ between the seminormed spaces $\mathcal{H}^n(B)$ and $\mathcal{H}^n(A)$. The assignment $A \mapsto \mathcal{H}^n(A)$, $\kappa \mapsto \mathcal{H}^n(\kappa)$ is evidently a contravariant functor from the category of Banach algebras to $\langle$**Ban**$\rangle$, the so-called *simplicial cohomology functor*. (In the same way, one can define a "larger" functor of simplicial cohomology, acting from the category of topological algebras to **Lin**.)

The introduced functor is denoted $\mathcal{H}^n(\,?\,)$, just as the functor from **PBAB** to $\langle$**Ban**$\rangle$, mentioned in 2.4.1. It will not lead to a confusion: one can easily show, that the former functor can be identified, up to the respective reversion of arrows, with the restriction of the latter to the category, consisting of objects of the form (A, A^*), and of morphisms of the form $(\kappa : A \to B, \kappa^* : B^* \to A^*)$.

Similarly, one can define the so-called *simplicial homology functor* $\mathcal{H}_n(\,?\,)$ from the category of $\widehat{\otimes}$-algebras to **TVS** and identify this functor (now without the reversion of arrows) with a suitable restriction of the homology functor $\mathcal{H}_n$ (see idem).

Now observe that, for any $\widehat{\otimes}$- (or $\overline{\otimes}$-) algebra A, the complex $\widetilde{C}(A)$ is linearly isomorphic to $\underset{\sim}{C}(A)^*$. Therefore Theorem 2.1.1 implies the following assertion which is, in fact, a particular case of Proposition 2.4.1.

PROPOSITION 2.6.1. *Let A be a Banach algebra, and $n > 0$ be an integer. The following conditions are equivalent:*

(i) *$\mathcal{H}_k(A) = 0$ for all $k \geqslant n$, and $\mathcal{H}_{n-1}$ is Hausdorff,*

(ii) *$\mathcal{H}^k(A) = 0$ for all $k \geqslant n$.*

Moreover, $\mathcal{H}_n(A) = 0$ for all $n \geqslant 0$ iff $\mathcal{H}^n(A) = 0$ for all $n \geqslant 0$.

We turn to algebras, which are the best from the point of view of the just introduced species of (co)homology.

DEFINITION 2.6.3. A Banach algebra A is called *simplicially amenable* if $\mathcal{H}^n(A)=0$ for all $n > 0$ (or, equivalently, $\mathcal{H}_n(A) = 0$ for all $n > 0$, and $\mathcal{H}_0(A)$ is Hausdorff).

Combining Theorems 2.4.19 (for $X = A^*$) and 2.4.20, we immediately have

PROPOSITION 2.6.2. *Every biflat (e.g., Johnson amenable or biprojective) Banach algebra is simplicially amenable.*

(Definition 2.6.3 can be extended to arbitrary Fréchet algebras, and the analogue of the Proposition 2.6.2 is still valid.)

We see that every nuclear C*-algebra, being Johnson amenable (Theorem 1.3.8) is certainly simplicially amenable. Another class of examples is provided be Christensen and Sinclair [24]: every C*-algebra A with $A^{\mathrm{tr}} = 0$ (see above) is simplicially amenable. In particular, for a Hilbert space H we have $\mathcal{H}^n(\mathcal{B}(H)) = \mathcal{H}_n(\mathcal{B}(H)) = 0$ (see also [218]). It is an open problem, whether every C*-algebra is simplicially amenable.

Note that the weakly amenable Banach algebras of 1.3.3 (which now could be defined as those with the vanishing $\mathcal{H}^1(\,\cdot\,)$) are not bound to be simplicially amenable. As an example, $\mathcal{H}^1(l_2) \neq 0$ whereas $\mathcal{H}^2(l_2) \neq 0$ (Aristov [3]).

2.6.2. *Cyclic (co)homology, Connes–Tzygan exact sequence and the excision theorem.* Let us return, under the assumptions of the previous subsection, to simplicial (co)chain complexes. Consider the operators $t_{\overline{n}} : C^n(A) \to C^n(A)$, defined by $t_{\overline{n}} f(a_0, \ldots, a_n) = (-1)^n f(a_1, \ldots, a_n, a_0)$, $n \geqslant 1$, and put $t_{\overline{0}} = \mathbf{1} : C^0(A) \to C^0(A)$. We shall write t instead of $t_{\overline{n}}$ if there is no danger of misunderstanding.

A simplicial cochain f is called *cyclic* if $tf = f$. We let $CC^n(A)$ denote the subspace of $C^n(A)$ formed by the cyclic cochains. (In particular, $CC^0(A) = C^0(A) = A^*$.) Note that for a Banach algebra A, $CC^n(A)$ is obviously closed in $C^n(A)$.

It is easy to see that the spaces $CC^n(A)$ form a subcomplex in $\widetilde{C}(A)$ denoted by $\widetilde{C}C(A)$; it is called the *cyclic cochain complex (for* A).

On the other hand, consider the operators $t_{\underline{n}} : C_n(A) \to C_n(A)$, well-defined by

$$t_{\underline{n}}(a_0 \otimes \cdots \otimes a_{n+1}) = (-1)^n a_{n+1} \otimes a_0 \otimes \cdots \otimes a_n;$$

and put $t_{\underline{0}} = \mathbf{1} : C_0(A) \to C_0(A)$. Let $CC_n(A)$ denote the quotient space of $C_n(A)$ modulo the closure of the linear span of elements of the form $x - t_{\underline{n}}x$, $n = 0, 1, \ldots$ (Thus $CC_0(A) = C_0(A)$.) It is easy to see that differentials in $\underset{\sim}{C}(A)$ induce, for all n, operators between CC_{n+1} and $CC_n(A)$. We obtain a quotient complex $\underset{\sim}{C}\,C(A)$ of $\underset{\sim}{C}(A)$, the so-called *cyclic chain complex* (for A).

DEFINITION 2.6.4 (cf. [31,213]). The n-th $\langle$ cohomology | homology $\rangle$ of the complex $\langle \widetilde{C}C(A) \mid \underset{\sim}{C}\,C(A) \rangle$ is called the *n-dimensional cyclic $\langle$ cohomology | homology $\rangle$ group of* A, and is denoted by $\langle \mathcal{H}C^n(A) \mid \mathcal{H}C_n(A) \rangle$.

Observe that the complex, dual to $\underset{\sim}{C}\,C(A)$, is isomorphic to $\widetilde{C}C(A)$ and hence, by virtue of Theorem 2.1.1, the vanishing of $\mathcal{H}C^n(A)$ for all $n \geqslant 0$ is equivalent to that of $\mathcal{H}C_n(A)$ for all $n \geqslant 0$.

If $\kappa : A \to B$ is as in 2.6.1, then $\kappa^n : C^n(B) \to C^n(A)$, $n = 0, 1, \ldots$ (see idem) obviously maps $CC^n(B)$ into $CC^n(A)$. Therefore a family of operators arises, which forms a morphism between the complexes $\widetilde{C}C(B)$ and $\widetilde{C}C(A)$, and hence generates, for any n, an operator between $\mathcal{H}C^n(B)$ and $\mathcal{H}C^n(A)$. Thus, as in the case of the simplicial cohomology, we obtain a functor from the category of Banach algebras to $\langle$**Ban**$\rangle$ (or from that of topological algebras to **Lin**). It is called the *cyclic cohomology functor* and denoted by $\mathcal{H}C^n(?)$. In the similar way, for any $n = 0, 1, \ldots$ one can define the *cyclic homology functor* $\mathcal{H}C_n(?)$.

REMARK. We leave outside of our account a later and more sophisticated version of cyclic cohomology, entire cyclic cohomology of Banach algebras. See Connes [32], and also Kastler [131].

* * *

How to compute cyclic (co)homology? An objective difficulty is the absence, at least at the moment, of workable expressions of these spaces in terms of derived functors – contrary to the case of the simplicial $\langle$ cohomology | homology $\rangle$, presented as $\langle {}_A\mathrm{Ext}^n_A(A, A^*) \mid {}^A\mathrm{Tor}^A_n(A, A) \rangle$ by virtue of Theorem 2.4.19. It is true, and in the next subsection we shall discuss that, that cyclic, as well as simplicial, (co)homology can be expressed in terms of Banach derived functors in some more sophisticated way; but so far this interpretation does not help in computing the relevant groups for concrete topological algebras.

Nevertheless, Connes [31] and Tzygan [213] discovered, in the context of (abstract) unital algebras, a powerful tool of another kind. This is a special exact sequence, connecting cyclic (co)homology with simplicial (co)homology. With its help, knowledge of the latter (co)homology (obtained, say, by means of the abovementioned Ext or Tor interpretations) gives valuable, and often exhaustive, information about the former (co)homology.

We proceed to describe the Banach version of this "Connes–Tzygan exact sequence". As we shall see, such a sequence exists for all unital and for many, but not all, non-unital Banach algebras. The relevant class of algebras was described in [94], but the same class already was defined earlier [217], in connection with another problem (see Theorem 2.6.3 below). To introduce it, we need some preparations.

Let A be a fixed Banach algebra. Consider the so-called *reduced simplicial cochain complex*

$$0 \longrightarrow C^0(A) \xrightarrow{\delta r^0} \cdots \longrightarrow C^n(A) \xrightarrow{\delta r^n} \cdots \qquad \big(\widetilde{C}R(A)\big)$$

with the same spaces as in $\widetilde{C}(A)$, but with the operators defined by

$$\delta r^n f(a_0, \ldots, a_{n+1}) = \sum_{k=0}^{n} (-1)^k f(a_0, \ldots, a_k a_{k+1}, \ldots, a_{n+1})$$

(that is, differing from δ^n by the absence of the last summand). Similarly, taking $\underset{\sim}{C}(A)$ and replacing the boundary operator d_n by dr_n, where the latter is well-defined by

$$dr_n(a_0 \otimes \cdots \otimes a_{n+1}) = \sum_{k=0}^{n} (-1)^k a_0 \otimes \cdots \otimes a_k a_{k+1} \otimes \ldots \otimes a_{n+1},$$

we obtain the so-called *reduced symplicial chain complex*, denoted by $\underset{\sim}{C}\,R(A)$. Note that dr_0 is just the product map for A, and δr^0 is its adjoint operator.

It is obvious that the complex, dual to $\underset{\sim}{C}\,R(A)$, is isomorphic to $\widetilde{C}R(A)$. This, together with the employment of the bar-resolution for the computing of the respective derived functors, easily gives

PROPOSITION 2.6.3. *The complex $\widetilde{C}R(A)$ is exact iff $\underset{\sim}{C}\,R(A)$ is exact, and both properties are equivalent to each of the properties*

$$\big({}_A \operatorname{Ext}^n(\mathbf{C}, \mathbf{C}) \mid \mathcal{H}^n(A, \mathbf{C}) \mid \operatorname{Tor}_n^A(\mathbf{C}, \mathbf{C}) \mid \mathcal{H}_n(A, \mathbf{C})\big) = 0$$

for any $n \geqslant 1$. (Here and below we write $\mathbf{C}$ instead of $\mathbf{C}_\infty$, i.e. equip it with the zero outer multiplication/s.)

DEFINITION 2.6.5 (cf. Wodzicki [217]). A Banach algebra A is called *homologically unital*, or, briefly, *H-unital*, if it satisfies the mentioned equivalent properties.

Of course, we could prolong the list of equivalent definitions of H-unitality, taking into account the relations

$$\begin{aligned}\mathcal{H}^{n+1}(A,\mathbf{C}) &= {}_A\mathrm{Ext}^{n+1}(\mathbf{C},\mathbf{C}) = \mathrm{Ext}_A^{n+1}(\mathbf{C},\mathbf{C}) = {}_A\mathrm{Ext}^n(\mathbf{C},A^*)\\ &= \mathrm{Ext}_A^n(\mathbf{C},A^*) = {}_A\mathrm{Ext}^n(A,\mathbf{C}) = \mathrm{Ext}_A^n(A,\mathbf{C});\end{aligned}$$

$n = 0, 1, 2, \ldots,$ and similar relations for the homology and Tor.

Theorems 2.3.6 and 2.3.10, and the indicated relations easily imply the following sufficient condition.

PROPOSITION 2.6.4. *A Banach algebra A is H-unital provided it is a flat left or right Banach A-module, and $A = \overline{A^2}$.*

In virtue of Proposition 1.1.9(ii), this happens rather frequently:

COROLLARY 2.6.1. *Suppose that A has a left or right b.a.i., or A is biflat. Then it is H-unital.*

Thus the class of H-unital Banach algebras includes all C*-algebras, L^1-algebras of locally compact groups, and at the same time tensor algebras, generated by dualities, and hence algebras of nuclear operators in Banach spaces with the approximation property (cf. Proposition 1.3.8 and its corollary). On the other hand, l_2 with coordinatewise multiplication and its "quantum version", the algebra $\mathcal{L}_2(H)$ of Hilbert–Schmidt operators on a Hilbert space H, are not H-unital: we have $H^1(\widetilde{C}R(l_1)) \simeq l_\infty/l_1$ and $H^1(\widetilde{C}R(\mathcal{L}_2(H)) \simeq \mathcal{B}(H)/\mathcal{N}(H)$ (see, e.g., [94]).

We come to the central point of our discussion.

THEOREM 2.6.1 (cf. [94]). *Let A be an H-unital Banach algebra. Then there exist, in the category* $\langle$**Ban**$\rangle$, *exact sequences*

$$\cdots \to \mathcal{HC}^n(A) \to \mathcal{H}^n(A) \to \mathcal{HC}^{n-1}(A) \to \mathcal{HC}^{n+1}(A) \to \mathcal{H}^{n+1}(A) \to \cdots \tag{2.10}$$

and

$$\cdots \leftarrow \mathcal{HC}_n(A) \leftarrow \mathcal{H}_n(A) \leftarrow \mathcal{HC}_{n-1}(A) \leftarrow \mathcal{HC}_{n+1}(A) \leftarrow \mathcal{H}_{n+1}(A) \leftarrow \cdots \tag{2.11}$$

These sequences are called the *Connes–Tzygan exact sequences* (for the cohomology or, respectively, homology of A).

REMARK. As a matter of fact, in the pioneering papers [31,213], and later in [94], these sequences were constructed in a certain canonical way, with uniquely determined operators. So, to be precise, those constructed sequences are actually named after Connes and

Tzygan. For the existence of such "canonical" sequences the condition of H-unitality turns out not only sufficient, but also necessary [94].

Knowing when Connes–Tzygan exact sequences exist, we now show how they work.

THEOREM 2.6.2 (cf. idem). *Let A be an H-unital simplicially amenable Banach algebra* (in particular, a biflat Banach algebra). *Then, up to topological isomorphism,*

$$\mathcal{H}C^n(A) = A^{\mathrm{tr}} \quad \text{and} \quad \mathcal{H}C_n(A) = A/\overline{[A,A]} \quad \text{for any even } n \geqslant 0,$$

and

$$\mathcal{H}C^n(A) = \mathcal{H}C_n(A) = 0 \quad \text{for any odd } n \geqslant 0.$$

(Here A^{tr} is the space of bounded traces on A (cf. 2.4.1), and $\overline{[A,A]}$ is the closure of the linear span of $\{ab - ba;\ a, b \in A\}$.)

Indeed, the vanishing of $\mathcal{H}^n(A)$; $n > 0$ in (2.10) implies the "Bott periodicity" $\mathcal{H}C^n(A) \simeq \mathcal{H}C^{n-2}(A)$. Thus, going to the left in (2.10), we see that the cyclic cohomology of A coincides with $\mathcal{H}C^0(A)$, that is with A^{tr}, in even dimensions and vanishes in odd dimensions. The assertion concerning homology can be proved by a similar argument.

EXAMPLE 2.6.1. Apart from biflat algebras, the conditions of this theorem are satisfied for C*-algebras "without bounded traces", that is, with $A^{\mathrm{tr}} = 0$ (cf. 2.5.1). Therefore for such algebras, in particular, for $\mathcal{K}(H)$ and $\mathcal{B}(H)$, we have $\mathcal{H}C^n(A) = \mathcal{H}C_n(A) = 0$ for all $n = 0, 1, 2, \ldots$ (cf. [24,218]).

EXAMPLE 2.6.2. For the biflat (moreover, biprojective) algebra $A = \mathcal{N}(E)$, where E has the approximation property, we have $A^{\mathrm{tr}} = \mathbf{C}$. Therefore

$$\mathcal{H}C^n(A) = \mathcal{H}C_n(A) = \begin{cases} \mathbf{C} & \text{for even } n, \\ 0 & \text{for odd } n. \end{cases}$$

There is another important tool, facilitating the computing of the cyclic homology for certain Banach algebras. It is the "excision theorem" of Wodzicki, which resembles the well known excision theorems in algebraic topology and algebraic K-theory.

THEOREM 2.6.3 [217]. *Let*

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

be an extension of Banach algebras (cf. 2.4.2), *which splits as a complex in* **Ban**. *If in addition, A is H-unital, then there is an exact sequence* (*in* **Lin**)

$$\begin{aligned} 0 \longleftarrow \mathcal{H}C_0(C) \longleftarrow \cdots \longleftarrow \mathcal{H}C_{n-1}(A) \longleftarrow \mathcal{H}C_n(C) \longleftarrow \\ \longleftarrow \mathcal{H}C_n(B) \longleftarrow \mathcal{H}C_n(A) \longleftarrow \mathcal{H}C_{n+1}(C) \longleftarrow \cdots \end{aligned}$$

(A similar exact sequence exists also in simplicial homology (cf. idem)).

REMARK. At the moment, let us make an exception from our general rule not to consider pure algebraic homology invariants. It is worth mentioning that Theorem 2.6.3 has a pure algebraic prototype (for k-algebras over an arbitrary field k) which was also established in [217]. This result was used in [218] for an elegant proof of a statement concerning stable C*-algebras, that is, algebras A, isomorphic to the (uniquely determined) C*-tensor product $A \otimes \mathcal{K}(H)$. Namely, if A is such an algebra, then its pure algebraic, as well as continuous cyclic homology vanishes in all dimensions (we emphasize that Example 2.6.1 above concerns only continuous (co)homology). In dimension zero that means exactly that every element of a stable C*-algebra is a finite sum of commutators (a fact, earlier proved by direct arguments in [49]). The argument of Wodzicki provides an instructive example of an interplay between cyclic and simplicial homology, and also between different dimensions of both of them. This interplay is achieved by the employment of several devices, notably the excision theorem and the functorial properties of the relevant homology constructions. Eventually the author obtains the isomorphisms $\mathcal{H}_n(A) \simeq \mathcal{H}C_n(A)$, and the Connes–Tzygan sequence delivers the final stroke.

A strong excision theorem for the so-called periodic version of the cyclic (co)homology for topological algebras was recently established by Cuntz [33].[19]

2.6.3. *Banach cyclic (co)homology in terms of some Ext/Tor.* Connes, in his study of the cyclic cohomology, discovered that these groups can be treated as particular cases of Ext, defined in some special category of so-called cyclic linear spaces [30]. The cyclic cohomology of an associative algebra A coincides with some $\operatorname{Ext}(\cdot, \cdot)$, where one of arguments is an object, associated with A, and the other argument is fixed. Such an approach leads to a natural general framework, which enables one to consider different species of cohomology – simplicial, cyclic, dihedral – from the unified point of view. Here we present the Banach version of these constructions [95].

Let $\mathcal{K}$ be a category with the countable set of objects $\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots\}$. Denote by $\underline{\mathcal{K}}$ the l_1-space constructed on the set of all morphisms of $\mathcal{K}$. It is evidently a Banach algebra under a multiplication uniquely defined as follows: the product of two morphisms is their composite, if this exists; otherwise, it is zero. With this a multiplication, $\underline{\mathcal{K}}$ is called the *Banach algebra associated to the category* $\mathcal{K}$.

With $\mathcal{K}$ fixed, let $\mathcal{F} : \mathcal{K} \to \mathbf{Ban}_1$ be a covariant functor with the values in the category $\mathbf{Ban}_1$ of Banach spaces and *contracting* operators; in the present context such a functor will be called a *$\mathcal{K}$-categorical Banach space*, or briefly a *$\mathcal{K}$-space*. Put $E^n(\mathcal{F}) = \mathcal{F}(\mathbf{n})$ and denote by $E(\mathcal{F})$ the l_1-sum of the Banach spaces $E^n(\mathcal{F})$, $n = 0, 1, \ldots$. This space has the structure of a left Banach $\underline{\mathcal{K}}$-module, said to be *associated with* $\mathcal{F}$, which is well-defined by putting, for $\xi : \mathbf{m} \to \mathbf{l}$ and $x \in E^n(\mathcal{F})$, $\xi \cdot x = [\mathcal{F}(\xi)](x)$ if $m = n$ and $\xi \cdot x = 0$ if $m \neq n$ $(l, m, n = 0, 1, \ldots)$. Moreover, if $\alpha : \mathcal{F} \to \mathcal{G}$ is a morphism (= natural transformation) of $\mathcal{K}$-spaces (as functors), then its components $\mathcal{F}(\mathbf{n}) \to \mathcal{G}(\mathbf{n})$ generate, in an obvious fashion, a contractive operator $\underline{\alpha} : E(\mathcal{F}) \to E(\mathcal{G})$, which is evidently a morphism of left $\underline{\mathcal{K}}$-modules. Further, the assignment $\mathcal{F} \mapsto E(\mathcal{F})$, $\alpha \mapsto \underline{\alpha}$ is an (injective) covariant functor from the category $[\mathcal{K}, \mathbf{Ban}_1]$ of $\mathcal{K}$-spaces and their morphisms to $\underline{\mathcal{K}}$-**mod**, which identifies the former

[19] See also later work of Brodski and Lykova.

category with a subcategory of the latter. (For an explicit description of that subcategory in terms of "geometrically essential modules", see [95].)

Similarly, every contravariant functor from $\mathcal{K}$ to $\mathbf{Ban}_1$, called now $\mathcal{K}^{\mathrm{op}}$-categorical Banach space, gives rise to a Banach right $\underline{\mathcal{K}}$-module, and the category $[\mathcal{K}^{\mathrm{op}}, \mathbf{Ban}_1]$ can be identified with a subcategory of $\mathbf{mod}$-$\underline{\mathcal{K}}$.

EXAMPLE 2.6.3. Consider the $\mathcal{K}$-space $\natural$, that assigns $\mathbf{C}$ to each $\mathbf{n}$ and $\mathbf{1}_{\mathbf{C}}$ to each morphism in $\mathcal{K}$. Then, $\underline{\natural}$, that is, the associated $\underline{\mathcal{K}}$-module, has the underlying space l_1 and the outer multiplication, defined for $\xi : \mathbf{m} \to \mathbf{n}$ and $x = (x_1, \dots, x_m, \dots)$ by $\xi \cdot x = (0, \dots, 0, x_m, 0, \dots)$, where now x_m is the n-th term of the latter sequence.

DEFINITION 2.6.6 (idem). Let $\mathcal{F}$ be a $\langle \mathcal{K} \mid \mathcal{K}^{\mathrm{op}} \rangle$-space, and $E(\mathcal{F})$ an associated Banach $\langle$ left | right $\rangle$ $\underline{\mathcal{K}}$-module. The complete seminormed space

$$\left\langle {}_{\underline{\mathcal{K}}}\mathrm{Ext}^n\left(\underline{\natural}, E(\mathcal{F})\right) \middle| \mathrm{Tor}_n^{\underline{\mathcal{K}}}\left(E(\mathcal{F}), \underline{\natural}\right)\right\rangle, \quad n = 0, 1, \dots,$$

is called the *n-dimensional* $\langle$ *$\mathcal{K}$-cohomology* | *$\mathcal{K}$-homology* $\rangle$ *group of* $\mathcal{F}$.

Now we turn to results which justify this terminology.

Recall the *standard simplicial category* Δ, perhaps, the best known of all categories with the countable set of objects [52]. If we take its formal definition, then its morphisms are given by the formal generators $\partial_n^i : \mathbf{n{-}1} \to \mathbf{n}$ ("face morphisms") and $\sigma_n^j : \mathbf{n{+}1} \to \mathbf{n}$) "degeneracy morphisms"), $i, j = 0, 1, \dots, n$, satisfying the relations

$$\begin{aligned}
\partial_{n+1}^j \partial_n^i &= \partial_{n+1}^i \partial_n^{j-1} \quad \text{for } i < j,\\
\sigma_n^j \sigma_{n+1}^i &= \sigma_n^i \sigma_{n+1}^{j-1} \quad \text{for } i \leqslant j,\\
\sigma_{n-1}^j \partial_n^i &= \begin{cases} \partial_{n-1}^i \sigma_{n-2}^{j-1} & \text{for } i < j,\\ \mathbf{1}_{\mathbf{n}-1} & \text{for } i = j \text{ and } i = j+1,\\ \partial_{n-1}^{i-1} \sigma_{n-2}^j & \text{for } i > j+1. \end{cases}
\end{aligned} \tag{2.12}$$

Restricting ourselves to the generators ∂_n^i and only taking, as relations, the first set of those in (2.1), we obtain a subcategory in Δ. The latter is called the *standard presimplicial category* and denoted by ∇.

The next category we need is the *standard cyclic category* Λ, introduced by Connes [30]. It has the same countable set of objects, but more morphisms: a new set of formal generators $\tau_n : \mathbf{n} \to \mathbf{n}$ is added to those in Δ. The relations consists of those in (2.12) and a new family

$$\begin{aligned}
\tau_n \partial_n^i &= \partial_n^{i-1} \tau_{n-1} \quad \text{for } 1 \leqslant i \leqslant n,\\
\tau_n \sigma_n^j &= \sigma_n^{j-1} \tau_{n+1} \quad \text{for } 1 \leqslant j \leqslant n,\\
(\tau_n)^{n+1} &= \mathbf{1}_n.
\end{aligned} \tag{2.13}$$

The subcategory of Λ that has the same objects but only ∂_n^i and τ_n as generating morphisms, is called the *standard precyclic category* and denoted by V (which is now Λ (lambda) formed upside down.).

In accordance with the already existing terminology, a $\mathcal{K}$-space in the case of $\mathcal{K} = \langle \Delta \mid \nabla \mid \Lambda \mid V \rangle$ is called a $\langle$ *cosimplicial* | *precosimplicial* | *cocyclic* | *precocyclic* $\rangle$ *Banach space;* and a $\mathcal{K}^{\mathrm{op}}$-space in the same case is called a $\langle$ *simplicial* | *presimplicial* | *cyclic* | *precyclic* $\rangle$ *Banach space.*

Given a cosimplicial Banach space $\mathcal{F}$, consider the sequence

$$0 \longrightarrow E^0 \xrightarrow{b^0} \cdots \longrightarrow E^{n-1} \xrightarrow{b^{n-1}} E^n \longrightarrow \cdots \qquad (\widetilde{C}(\mathcal{F})),$$

where $E^n = \mathcal{F}(\mathbf{n})$, and

$$b^{n-1} = \sum_{k=0}^{n} (-1)^k \mathcal{F}(\partial_n^i).$$

It follows from the first family of relations in (2.12) that this is a (cochain) complex in **Ban**. If $\mathcal{F}$ is in fact (not only a cosimplicial, but) cocyclic Banach space, $\widetilde{C}(\mathcal{F})$ has a subcomplex

$$0 \longrightarrow EC^0 \xrightarrow{bc^0} \cdots \longrightarrow EC^{n-1} \xrightarrow{bc^{n-1}} EC^n \longrightarrow \cdots \qquad (\widetilde{C}C(\mathcal{F}))$$

in which $EC^n = \{x \in E^n\colon\ x = (-1)^n \mathcal{F}(\tau_n)x\}$.

Since in these construction the degeneracy morphisms take no part, $\langle \widetilde{C}(\mathcal{F}) \mid \widetilde{C}C(\mathcal{F}) \rangle$ is meaningful also for the "merely" $\langle$ precosimplicial | precocyclic $\rangle$ Banach spaces.

Similarly, if a $\langle$ simplicial | cyclic $\rangle$, or "merely" $\langle$ presimplicial | precyclic $\rangle$ Banach space $\mathcal{F}$ is given, one can define by an obvious analogy the chain complex $\langle \underset{\sim}{C}(\mathcal{F}) \mid \underset{\sim}{C}C(\mathcal{F}) \rangle$.

Now, "forgetting" for a while the general Definition 2.6.6, we accept

DEFINITION 2.6.7. Let $\mathcal{F}$ be a $\langle$ cosimplicial (or precosimplicial) | cocyclic (or precocyclic) $\rangle$ Banach space; then the n-th cohomology space of the complex $\langle \widetilde{C}(\mathcal{F}) \mid \widetilde{C}C(\mathcal{F}) \rangle$ is called the *n-dimensional* $\langle$ *simplicial* | *cyclic* $\rangle$ cohomology group of $\mathcal{F}$ and denoted by $\langle \mathcal{H}^n(\mathcal{F}) \mid \mathcal{H}C^n(\mathcal{F}) \rangle$.

Similarly, the *n-dimensional homology group* of $\langle$ simplicial (presimplicial) | cyclic (precyclic) $\rangle$ Banach space $\mathcal{F}$ is defined as the n-th homology of $\langle \underset{\sim}{C}(\mathcal{F}) \mid \underset{\sim}{C}C(\mathcal{F}) \rangle$; it is denoted by $\langle \mathcal{H}_n(\mathcal{F}) \mid \mathcal{H}C_n(\mathcal{F}) \rangle$.

We single out the eminently important

EXAMPLE 2.6.4. Let A be a Banach algebra, for which we suppose the multiplicative inequality $\|ab\| \leqslant \|a\|\|b\|$ is satisfied. Let $\mathcal{F}\colon V \to \mathbf{Ban}_1$ be a covariant functor, well-defined by $\mathbf{n} \mapsto C^n(A)$, $\partial_n^i \mapsto d_n^i\colon C^{n-1}(A) \to C^n(A)$, and $\tau_n \mapsto t_n\colon C^n(A) \to C^n(A)$, where $d_n^i f(a_0, \ldots, a_n) = f(a_0, \ldots, a_i a_{i+1}, \ldots, a_n)$ for $i < n$, $d_n^n f(a_0, \ldots, a_n) =$

$f(a_n a_0, a_1, \ldots, a_{n-1})$ and $t_n f(a_0, \ldots, a_n) = f(a_1, \ldots, a_n, a_0)$. If, in addition, A has an identity e, we extend $\mathcal{F}$ to a functor (again denoted by $\mathcal{F}$) from Λ to $\mathbf{Ban}_1$, joining the assignments $\sigma_n^j \mapsto s_n^j : C^{n+1}(A) \to C^n(A)$, where $s_n^j f(a_0, \ldots, a_n) = f(a_0, \ldots, a_j, e, a_{j+1}, \ldots, a_n)$. It is obvious that $\mathcal{F}$ is precosimplicial or, in the unital case, cosimplicial Banach space, and the restriction $\mathcal{F}$ to ∇ (or, respectively, to Δ) is a precosimplicial (or cosimplicial) Banach space. One can easily check that the space $\langle \mathcal{H}^n(\mathcal{F}) \mid \mathcal{HC}^n(\mathcal{F}) \rangle$ coincides with the "usual" n-dimensional $\langle$ simplicial | cyclic $\rangle$ cohomology group $\langle \mathcal{H}^n(A) \mid \mathcal{HC}^n(A) \rangle$.

Similarly, the same Banach algebra A gives rise to a precyclic, and in the unital case a cyclic Banach space $\mathcal{F}$ with $\mathcal{F}(\mathbf{n}) = A \mathbin{\widehat{\otimes}} \cdots \mathbin{\widehat{\otimes}} A$ ($n+2$ factors) and properly chosen operators, assigned to the generator morphisms; in particular, $\mathcal{F}(\tau_j)$ takes $a_0 \otimes \cdots \otimes a_{n+1}$ into $a_{n+1} \otimes a_0 \otimes \cdots \otimes a_n$. The $\langle$ simplicial | cyclic $\rangle$ homology of this cyclic Banach space coincides with $\langle \mathcal{H}_n(A) \mid \mathcal{HC}_n(A) \rangle$.

We now arrive at the main result which shows that cyclic, as well as simplicial (co)homology in the sense of Definition 2.6.7 is but a particular case of $\mathcal{K}$-categorical (co)homology in the sense of Definition 2.6.6.

THEOREM 2.6.4 [95]. *Let $\mathcal{F}$ be a $\langle$ cosimplicial (or precosimplicial) | cocyclic (or precocyclic) $\rangle$ Banach space; then, up to a topological isomorphism,*

$$\mathcal{H}^n(\mathcal{F}) = {}_{\underline{\mathcal{K}}}\mathrm{Ext}^n\big(\underline{\natural}, E(\mathcal{F})\big), \quad n = 0, 1, \ldots,$$

where $\mathcal{K} = \langle \Delta \text{ (or } \nabla) \mid \Lambda \text{ (or } V) \rangle$.

Similarly, if in the assumption of the theorem we have contravariant, instead of covariant, functors (that is, "simplicial" instead of "cosimplicial" etc.), then, for the same respective $\mathcal{K}$, we have the "twin" formula

$$\mathcal{H}_n(\mathcal{F}) = \mathrm{Tor}_n^{\underline{\mathcal{K}}}\big(E(\mathcal{F}), \underline{\natural}\big).$$

The proof is based on the computing of the respective Ext (or Tor) spaces with the help of a special projective resolution of the $\mathcal{K}$-module $\underline{\natural}$. If $\mathcal{K} = \langle \Delta \mid \nabla \rangle$, then the $\underline{\mathcal{K}}$-modules, participating in this resolution, are just the closed left ideals $\underline{\mathcal{K}}_n$ in $\underline{\mathcal{K}}$, generated by $\mathbf{1}_n : \mathbf{n} \to \mathbf{n}$ as an (idempotent) element of $\underline{\mathcal{K}}$. If $\mathcal{K} = \Lambda$, the respective modules have the same underlying spaces as $\underline{\Delta}_n$ and the $\underline{\Lambda}$-module structure, extending in some suitable way the $\underline{\Delta}$-module structure of $\underline{\Delta}_n$. (The same holds for $\mathcal{K} = V$, if we replace Δ by ∇.) See [95] for the formulae for the differentials and other details.

The Connes–Tzygan exact sequence has its proper generalization to the case of arbitrary (co)cyclic Banach spaces:

THEOREM 2.6.5 (idem). *Let $\mathcal{F}$ be a cocyclic Banach space. Then there exists the exact sequence*

$$\cdots \to \mathcal{HC}^n(\mathcal{F}) \to \mathcal{H}^n(\mathcal{F}) \to \mathcal{HC}^{n-1}(\mathcal{F}) \to \mathcal{HC}^{n+1}(\mathcal{F}) \to \mathcal{H}^{n+1}(\mathcal{F}) \to \cdots$$

Moreover, if $\mathcal{F}$ is a precocyclic Banach space ("only"), then such a sequence exists provided ${}_{\underline{V}}\operatorname{Ext}^n(\underline{\mathbf{C}}, E(\mathcal{F})) = 0$ *where* $\underline{\mathbf{C}}$ *is the left* $\underline{V}$*-module associated with the functor, well-defined by the assignments* $\mathbf{0} \mapsto \mathbf{C}$ *and* $\mathbf{n} \mapsto \{0\}$ *for* $n > 0$.

A parallel statement is valid for cyclic Banach spaces, homology and Tor.

In fact, the indicated Ext (or the respective Tor) space vanishes in the case when a given $\mathcal{F}: V \to \mathbf{Ban}_1$ can be extended to the larger category Λ; so the first assertion of the theorem can be obtain as a corollary of the second. Thus the mentioned condition on Ext (Tor) plays the same role as the condition of H-unitality does in the "usual" cyclic (co)homology. Moreover, as in the latter case, such a condition turns out to by not only sufficient but also necessary if we shall consider a certain canonical way of constructing the relevant exact sequences.

Recently, some problems in topology and algebra led to the introduction, in the pure algebraic context, of the so-called dihedral (co)homology, based on the notion of the standard dihedral category Ξ (cf. [135]); the latter has the same objects as Λ but a larger set of morphisms. The Banach version of the dihedral cohomologies was studied by Mel'nikov [151]. He established that Banach dihedral cohomology, defined by means of some standard complex (cf. $\widetilde{C}(\mathcal{F})$ above), happens to be the particular case of a $\mathcal{K}$-categorical cohomology for $\mathcal{K} = \Xi$. (Thus he extended Theorem 2.6.4 from the cases of Δ and Λ to that of Ξ.) Besides, he has shown that the cyclic cohomology of a dihedral Banach space $\mathcal{F}$ decomposes into the direct sum of the dihedral cohomology of $\mathcal{F}$ and that of some "twin" dihedral Banach space $\mathcal{F}'$.

References

[1] C.A. Akemann, G.A. Elliott, G.K. Pedersen and J. Tomiyama, *Derivations and multipliers of* C*-*algebras,* Amer. J. Math. **98** (1976), 679–708.

[2] C.A. Akemann and G.K. Pedersen, *Central sequences and inner derivations of separable* C*-*algebras,* Amer. J. Math.

[3] O.Yu. Aristov, *Non-vanishing of the second cohomology of certain weakly amenable Banach algebras,* Vest. Mosk. Univ. Ser. Mat. Mekh. **4** (1993), 71–74 (in Russian).

[4] O.Yu. Aristov, *The global dimension theorem for non-unital and some other separable* C*-*algebras,* Mat. Sb. **186** (9) (1995), 3–18 (in Russian).

[5] W.G. Bade, *The Wedderburn decomposition for quotient algebras arising from sets of non-synthesis,* Proceedings of the Centre for Math. Analysis, Australian National University, Vol. 21 (1989), 25–31.

[6] W.G. Bade and P.C. Curtis, *Homomorphisms of commutative Banach algebras,* Amer. J. Math. **82** (1960), 589–608.

[7] W.G. Bade and P.C. Curtis, *The Wedderburn decomposition of commutative Banach algebras,* Amer. J. Math. **82** (1960), 851–866.

[8] W.G. Bade, P.C. Curtis and H.G. Dales, *Amenability and weak amenability for Beurling and Lipschitz algebras,* Proc. London Math. Soc. **55** (1987), 359–377.

[9] B.Blackadar, *K-theory for Operator Algebras,* Springer, New York (1986).

[10] D.P. Blecher and V.I. Paulsen, *Tensor products of operator spaces,* J. Funct. Anal. **99** (1991), 262–292.

[11] F.F. Bonsall and J. Duncan, *Complete Normed Algebras,* Ergebnisse der Mathematik und ihrer Grenzgebiete, B. 80, Springer, New York (1973).

[12] S. Bowling and J. Duncan, *First order cohomology of Banach semigroup algebras,* Preprint.

[13] O. Bratelli and D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics,* I, Springer, New York (1979).

[14] J.W. Bunce, *Characterization of amenable and strongly amenable* C*-*algebras*, Pacific J. Math. **43** (1972), 563–572.
[15] J.W. Bunce, *Finite operators and amenable* C*-*algebras*, Proc. Amer. Math. Soc. **56** (1976), 145–151.
[16] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton Univ. Press, Princeton (1956).
[17] E. Christensen, *Perturbations of operator algebras*, Part I, Invent. Math. **43** (1977), 1–13; Part II, Indiana Univ. Math. J. **26** (1977), 891–904.
[18] E. Christensen, *Derivations of nest algebras*, Math. Ann. **229** (1977), 155–161.
[19] E. Christensen, *Derivations and their relation to perturbations of operator algebras*, Operator Algebras and Applications, R.V. Kadison, ed., Proc. of Symp. in Pure Math., Vol. 38, Part 2, Amer. Math. Soc., Providence, RI (1982), 261–274.
[20] E. Christensen, E.G. Effros and A.M. Sinclair, *Completely bounded multilinear maps and* C*-*algebraic cohomology*, Invent. Math. **90** (1987), 279–296.
[21] E. Christensen and D.E. Evans, *Cohomology of operator algebras and quantum dynamical semigroups*, J. London Math. Soc. **20** (1979), 358–368.
[22] E. Christensen, F. Pop, A.M. Sinclair, R.R. Smith: *On the cohomology groups of certain finite von Neumann algebras*, Preprint.
[23] E. Christensen and A.M. Sinclair, *Representations of completely bounded multilinear operators*, J. Funct. Anal. **72** (1987), 151–181.
[24] E. Christensen and A.M. Sinclair, *On the vanishing of* $\mathcal{H}^n(A, A^*)$ *for certain* C*-*algebras*, Pacific J. Math. **137** (1989), 55–63.
[25] E. Christensen and A.M. Sinclair, *A cohomological characterization of approximately finite dimensional von Neumann algebras*, Københavns Univ., Mat. Inst. Preprint Series, No. 11 (1994).
[26] E. Christensen and A.M. Sinclair, *On the Hochschild cohomology of von Neumann algebras*, Københavns Univ., Mat. Inst. Preprint Series, No. 32 (1988).
[27] J. Cigler, V. Losert and P. Michor, *Banach Modules and Functors on Categories of Banach Spaces*, Marcel Dekker, New York (1979).
[28] A. Connes, *Classification of injective factors*, Ann. of Math. **104** (1976), 73–115.
[29] A. Connes, *On the cohomology of operator algebras*, J. Funct. Anal. **28** (1978), 248–253.
[30] A. Connes, *Cohomologie cyclique et foncteurs* Ext^n, C. R. Acad. Sci. Paris Sér I Math. **296** (1983), 953–958.
[31] A. Connes, *Non-commutative differential geometry*, Part I and II, Preprint IHES (1983); Publ. Math. **62** (1986), 44–144.
[32] A. Connes, *Entire cyclic cohomology of Banach algebras and characters of* θ*-summable Fredholm modules*, K-theory **1** (1988), 519–548.
[33] J. Cuntz, *Excision in periodic cyclic theory for topological algebras*, Forschergruppe, Topologie und nichtkommutative geometrie, Preprint No. 127 (1994).
[34] P.C. Curtis, *Amenable and weakly amenable commutative Banach Algebras*, Linear and Complex Analysis Problem Book, Part I, Lecture Notes in Math., Vol. 1573, Springer (1994), 83–84.
[35] H.G. Dales, *Automatic continuity, a survey*, Bull. London Math. Soc. **10** (1978), 129–183.
[36] H.G. Dales, *Banach Algebras and Automatic Continuity*, London Mathematical Society Monographs, Oxford University Press, to appear.
[37] K.R. Davidson, *Nest Algebras*, Pitman Research Notes in Math. Series, No. 191, Longman, Harlow (Essex, UK).
[38] M. Despic and F. Ghahramani, *Weak amenability of group algebras and non-amenability of measure algebras*, Preprint.
[39] R.S. Doran and J. Wichmann, *Approximate Identities and Factorization in Banach Modules*, Springer, Berlin (1979).
[40] R.G. Douglas, C*-*algebra Extensions and* K*-homology*, Princeton Univ. Press, Princeton (1980).
[41] S. Drury, *On non-triangular sets in tensor algebras*, Studia Math. **34** (1970), 253–263.
[42] N. Dunford, *Spectral operators*, Pacific J. Math. **4** (1954), 321–354.
[43] E.G. Effros, *Advances in quantized functional analysis*, Proc. of the ICM, Berkeley (1986).
[44] E.G. Effros, *Amenability and virtual diagonals for von Neumann algebras*, J. Funct. Anal. **78** (1988), 137–153.

[45] E.G. Effros and A. Kishimoto, *Module maps and Hochschild-Johnson cohomology,* Indiana Univ. Math. J. **36** (1987), 257–276.
[46] E.G. Effros and Z.-J. Ruan, *A new approach to operator spaces,* Canad. Math. Bull. **34** (1991), 329–337.
[47] S. Eilenberg and J.C. Moore, *Foundation of Relative Homological Algebra,* Mem. Amer. Math. Soc., No. 55, 1965.
[48] J. Eschmeier and M. Putinar, *Spectral theory and sheaf theory,* III, J. Reine Angew. Math. **354** (1984), 150–163.
[49] T. Fack, *Finite sums of commutators in* C^**-algebras,* Ann. Inst. Fourier (Grenoble), **32** (1) (1982), 129–137.
[50] C. Faith, *Algebra, Rings, Modules and Categories* I, Springer, New York (1973).
[51] C. Feldman, *The Wedderburn principal theorem in Banach algebras,* Proc. Amer. Math. Soc. **2** (1951), 771–777.
[52] P. Gabriel and M. Zisman, *Calculus of Fractions and Homotopy Theory,* Springer, Berlin (1967).
[53] F. Gilfeather, *Derivations on certain CSL-algebras,* J. Operator Theory **11** (1984), 145–156.
[54] F. Gilfeather, A. Hopenwasser and D.R. Larson, *Reflexive algebras with finite width lattices: tensor products, cohomology, compact perturbations,* J. Funct. Anal. **55** (1984), 176–199.
[55] F. Gilfeather and R.R. Smith, *Operator algebras with arbitrary Hochschild cohomology,* Cont. Math. **120** (1991), 33–40.
[56] F. Gilfeather and R.R. Smith, *Cohomology for operator algebras: cones and suspensions,* Proc. London Math. Soc. **65** (1992), 175–198.
[57] F. Gilfeather and R.R. Smith, *Cohomology for operator algebras: joins,* Amer. J. Math. **116** (1994), 541–561.
[58] A.M. Gleason, *Finitely generated ideals in Banach algebras,* J. Math. Mech. **13** (1964), 125–132 (in English) = Matematika **9** (1965), 119–127 (in Russian).
[59] Yu.O. Golovin, *Homological properties of Hilbert modules over nest operator algebras,* Mat. Zametki **41** (6) (1987), 769–775 (in Russian) = Math. Notes (1987), 433–437 (in English).
[60] Yu.O. Golovin, *A criterion for the spatial projectivity of an indecomposable* CSL*-algebra of operator,* Uspekhi Mat. Nauk **49** (4) (1994), 161–162 (in Russian) = Russian Math. Surveys **49** (4) (1994), 161–162 (in English).
[61] Yu.O. Golovin, *Spatial homological dimension of operator-algebraic cones, suspensions and some joins,* Commun. Algebra **23** (9) (1995), 3447–3455.
[62] Yu.O. Golovin, *The spatial flatness and injectivity for an indecomposable* CSL*-algebra,* Preprint.
[63] Yu.O. Golovin and A.Ya. Helemskii, *The homological dimension of certain modules over a tensor product of Banach algebras,* Vest. Mosk. Univ. Ser. Mat. Mekh. **1** (1977), 54–61 = Moscow Univ. Math. Bull. **32** (1977), 46–52.
[64] N. Grønbaek, *Amenability of weighted convolution algebras on locally compact groups,* Trans. Amer. Math. Soc., Preprint.
[65] N. Grønbaek, *Amenability of weighted discrete convolution algebras on cancellative semigroups,* Proc. Royal Soc., Edinburgh **110A** (1988), 351–360.
[66] N. Grønbaek, *A characterization of weakly amenable Banach algebras,* Studia Math. **94** (1989), 149–162.
[67] N. Grønbaek, *Morita equivalence for Banach algebras,* J. Pure Appl. Algebra **99** (1995), 183–219.
[68] N. Grønbaek, *Morita equivalence for self-induced Banach algebras,* Københavns Univ., Mat. Inst. Preprint Series, No. 24 (1994).
[69] N. Grønbaek, *Amenability and weak amenability of tensor algebras and algebras of nuclear operators,* Københavns Univ., Mat. Inst. Preprint Series, No. 23 (1989).
[70] N. Grønbaek, B.E. Johnson and G.A. Willis, *Amenability of Banach algebras of compact operators,* Københavns Univ. Mat. Inst. Preprint Series, No. 13 (1992).
[71] A. Grothendieck, *Produits Tensoriels Topologiques et Espaces Nucleaires,* Mem. Amer. Math. Soc., No. 16 (1955).
[72] A. Guichardet, *Sur l'homologie et la cohomologie des algebres de Banach,* C. R. Acad. Sci. Paris **262** (1966), A38–41.
[73] R.N. Gumerov, *Homological dimension of radical algebras of Beurling type with rapidly decreasing weight,* Vest. Mosk. Univ. Ser. Mat. Mekh. **5** (1988), 18–22 (in Russian).

[74] R.C. Gunning and H. Rossi, *Analytic Functions of Several Complex Variables,* Prentice-Hall, Englewood Cliffs, NJ (1965).

[75] U. Haagerup, *All nuclear* C^**-algebras are amenable,* Invent. Math. **74** (1983), 305–319.

[76] P. de la Harpe, *Moyennabilité du groupe unitaire et propriété P de Schwartz des algébres de von Neumann,* Algebres d'Operateurs Seminaire, Les Plans-sur-Bex, Suisse, 1978, P. de la Harpe, ed., Lecture Notes in Math., Vol. 725, Springer, Berlin (1979), 220–227.

[77] A.Ya. Helemskii, *Annihilator extensions of commutative Banach algebras,* Izv. Akad. Nauk SSSR Ser. Mat. **29** (1965), 945–956 (in Russian).

[78] A.Ya. Helemskii, *Strong decomposability of finite-dimensional extensions of* C^**-algebras,* Vest. Mosk. Univ. Ser. Mat. Mekh. **4** (1967), 50–52 (in Russian).

[79] A.Ya. Helemskii, *The homological dimension of normed modules over Banach algebras,* Mat. Sb. **81** (123) (1970), 430–444 (in Russian) = Math. USSR Sb. **10** (1970), 399–411 (in English).

[80] A.Ya. Helemskii, *The homological dimension of Banach algebras of analytic functions,* Mat. Sb. **83** (125) (1970), 222–233 (in Russian) = Math. USSR Sb. **12** (1970), 221–233 (in English).

[81] A.Ya. Helemskii, *A description of relatively projective ideals in the algebra* $C(\Omega)$, Doklad. Akad. Nauk SSSR **195** (1970), 1286–1289 (in Russian) = Soviet Math. Dokl. **1** (1970), 1680–1683 (in English).

[82] A.Ya. Helemskii, *The periodic product of modules over Banach algebras,* Funct. Anal. i Pril. **5** (1971), 95–96 (in Russian) = Functional Anal. Appl. **5** (1971), 84–85 (in English).

[83] A.Ya. Helemskii, *A certain class of flat Banach modules and its applications,* Vest. Mosk. Univ. Ser. Mat. Mekh. **27** (1) (1972), 29–36 (in Russian).

[84] A.Ya. Helemskii, *A certain method of computing and estimating the global homological dimension of Banach algebras,* Mat. Sb. **87** (129) (1972), 122–135 (in Russian) = Math. USSR Sb. **16** (1972), 125–138 (in English).

[85] A.Ya. Helemskii, *The global dimension of a Banach function algebra is different from one,* Funct. Anal. i Pril. **6** (1972), 95–96 (in Russian) = Functional Anal. Appl. **6** (1972), 166–168 (in English).

[86] A.Ya. Helemskii, *The lowest values taken by the global homological dimensions of functional Banach algebras,* Trudy Sem. Petrovsk. **3** (1978), 223–242 (in Russian) = Amer. Math. Soc. Trans. **124** (1984), 75–96 (in English).

[87] A.Ya. Helemskii, *Homological methods in the holomorphic calculus of several operators in Banach space after Taylor,* Uspekhi Matem. Nauk **36** (1981), 127–172 (in Russian) = Russian Math. Surveys **36** (1) (1981) (in English).

[88] A.Ya. Helemskii, *Flat Banach modules and amenable algebras,* Trudy MMO **47** (1984), 179–218 (in Russian) = Trans. Moscow Math. Soc. (1985), 199–244 (in English).

[89] A.Ya. Helemskii, *The Homology of Banach and Topological Algebras,* Kluwer, Dordrecht (1989).

[90] A.Ya. Helemskii, *Some remarks about ideas and results of topological homology,* Proc. Centre for Math. Anal., Australian National University, Vol. 21 (1989), 203–238.

[91] A.Ya. Helemskii, *Homological algebra background of the "amenability after A. Connes": injectivity of the predual bimodule,* Mat. Sb. **180** (1989), 1680–1690 (in Russian).

[92] A.Ya. Helemskii, *Homology in Banach and polynormed algebras: some results and problems,* Operator Theory, Advances and Applications, Vol. 42, Birkhäuser, Basel (1990), 195–208.

[93] A.Ya. Helemskii, *From the topological homology: algebras with various conditions of the homological triviality,* Algebraic Problems of the Analysis and Topology, Voronezh University Press (1990), 77–96 (in Russian) = Lecture Notes in Math., Vol. 1520, Springer (1992), 19–40 (in English).

[94] A.Ya. Helemskii, *Banach cyclic (co)homology and the Connes–Tzygan exact sequence,* J. London Math. Soc. (2) **46** (1992), 449–462.

[95] A.Ya. Helemskii, *Banach cyclic (co)homology in terms of Banach derived functors,* Algebra i Analiz **3** (1991), 213–228 (in Russian) = St. Petersburg Math. J. **3** (1992), 1149–1164 (in English).

[96] A.Ya. Helemskii, *Banach and Locally Convex Algebras, General Theory, Representations, Homology,* Oxford University Press, Oxford (1993).

[97] A.Ya. Helemskii, *A description of spatially projective von Neumann algebras,* J. Operator Theory **32** (1994), 381–398.

[98] A.Ya. Helemskii, *The spatial flatness and injectivity of Connes operator algebras,* Extracta Math. **9** (1) (1994), 75–81.

[99] A.Ya. Helemskii, *31 problems on the homology of the algebras of analysis,* Linear and Complex Analysis Problem Book, Part I, Lecture Notes in Math., Vol. 1573, Springer, Berlin (1994), 54–78.

[100] A.Ya. Helemskii, *A homological characterization of type I factors,* Doklady RAN **344** (4) (1995).

[101] A.Ya. Helemskii and Z.A. Lykova, *On the normal version of the simplicial cohomology of operator algebras,* Bull. Austral. Math. Soc. **48** (1993), 407–410.

[102] A.Ya. Helemskii and M.V. Sheinberg, *Amenable Banach algebras,* Funct. Anal. i Pril. **13** (1979), 42–48 (in Russian) = Functional Anal. Appl. **13** (1979), 32–37 (in English).

[103] G. Hochschild, *On the cohomology groups of an associative algebra,* Ann. of Math. **46** (1945), 58–76.

[104] G. Hochschild, *Relative homological algebra,* Trans. Amer. Math. Soc. **82** (1956), 246–249.

[105] K. Jarosz, *Perturbations of Banach Algebras,* Springer, Berlin (1980).

[106] B.E. Johnson, *The Wedderburn decomposition of Banach algebras with finite-dimensional radical,* Amer. J. Math. **90** (1968), 866–876.

[107] B.E. Johnson, *An introduction to the theory of centralizers,* Proc. London Math. Soc. **14** (1964), 299–320.

[108] B.E. Johnson, *Approximate diagonals and cohomology of certain annihilator Banach algebras,* Amer. J. Math. **94** (1972), 685–698.

[109] B.E. Johnson, *Cohomology in Banach Algebras,* Mem. Amer. Math. Soc., No. 127 (1972).

[110] B.E. Johnson, *A class of II_1 factors without property P but with zero second cohomology,* Arkiv för Math. **12** (1974), 153–159.

[111] B.E. Johnson, *Introduction to cohomology in Banach algebras,* Algebras in Analysis, J.H. Williamson, ed., Acad. Press, London (1975), 84–99.

[112] B.E. Johnson, *Perturbations of Banach algebras,* Proc. London Math. Soc. **34** (1977), 439–458.

[113] B.E. Johnson, *Stability of Banach algebras,* Lecture Notes in Math., Vol. 660, Springer (1978), 109–114.

[114] B.E. Johnson, *Low dimensional cohomology of Banach algebras,* Operator Algebras and Applications, R.V. Kadison, ed., Proc. of Symp. in Pure Math., Providence, Amer. Math. Soc., Part 2 (1982), 253–259.

[115] B.E. Johnson, *Weak amenability of group algebras,* Bull. London Math. Soc. **23** (1991), 281–284.

[116] B.E. Johnson, R.V. Kadison and J.R. Ringrose, *Cohomology of operator algebras,* III, *Reduction to normal cohomology,* Bull. Soc. Math. France **100** (1972), 73–96.

[117] B.E. Johnson and S.K. Parrot, *Operators commuting with a von Neumann algebra modulo the set of compact operators,* J. Funct. Analysis **11** (1972), 39–61.

[118] L. Kadison, *On the cyclic cohomology of nest algebras and a spectral sequence induced by a subalgebra in Hochschild cohomology,* C. R. Acad. Sci. Paris, Serie I **311** (1990), 247–252.

[119] R.V. Kadison, *Derivations of operator algebras,* Ann. of Math. **83** (1966), 280–293.

[120] R.V. Kadison and D. Kastler, *Perturbations of von Neumann algebras* I, *Stability of type,* Amer. J. Math. **94** (1972), 38–54.

[121] R.V. Kadison and J.R. Ringrose, *Cohomology of operator algebras* I, *Type* I *von Neumann algebras,* Acta Math. **126** (1971), 227–243.

[122] R.V. Kadison and J.R. Ringrose, *Cohomology of operator algebras* II, *Extended cobounding and the hyperfinite case,* Ark. Mat. **9** (1971), 55–63.

[123] R.V. Kadison and J.R. Ringrose, *Fundamentals of the Theory of Operator Algebras,* Vol. II, Academic Press, London (1986).

[124] J.P. Kahane, *Séries de Fourier Absolument Convergentes,* Ergebnisse der Mathematik und ihrer Grenzgebiete, B. 50, Springer, Berlin (1970).

[125] S.I. Kaliman and Yu.V. Selivanov, *On the cohomology of operator algebras,* Vest. Mosk. Univ. Ser. Mat. Mekh. **29** (5) (1974), 24–27 (in Russian) = Moscow Univ. Math. Bull. **29** (5) (1974), 20–22 (in English).

[126] E. Kallin, *A non-local function algebra,* Proc. Nat. Acad. Sci. **49** (1963), 821–824.

[127] H. Kamowitz, *Cohomology groups of commutative Banach algebras,* Trans. Amer. Math. Soc. **102** (1962), 352–372.

[128] H. Kamowitz and S. Scheinberg, *Derivations and automorphisms of $L^1[0,1]$,* Trans. Amer. Math. Soc. **135** (1969), 415–427.

[129] I. Kaplansky, *On the dimension of modules and algebras,* X, Nagoya Math. J. **13** (1958), 85–88.

[130] A.M. Karyaev and N.V. Yakovlev, *On the automatic complementability of certain classes of annihilator ideals in Banach algebras,* Vest. Mosk. Univ. Ser. Mat. Mekh. **3** (1990), 95–98 (in Russian).

[131] D. Kastler, *Entire cyclic cohomology of $\mathbf{Z}/2$-graded Banach Algebras,* K-theory **1** (1989), 485–509.

[132] A.A. Kirillov and A.D. Gvishiani, *Theorems and Problems in Functional Analysis,* Nauka, Moscow (1979) (in Russian).

[133] A.S. Kleshchev, *The homological dimension of Banach algebras of smooth functions is infinity,* Vest. Mosk. Univ. Ser. Mat. Mekh. **6** (1988), 57–60 (in Russian).

[134] L. Koszul, *Homologie et cohomology des algébres de Lie,* Bull. Soc. Math. France **78** (1950), 65–127.

[135] R.L. Krasauskas, S.V. Lapin and Yu.P. Solov'ev, *Dihedral homology and cohomology. Basic concepts and constructions,* Mat. Sb. **133** (175) (1987), 24–48 (in Russian) = Math. USSR Sb. **61** (1988) (in English).

[136] A.N. Krichevets, *On the connection of homological properties of some Banach modules and questions of the geometry of Banach spaces,* Vest. Mosk. Univ. Ser. Mat. Mekh. **2** (1981), 55–58 (in Russian) = Moscow Univ. Math. Bull. **36** (1981), 68–72 (in English).

[137] A.N. Krichevets, *Calculation of the global dimension of tensor products of Banach algebras and a generalization of a theorem of Phillips,* Mat. Zametki **31** (1982), 187–202 = Math. Notes **31** (1982), 95–104.

[138] A.N. Krichevets, *On the homological dimension of* $C(\Omega)$, VINITI Preprint No. 9012-V86 (1986) (in Russian).

[139] E.Sh. Kurmakaeva, *The impact of the topology of* Ω *on the strict homological dimension of* $C(\Omega)$, Mat. Zametki **55** (3) (1994), 76–83 (in Russian).

[140] E.C. Lance, *Cohomology and perturbations of nest algebras,* Proc. London Math. Soc. **43** (1981), 334–356.

[141] A.T. Lau, *Characterizations of amenable Banach algebras,* Proc. Amer. Math. Soc. **70** (1978), 156–160.

[142] G. Lindblad, *Dissipative operators and cohomoly of operator algebras,* Lett. Math. Phis. **1** (1976), 219–224.

[143] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces II, Sequences Spaces,* Ergebnisse Math. 92, Springer, New York (1977).

[144] J.-L. Loday, *Cyclic Homology,* Springer, Berlin (1992).

[145] Z.A. Lykova, *On homological characteristics of operator algebras,* Vest. Mosk. Univ. Ser. Mat. Mekh. **1** (1986), 8–13 (in Russian).

[146] Z.A. Lykova, *Structure of Banach algebras with trivial central cohomology,* J. Operator Theory **28** (1992), 147–165.

[147] Z.A. Lykova, *On non-zero values of the central homological dimension of* C^**-algebras,* Glasgow Math. J. **34** (1992), 369–378.

[148] Z.A. Lykova, *The lower estimate of the global homological dimension of infinite-dimensional* CCR-*algebras,* Operator Algebras and Topology, W.B. Arveson, A.S. Mishchenko, M. Putinar, M.A. Rieffel, S. Stratila, eds, Longman (1992), 93–129.

[149] S. MacLane, *Homology,* Die Grundlehren der Mathematischen Wissenschaften 114, Springer, Berlin (1963).

[150] A. Mallios, *Topological Algebras, Selected Topics,* North-Holland, Amsterdam (1986).

[151] S.S. Mel'nikov, *On dihedral Banach cohomologies,* Uspekhi Mat. Nauk **48** (6) (1993), 161–162 (in Russian) = Russian Math. Surveys **48** (6) (1993), 171–172 (in English).

[152] W. Moran, *The global dimension of* $C(X)$, J. London Math. Soc. (2) **17** (1978), 321–329.

[153] J.P. Nielsen, *Cohomology of some non-selfadjoint operator algebras,* Math. Scand. **47** (1980), 150–156.

[154] O.S. Ogneva, *Coincidence of homological dimensions of Fréchet algebra of smooth functions on a manifold with the dimension of the manifold,* Funct. Anal i Pril. **20** (1986), 92–93 (in Russian).

[155] O.S. Ogneva, *On homological dimensions of certain algebras of test functions and distributions on Abelian compactly-generated Lie groups,* VINITI, No. 4591–B.86 (1986) (in Russian).

[156] O.S. Ogneva and A.Ya. Helemskii, *Homological dimensions of certain algebras of fundamental (test) functions,* Mat. Zametki **35** (1984), 177–187 (in Russian).

[157] A.Yu. Ol'shanskii, *On the question of the existence of an invariant mean on a group,* Uspekhi Mat. Nauk **35** (1980), 199–200 (in Russian).

[158] T.W. Palmer, *Banach Algebras and the General Theory of* **-Algebras,* Vol. I, Cambridge University Press, Cambridge (1994).

[159] T.W. Palmer, *Banach Algebras and the General Theory of* **-Algebras,* Vol. II, to appear.

[160] A.L.T. Paterson, *Amenability,* AMS, Providence, RI (1988).

[161] A.L.T. Paterson, *Invariant mean characterizations of amenable operator algebras,* Preprint.

[162] A.L.T. Paterson, *Nuclear* C^**-algebras have amenable unitary groups,* Preprint.

[163] A.L.T. Paterson, *Virtual diagonals and n-amenability for Banach algebras,* Preprint.

[164] A.L.T. Paterson and R.R. Smith, *Higher-dimensional virtual diagonals and ideal cohomology for triangular algebras,* Preprint.

[165] G.K. Pedersen, *C^*-Algebras and Their Automorphism Groups,* Academic Press, London (1979).

[166] J. Phillips and I. Raeburn, *Perturbations of C^*-algebras,* II, Proc. London Math. Soc. (3) **43** (1981), 46–72.

[167] J. Phillips and I. Raeburn, *Central cohomology of C^*-algebras,* J. London Math. Soc. (2) **28** (1983), 363–372.

[168] S. Popa, *The commutant modulo the set of compact operators of a von Neumann algebra,* J. Funct. Anal. **71** (1987), 393–408.

[169] S. Pott, *An account on the Global Homological Dimension Theorem of A.Ya. Helemskii,* The Univ. of Leeds, Preprint (1994).

[170] L.I. Pugach, *Flat ideals of Banach algebras,* Vest. Mosk. Univ. Ser. Mat. Mekh. **1** (1981), 59–63 (in Russian) = Moscow Univ. Math. Bull. **36** (1981), 70–74 (in English).

[171] L.I. Pugach, *Projective and flat ideals of function algebras and their connection with analytic structure,* Mat. Zametki **31** (1982), 239–245 (in Russian).

[172] L.I. Pugach, *Homological properties of functional algebras and analytic polydiscs in their maximal ideal spaces,* Rev. Roumaine Math. Pure Appl. **31** (1986), 347–356 (in Russian).

[173] L.I. Pugach, *Projective ideals of function algebras, and the Shilov boundary,* Problems of Group Theory and Homological Algebra, Yaroslavl Univ. Press (1987), 22–24 (in Russian).

[174] M. Putinar, *Functional calculus with sections of an analytic space,* J. Operator Theory **4** (1980), 297–306.

[175] F. Radulescu, *Vanishing of $H^2_w(M, K(H))$ for certain finite von Neumann algebras,* Trans. Amer. Math. Soc. **326** (1994), 569–584.

[176] I. Raeburn and J.L. Taylor, *Hochschild cohomology and perturbations of Banach algebras,* J. Funct. Anal. **25** (1977), 258–266.

[177] T.T. Read, *The power of a maximal ideal in a Banach algebra and analytic structure,* Trans. Amer. Math. Soc. **161** (1971), 235–248.

[178] M.A. Rieffel, *Induced Banach representations of Banach algebras and locally compact groups,* J. Funct. Anal. **1** (1967), 443–491.

[179] J.R. Ringrose, *On some algebras of operators,* Proc. London Math. Soc. **15** (3) (1965), 61–83.

[180] J.R. Ringrose, *Cohomology of operator algebras,* Lectures in operator Algebras, Lecture Notes in Math., Vol. 247, Springer, Berlin (1972), 355–434.

[181] J.R. Ringrose, *Cohomology theory for operator algebras,* Operator Algebras and Applications, R.V. Kadison, ed., Proc. of Symp. in Pure Math., Part 2, Providence, Amer. Math. Soc. (1982), 229–252.

[182] R. Rochberg, *The disc algebra is rigid,* Proc. London Math. Soc. **39** (3) 1979, 93–118.

[183] J. Rosenberg, *Homological invariants of extensions of C^*-algebras,* Operator Algebras and Applications, R.V. Kadison, ed., Proc. of Symp. in Pure Math., Vol. 38, Part I, Amer. Math. Soc., Providence, RI (1982), 35–75.

[184] J. Rosenberg, *Algebraic K-Theory and its Applications,* Springer, Berlin (1994).

[185] Z.-J. Ruan, *Subspaces of C^*-algebras,* J. Funct. Anal. **76** (1988), 217–230.

[186] V. Runde, *A functorial approach to weak amenability for commutative Banach algebras,* Glasgow Math. J. **34** (1992).

[187] S. Sakai, *Derivations of W^*-algebras,* Ann. of Math. **83** (1966), 273–279.

[188] H. Schaeffer, *Topological Vector Spaces,* Springer, New York (1971).

[189] Yu. V. Selivanov, *The values assumed by the global dimension in certain classes of Banach algebras,* Vest. Mosk. Univ. Ser Mat. Mekh. **30** (1) (1975), 37–42 (in Russian) = Moskow Univ. Math. Bull. **30** (1) (1975), 30–34 (in English).

[190] Yu.V. Selivanov, *Homological dimension of cyclic Banach modules and homological characterization of metrizable compacta,* Mat. Zametki **17** (1975), 301–305 (in Russian) = Math. Notes **17** (1975), 174–176 (in English).

[191] Yu.V. Selivanov, *Biprojective Banach algebras, their structure, cohomologies, and connection with nuclear operators,* Funct. Anal. i Pril. **10** (1976), 89–90 (in Russian) = Functional Anal. Appl. **10** (1976), 78–79 (in English).

[192] Yu.V. Selivanov, *Banach algebras of small global dimension zero,* Uspekhi Mat. Nauk **31** (1976), 227–228 (in Russian).

[193] Yu.V. Selivanov, *Biprojective Banach algebras,* Izv. Akad. Nauk SSSR Ser. Mat. **43** (1979), 1159–1174 (in Russian) = Math. USSR Izvestija **15** (1980), 387–399 (in English).

[194] Yu.V. Selivanov, *Homological "additivity formulae" for Banach algebras,* Int. Conf. on Algebra (Barnaul 20–25 Aug., 1991), Novosibirsk (1991), p. 78 (in Russian).

[195] Yu.V. Selivanov, *Homological characterizations of the approximation property for Banach spaces,* Glasgow Math. J. **34** (1992), 229–239.

[196] Yu.V. Selivanov, *Computing and estimating the global dimension in certain classes of Banach algebras,* Math. Scand. **72** (1993), 85–98.

[197] Yu.V. Selivanov, *Projective ideals and estimating the global dimension of commutative Banach algebras,* Third Int. Conf. on Algebra in Honour of M. I. Kargopolov: Abstract of Reports, Krasnojarsk (1993), 297–298 (in Russian).

[198] Yu.V. Selivanov, *Projective Fréchet modules with the approximation property,* Uspekhi Mat. Nauk **50** (1) (1995), 209–210 (in Russian).

[199] Yu.V. Selivanov, *Fréchet algebras of global dimension zero,* Algebra, Proc. of the 3rd Int. Conf. on Algebra held in Krasnoyarsk, Russia, 23–28 August, 1993, Walter de Gruyter (1996), 225–236.

[200] Yu.V. Selivanov, *Weak homological bidimension and its values in the class of biflat Banach algebras,* Extracta Math. (to appear).

[201] M.V. Sheinberg, *Homological properties of closed ideals that have a bounded approximate unit,* Vest. Mosk. Univ. Ser. Mat. Mekh. **27** (4) (1972), 39–45 (in Russian) = Moscow Univ. Math. Bull. **27** (1972), 103–108 (in English).

[202] M.V. Sheinberg, *On the relative homological dimension of group algebras of locally compact groups,* Izv. Akad. Nauk SSSR Ser. Mat. **37** (1973), 308–318 (in Russian) = Math. USSR Izvestija **7** (1973), 307–317 (in English).

[203] M.V. Sheinberg, *A characterization of the algebra* $C(\Omega)$ *in terms of cohomology groups,* Uspekhi Mat. Nauk **32** (1977), 203–204 (in Russian).

[204] A.M. Sinclair and R.R. Smith, *Hochschild Cohomology of von Neumann Algebras,* Cambridge University Press, Cambridge (1995).

[205] I.M. Singer and J. Wermer, *Derivations on commutative normed algebras,* Math. Ann. **129** (1955), 260–264.

[206] J.L. Taylor, *A joint spectrum for several commuting operators,* J. Funct. Anal. **6** (1970), 172–191.

[207] J.L. Taylor, *Analytic functional calculus for several commuting operators,* Acta Math. **125** (1970), 1–40.

[208] J.L. Taylor, *A general framework for a multi-operator functional calculus,* Adv. Math. **9** (1972), 183–252.

[209] J.L. Taylor, *Homology and cohomology for topological algebras,* Adv. Math. **9** (1972), 137–182.

[210] J.L. Taylor, *Functions of several noncommuting variables,* Bull. Amer. Math. Soc. **79** (1973), 1–34.

[211] M.P. Thomas, *The image of a derivation is contained in the radical,* Ann. of Math. **128** (1988), 435–460.

[212] J.Tomiyama, *On the projection of norm one in* W^**-algebras,* Proc. Japan. Acad. **33** (1957), 608–612.

[213] B.L. Tzygan, *Homology of matrix Lie algebras over rings and Hochschild homology,* Uspekhi Mat. Nauk **38** (1983), 217–218 (in Russian).

[214] N.Th. Varopoulos, *Tensor algebras and harmonic analysis,* Acta Math. **119** (1967), 51–112.

[215] S.Wasserman, *On tensor products of certain group* C^**-algebras,* J. Funct. Anal. **23** (1976), 239–254.

[216] J.G. Wendel, *Left centralizers and isomorphisms of group algebras,* Pacific J. Math. **2** (1952), 251–261.

[217] M. Wodzicki, *The long exact sequence in cyclic homology associated with an extension of algebras,* C. R. Acad. Sci. Paris **306** (1988), 399–403.

[218] M. Wodzicki, *Vanishing of cyclic homology of stable* C^**-algebras,* C. R. Acad. Sci. Paris **307** (1988), 329–334.

[219] M. Wodzicki, *Homological properties of rings of functional-analytic type,* Proc. National Academy Science, USA, Preprint.

[220] N.V. Yakovlev, *Examples of Banach algebras with a radical that can not be complemented as a Banach space,* Uspekhi Mat. Nauk **44** (5) (1989), 185–186 (in Russian) = Russian Math Surveys **44** (5) (1989), 224–225 (in English).

Section 2D
Model Theoretic Algebra

This Page Intentionally Left Blank

Stable Groups

Frank Wagner*
Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, United Kingdom
E-mail: wagner@maths.ox.ac.uk

Contents

*Heisenberg-Stipendiat der Deutschen Forschungsgemeinschaft (Wa 899/2-1). Current address: Institut Girard Desargues, Université Claude Bernard (Lyon-1), Mathématiques, bâtiment 101, 43, boulevard du 11 novembre 1918, 69622 Villeurbanne-cedex, France. E-mail: wagner@desargues.univ-lyon1.fr.

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

1. Introduction

Stability theory – or, as it was then called, classification theory – has been invented in the early 1970's by Shelah [165] as a tool for the classification of the models of a complete first-order theory. Through the work of Zil'ber [200], however, it soon became apparent that the notion has a distinctly algebro-geometric flavour, and that many of the methods and results from algebraic geometry should generalize to the wider setting of stable theories, and that algebraic objects, in particular groups and fields, would play a central part in the development of the theory.

Roughly speaking, stability theory provides a notion of independence which generalizes algebraic independence; in nice cases (superstable structures) there also is a dimension, called *rank*, which takes ordinal values. Thus we may obtain a whole hierarchy of simultaneous finitely-valued dimension theories, one at each level ω^α, each coarser than the preceding one, allowing us to rank and compare definable sets of entirely different magnitude, and study their interactions.

A structure is *strongly minimal* if every definable subset is either finite or cofinite, uniformly in parameters. These are the simplest stable structures; examples are

(1) a set without structure (with only equality),
(2) a vector space over a division ring D (with addition and unary functions for scalar multiplication by d, for every $d \in D$),
(3) an algebraically closed field (with addition and multiplication).

If A is a subset of a structure $\mathfrak{M}$, the *algebraic closure* $\mathrm{acl}(A)$ of A is the set of elements of $\mathfrak{M}$ which lie in some finite A-definable set; if $\mathfrak{M}$ is strongly minimal, algebraic closure induces a combinatorial pre-geometry on $\mathfrak{M}$, a dependence relation (x depends on y over A if $x \in \mathrm{acl}(A, y) - \mathrm{acl}(A)$), and a dimension. (In the examples above, this dependence relation is just equality, linear and algebraic dependence, respectively, and dimension is cardinality, vector space dimension, and transcendence degree.) $\mathfrak{M}$ is *locally modular* if $\dim(A \cup B) + \dim(A \cap B) = \dim(A) + \dim(B)$ whenever $A, B \subseteq \mathfrak{M}$ are algebraically closed with $A \cap B \neq \emptyset$. Zil'ber conjectured:

TRICHOTOMY CONJECTURE. Every strongly minimal structure either has degenerate pre-geometry (if a depends on $A \cup B$, it depends on A or on B), or is bi-interpretable with a vector space over a division ring, or with an algebraically closed field.

He then proceeded to show that a locally finite locally modular strongly minimal set interprets a definable Abelian strongly minimal group (the *group configuration theorem* [205]), and that under some circumstances certain automorphism groups of a structure are interpretable in it (the *binding groups* [202]). Both of these theorems were generalized by Hrushovski to a much wider context [80]. As groups interpretable in a stable structure are again stable, this forged a link between abstract model theory and the study of stable groups.

In 1988 Hrushovski [87] constructed a new kind of strongly minimal set which is neither locally modular (ruling out degenerate pre-geometries and vector spaces) nor interprets any group at all, thus refuting the Trichotomy Conjecture. However, the trichotomy was shown to hold for strongly minimal subsets of many algebraically interesting structures (algebraically closed fields, differentially closed fields, separably closed fields, existentially closed difference fields) [117,43]; in fact in [95,96] Hrushovski and Zil'ber created an abstract framework (called *Zariski Geometries*) in which the Trichotomy Conjecture is true, leading to applications in algebraic geometry and number theory (Hrushovski's proofs of the Mordell–Lang and Manin–Mumford conjectures, see Section 7).

Let us call a structure $\mathfrak{M}$ *almost strongly minimal* if it has a definable strongly minimal subset X such that $\mathfrak{M} = \mathrm{acl}(X)$. Then a particular consequence of the Trichotomy Conjecture would be

CHERLIN'S CONJECTURE [56]. An almost strongly minimal simple group is an algebraic group over an algebraically closed field.

This conjecture is still open; if true, it would mean that from the existence of a dimension function on the definable sets of a simple group one can recover the Zariski topology. In analysing a minimal counter-example to the conjecture [56,121,59,39], one quickly came up with two main obstacles: the possible existence of an almost strongly minimal simple group whose soluble subgroups are all nilpotent-by-finite (a *bad group*), and the possible existence of an almost strongly minimal field with a definable infinite proper multiplicative subgroup (a *bad field*). Recent results by Poizat [156] indicate that bad fields may exist; as for the existence of bad groups, it seems likely that they may be constructed using the methods of Olshanski'i [97] – but proving almost strong minimality for such a group would be very difficult.

Struck by the similarity between almost strongly minimal simple groups and finite simple groups of Lie type, Borovik and Nesin [37] and their collaborators developed a programme to prove Cherlin's Conjecture under the additional assumption that there were no bad groups or fields interpretable, which proceeds in analogy (albeit on a much smaller scale) to the Lie case of the classification of finite simple groups. This has been quite successful; the cases of even characteristic and large groups in odd characteristic have almost been dealt with. However, a lot of work remains to be done in the case of small groups of odd characteristic.

As I see it, the interest in stable groups is thus threefold. Firstly, they serve as a tool for the model-theoretic analysis of arbitrary structures, via the group configuration and the binding group. Secondly, they form an interesting class of infinite groups which is amenable to group-theoretic investigation. And thirdly, they encompass various important specific algebraic structures, whose detailed model-theoretic study has been very fruitful.

After a quick review of model theory, we shall define model-theoretic stability in Section 2 and survey its main properties. In Section 3 we shall derive the various chain conditions on subgroups implied by stability, in particular the chain condition on centralizers, and study groups with chain conditions. Section 4 introduces generic types; we shall see that type-definable groups, or generically given groups, are embedded into definable groups. Superstable groups and fields are dealt with in Section 5, which also contains Zil'ber's Indecomposability Theorem. In Section 6 we look at groups of finite Morley rank and Cherlin's Conjecture. One-based groups (satisfying a strong form of local modularity) are treated in Section 7, the group configuration and the binding group in Section 8. We discuss additional results and recent developments in Section 9.

2. Model-theoretic preliminaries

In a model-theoretic analysis of a structure, it is important to specify the language used. For instance, the complex numbers in the ring language $\{0, 1, +, -, *\}$ will have very different properties to the complex numbers in the ring language enriched by the modulus function; less trivially, a vector space V over a field K in the module language $\{0, +, \lambda_k\colon k \in K\}$ is very different from the two-sorted structure (V, K), where we have addition on V, addition and multiplication on K, and an application function $(k, v) \mapsto kv$. As we shall be dealing with groups, our structures will usually have a symbol $*$ for multiplication, but may carry additional constants, functions and relations. A structure for the language $\mathcal{L}$ is an $\mathcal{L}$-*structure*; given an $\mathcal{L}$-structure $\mathfrak{M}$ and a subset A of $\mathfrak{M}$, we can expand the language by adding constants for elements of A. This new language is denote by $\mathcal{L}(A)$; the expansion is called *inessential* and is usually harmless.

Having fixed the language $\mathcal{L}$, we can define the $\mathcal{L}$-*terms* recursively: any constant or variable is a term, and if $t_1, \ldots, t_n$ are terms, so is $f(t_1, \ldots, t_n)$, for any n-ary function $f \in \mathcal{L}$. An *atomic* formula is one of the form $R(t_1, \ldots, t_n)$, where R is an n-ary relation symbol (for instance, equality is a binary relation), and $t_1, \ldots, t_n$ are again terms. More general formulas are obtained by means of Boolean combinations and quantification: if φ and ψ are formulas, so are $\neg\varphi$, $\phi \wedge \psi$, $\phi \vee \psi$, $\exists x\varphi$ and $\forall x\varphi$. A *sentence* is a formula whose variables are all in the scope of some quantifier; we say that $\mathfrak{M}$ *satisfies* a sentence φ (denoted $\mathfrak{M} \models \varphi$) if the natural interpretation of φ in $\mathfrak{M}$ holds. A *theory* is a collection of $\mathcal{L}$-sentences; an $\mathcal{L}$-structure $\mathfrak{M}$ is a *model* for some theory T if every $\varphi \in T$ is satisfied by $\mathfrak{M}$. A theory is *complete* if it implies every sentence or its negation; given an $\mathcal{L}$-structure $\mathfrak{M}$ we obtain a complete theory $\mathrm{Th}(\mathfrak{M})$ as the set of all $\mathcal{L}$-sentences satisfied by (or *true in*) $\mathfrak{M}$. Two $\mathcal{L}$-structures are *elementarily equivalent* if they have the same theory; a substructure $\mathfrak{M}$ of an $\mathcal{L}$-structure $\mathfrak{N}$ is *elementary*, denoted $\mathfrak{M} \prec \mathfrak{N}$, if it is $\mathcal{L}(\mathfrak{M})$-elementarily equivalent to $\mathfrak{N}$. The theorem of Löwenheim–Skolem states that every infinite $\mathcal{L}$-structure $\mathfrak{M}$ has an elementary substructure in every infinite cardinality between $|\mathcal{L}|$ and $|\mathfrak{M}|$, and an elementary superstructure in every cardinality greater than $|\mathfrak{M}|$. Hence the best one can hope for is that a theory describes its models up to cardinality.

DEFINITION 1. A theory T is κ-*categorical* if all its models of cardinality κ are isomorphic; it is *totally categorical* if it is categorical in all infinite cardinalities.

A famous theorem by Morley [120] states that a countable theory is ω_1-categorical if and only if it is categorical in every uncountable cardinality. A dense linear order is ω-categorical, a vector space over a finite field is totally categorical, and an algebraically closed field is uncountably categorical (as is any strongly minimal, or almost strongly minimal, set).

If $\mathfrak{M}$ is an $\mathcal{L}$-structure and $n < \omega$, a subset $X \subseteq \mathfrak{M}^n$ is *definable* if it is of the form $\varphi(\bar{x})^{\mathfrak{M}} = \{\overline{m} \in \mathfrak{M}^n\colon \mathfrak{M} \models \varphi(\overline{m})\}$ for some $\mathcal{L}$-formula $\varphi(\bar{x})$. Examples of definable subsets of a group are the centralizer of an element g (via the formula $xg = gx$), the subset of elements of order dividing n (defined by the formula $x * \cdots * x = 1$, abbreviated to $x^n = 1$), and the centre $Z(G)$ (given by $\forall y\ xy = yx$). Subsets not usually definable are those which involve a recursive process, like the group generated by a finite tuple, or the commutator subgroup (unless there is a bound on the length of a sum of commutators needed to express an element of G'). A set is definable *over* A if it can be expressed by a formula with parameters from A.

A subset $X \subseteq \mathfrak{M}^n$ is *type-definable* if it is given by an infinite intersection of definable sets; some care is necessary when dealing with type-definable objects. For instance, the intersection of the definable subgroups $n\mathbb{Z}$ of $\mathbb{Z}$ (for $n < \omega$) is the identity $\{1\}$. However, this is a particularity of the model $\mathbb{Z}$: in every group G elementarily equivalent to $\mathbb{Z}$, the subgroup given by $\bigcap_{n<\omega} nG$ (i.e. type-defined by the same formulas) has index at most the size of the continuum in G.

An important model-theoretic (and algebraic) construction is the *ultraproduct*. Let I be an index set, and $\mathfrak{M}_i$ an $\mathcal{L}$-structure for every $i \in I$. An *ultrafilter* $\mathcal{U}$ on I is a maximal collection of subsets of I closed under finite intersections and supersets and not containing $\emptyset$; for every subset $J \subseteq I$ either $J \in \mathcal{U}$ or $I - J \in \mathcal{U}$. The ultraproduct $\prod_I \mathfrak{M}_i/\mathcal{U}$ is the $\mathcal{L}$-structure with domain $\prod_I \mathfrak{M}_i/\sim$, where $\sim$ is the equivalence relation given by

$$(m_i\colon i \in I) \sim (n_i\colon i \in I) \quad \text{iff} \quad \{i \in I\colon m_i = n_i\} \in \mathcal{U}.$$

Put $(m_i\colon i \in I)/\sim\; = [m_i\colon i \in I]$. Functions on $\prod_I \mathfrak{M}_i/\mathcal{U}$ are defined via

$$f\big([m_i^1\colon i \in I], \ldots, [m_i^n\colon i \in I]\big) = \big[f\big(m_i^1, \ldots, m_i^n\big)\colon i \in I\big],$$

and relations by

$$\big([m_i^1\colon i \in I], \ldots, [m_i^n\colon i \in I]\big) \in R \Leftrightarrow \big\{i \in I\colon \big(m_i^1, \ldots, m_i^n\big) \in R^{\mathfrak{M}_i}\big\} \in \mathcal{U}.$$

This turns $\prod_I \mathfrak{M}_i/\mathcal{U}$ into an $\mathcal{L}$-structure; Łos' Theorem states that

$$\prod_i \mathfrak{M}_i/\mathcal{U} \models \varphi[\overline{m}_i\colon i \in I]$$

for any formula $\varphi(\bar{x})$ if and only if $\{i \in I\colon \mathfrak{M}_i \models \varphi(\overline{m}_i)\} \in \mathcal{U}$. In particular, if all $\mathfrak{M}_i$ are models of some theory T, so is their ultraproduct. If all $\mathfrak{M}_i$ are equal to $\mathfrak{M}$, we call $\prod_I \mathfrak{M}/\mathcal{U}$ an *ultrapower* of $\mathfrak{M}$; in this case the diagonal map induces an elementary embedding of $\mathfrak{M}$ into $\prod_I \mathfrak{M}/\mathcal{U}$. For instance, if every $\mathfrak{M}_i$ is a finite group (field), then $\prod_I \mathfrak{M}_i/\mathcal{U}$ is a *pseudofinite* group (field); every sentence true in a pseudofinite structure is true in infinitely many finite structures. Łos' Theorem has an important consequence, the

COMPACTNESS THEOREM. *Let Σ be a set of sentences such that any finite subset of Σ has a model. Then Σ has a model.*

PROOF. For any finite subset i of Σ consider a model $\mathfrak{M}_i$ of i. If I is the collection of all finite subsets of Σ and $\mathcal{U}$ is an ultrafilter on I containing all sets of the form $\{j \in I\colon j \supseteq i\}$ for $i \in I$, then $\prod_I \mathfrak{M}_i/\mathcal{U} \models \Sigma$. □

If $\overline{m}$ is a tuple in $\mathfrak{M}$ and A is a subset of $\mathfrak{M}$, the *type* of $\overline{m}$ over A, denoted $\text{tp}(\overline{m}/A)$, is the set of all formulas with parameters from A satisfied by $\overline{m}$ in $\mathfrak{M}$ (i.e. the smallest type-definable set over A containing $\overline{m}$). More generally, a partial n-type over $A \subseteq \mathfrak{M}$ is a consistent collection of formulas with parameters in A and free variables $x_1, \ldots, x_n$; it is *complete* if it is maximal. We denote the collection of complete n-types over A by $S_n(A)$, and put $S(A) = \bigcup_{n<\omega} S_n(A)$. Clearly $\mathrm{tp}(\overline{m}/A)$ is a complete type for any $\overline{m} \in \mathfrak{M}$ and $A \subseteq \mathfrak{M}$; while not every type over $A \subseteq \mathfrak{M}$ must be realized in $\mathfrak{M}$ itself, it is always realized in some elementary extension of $\mathfrak{M}$ by compactness. If π is a partial type and the ambient model $\mathfrak{M}$ is clear, we write $\overline{m} \models \pi$ if $\mathfrak{M} \models \varphi(\overline{m})$ for all $\varphi(\bar{x}) \in \pi(\bar{x})$, and say that $\overline{m}$ *realizes* π (in $\mathfrak{M}$).

DEFINITION 2. A structure $\mathfrak{M}$ is κ*-saturated* if every type over fewer than κ parameters in $\mathfrak{M}$ is realized in $\mathfrak{M}$; it is *saturated* if it is $|\mathfrak{M}|$-saturated.

Saturated structures resemble compact structures in topology; in a saturated structure type-definable sets (given by small intersections) are well-behaved. It can be shown that every model of a stable theory embeds elementarily into a saturated structure.

We now turn towards model-theoretic stability. To simplify notation, we fix a big, saturated model $\mathfrak{C}$, the *monster model*, of the theory in question; all elements and parameter sets will come from $\mathfrak{C}$ and all models will be elementary submodels of $\mathfrak{C}$, of smaller cardinality. (In the case of a stable group, we may write $\mathfrak{G}$ instead of $\mathfrak{C}$.) As every model of the theory of cardinality less than $|\mathfrak{C}|$ embeds elementarily into $\mathfrak{C}$, this is no restriction. We shall just write $\models \varphi$ for $\mathfrak{C} \models \varphi$.

DEFINITION 3. A sequence $(\overline{m}_i \colon i \in I)$ (where I is a totally ordered index set) is *indiscernible* over a set A if $\mathrm{tp}(\overline{m}_{i_1}, \ldots, \overline{m}_{i_k}/A)$ depends only on the order type of $(i_1, \ldots, i_k)$; the sequence is an *indiscernible set* if $\mathrm{tp}(\overline{m}_{i_1}, \ldots, \overline{m}_{i_k}/A)$ depends only on k. A formula $\varphi(\bar{x}, \overline{m})$ *divides* over A if there is an indiscernible sequence $(\overline{m}_i \colon i < \omega)$ over A with $\overline{m} = \overline{m}_0$ and $\bigwedge_{i<\omega} \varphi(\bar{x}, \overline{m})$ inconsistent. A partial type π *forks* over A if it implies a finite disjunction of formulas, each of which divides over A. We say that $\bar{a}$ *is independent of* B *over* A, and write $\bar{a} \mathop{\smile\!\!\!\!\!|}_A B$, if $\mathrm{tp}(\bar{a}/AB)$ does not fork over A, and $C \mathop{\smile\!\!\!\!\!|}_A B$ if $\bar{c} \mathop{\smile\!\!\!\!\!|}_A B$ for every finite tuple $\bar{c} \in C$.

A formula $\varphi(\bar{x}, \bar{y})$ has the *order property* if there is a sequence $(\bar{a}_i, \bar{b}_i \colon i < \omega)$ such that $\models \varphi(\bar{a}_i, \bar{b}_j)$ iff $i < j$.

A partial type $\pi(\bar{x})$ over A is $\varphi(\bar{x}, \bar{y})$*-definable* if there is a formula with free variables $\bar{y}$, usually denoted $d_\pi \bar{x}\, \varphi(\bar{x}, \bar{y})$, such that $\models d_\pi \bar{x}\, \varphi(\bar{x}, \bar{a})$ iff $\varphi(\bar{x}, \bar{a}) \in \pi$, for any $\bar{a} \in A$. A type is *definable* if it is φ-definable for all φ.

A complete theory T is λ*-stable* if $|S(\mathfrak{M})| = \lambda$ for every model $\mathfrak{M}$ of cardinality λ; it is *stable* if it is stable for some λ. Finally, T is *superstable* if it is λ-stable for all sufficiently large λ.

It is easy to see that two φ-definitions $d_p\bar{x}\, \varphi(\bar{x}, \bar{y})$ and $d'_p\bar{x}\, \varphi(\bar{x}, \bar{y})$ for a type p over a model $\mathfrak{M}$ are equivalent, as $\mathfrak{M} \models \forall \bar{y}\, [d_p\bar{x}\, \varphi(\bar{x}, \bar{y}) \leftrightarrow d'_p\bar{x}\, \varphi(\bar{x}, \bar{y})]$. This need not be the case, however, for types over arbitrary sets.

THEOREM 2.1. *The following are equivalent:*

(i) $\mathrm{Th}(\mathfrak{C})$ *is stable.*

(ii) *Every complete type (over every parameter set) is definable.*
(iii) *Every infinite indiscernible sequence is an indiscernible set.*
(iv) *No formula has the order property.*

A partial type π in a stable theory forks over A if and only if it implies some formula which divides over A. A stable theory T is λ-stable for every λ with $\lambda^{|T|} = \lambda$. If T is superstable, it is λ-stable for all $\lambda \geqslant 2^{|T|}$. If T is countable and ω-stable, it is stable for all infinite λ (and in particular superstable).

If T is stable, then $\perp$ satisfies:

(1) Existence: *If $p \in S(A)$ and $A \subseteq B$, then there is an extension $q \in S(B)$ of p which does not fork over A.*
(2) Symmetry: *$A \perp_B C$ if and only if $C \perp_B A$.*
(3) Transitivity: *$A \perp_B C$ and $A \perp_{BC} D$ if and only if $A \perp_B CD$.*
(4) Finite Character: *$A \not\perp_B C$ if and only if there are finite $\bar{a} \in A$ and $\bar{c} \in C$ with $\bar{a} \not\perp_B \bar{c}$.*
(5) Local Character: *For every $p \in S(A)$ there is $A_0 \subseteq A$ with $|A_0| \leqslant |T|$, such that p does not fork over A_0.*
(6) Boundedness: *If $p \in S(A)$ and $A \subseteq B$, then p has at most $2^{|T|}$ non-forking extensions to B.*

In fact, if $\mathfrak{M}$ is a model of a stable theory, $p \in S(\mathfrak{M})$ and $A \supseteq \mathfrak{M}$, then p has a unique non-forking extension $q \in S(A)$; it is characterized by $\varphi(\bar{x}, \bar{a}) \in q$ iff $\models d_p\bar{x}\, \varphi(\bar{x}, \bar{a})$ for any $\varphi \in \mathcal{L}$ and $\bar{a} \in A$. So q has the same φ-definition as p.

T is superstable if and only if it is stable and for every $p \in S(A)$ there is a finite $A_0 \subseteq A$ such that p does not fork over A_0.

By the existence axiom, given any type $p \in S(A)$ we can find an indiscernible sequence $(\bar{a}_i\colon i < \omega)$ of realizations of p such that $\bar{a}_i \perp_A \{\bar{a}_j\colon j < i\}$. Such a sequence is called a *Morley sequence* for p; as it forms an indiscernible set, we have $\bar{a}_i \perp_A \{\bar{a}_j\colon j \neq i\}$. All Morley sequences for p are isomorphic over A.

Having defined independence, we now aim for dimensions. For convenience, denote by On^+ the collection of ordinals together with an additional symbol ∞ which is greater than all ordinals.

DEFINITION 4. Let $\varphi(\bar{x}, \bar{y})$ be a formula. $D(., \varphi, \omega)$ is the smallest function from the collection of all partial types in $\bar{x}$ to ω satisfying: if $\pi(\bar{x})$ is a partial type over A, then $D(\pi(\bar{x}), \varphi, \omega) \geqslant n + 1$ if there is an instance $\varphi(\bar{x}, \bar{a})$ which forks over A, such that $D(\pi(\bar{x}) \cup \{\varphi(\bar{x}, \bar{a})\}, \varphi, \omega) \geqslant n$.

The *Shelah rank* D is the smallest function from the collection of all partial types to On^+ satisfying for all ordinals α: if $\pi(\bar{x})$ is a partial type over A, then $D(\pi) \geqslant \alpha + 1$ if there is a formula $\varphi(\bar{x})$ which forks over A, such that $D(\pi \cup \{\varphi\}) \geqslant \alpha$.

The *Morley rank* RM is the smallest function from the collection of all partial types to On^+ which satisfies for all ordinals α: if $\pi(\bar{x})$ is a partial type over A, then $RM(\pi) \geqslant \alpha + 1$ if there is $B \supseteq A$ and pairwise contradictory extensions q_i for $i < \omega$ of π to B, such that $RM(q_i) \geqslant \alpha$ for all $i < \omega$.

The *Lascar rank* U is the smallest function from the collection of all complete types to On^+ satisfying for all ordinals α: $U(p) \geqslant \alpha + 1$ if p has a forking extension q with $U(q) \geqslant \alpha$.

If R is any rank, we shall abbreviate $R(\mathrm{tp}(a/B))$ as $R(a/B)$. It is clear from the definition that the Lascar, Shelah and Morley ranks are invariant under definable maps with finite fibres. If $RM(\pi) < \infty$, we define the *Morley degree* $DM(\pi)$ to be the maximal number of pairwise inconsistent extensions of π of Morley rank $RM(\pi)$; this is finite. It is easy to see that a set is strongly minimal if and only if it has Morley rank and degree 1. More generally, an almost strongly minimal set, or an ω_1-categorical countable theory, have finite Morley rank.

THEOREM 2.2. *If T is stable, then $D(\pi, \varphi, \omega)$ is finite for all π and φ. Moreover, if $A \subseteq B$ and $q \in S(B)$ extends $p \in S(A)$, then q does not fork over A iff $D(q, \varphi, \omega) = D(p, \varphi, \omega)$ for all φ.*

For any type p we have $U(p) \leqslant D(p) \leqslant RM(p)$. Furthermore $D(p) = \min\{D(\varphi)\colon \varphi \in p\}$ and $RM(p) = \min\{RM(\varphi)\colon \varphi \in p\}$; conversely $D(\varphi) = \max\{D(p)\colon \varphi \in p\}$ and $RM(\varphi) = \max\{RM(p)\colon \varphi \in p\}$, and both maxima are attained.

T is superstable iff $U(p) < \infty$ for all types p, or equivalently iff $D(p) < \infty$ for all types p. A countable theory is ω-stable iff $RM(p) < \infty$ for all types p.

If q extends p and $U(q) < \infty$, then q is a nonforking extension of p if and only if $U(p) = U(q)$. An analogous assertion holds for Shelah and Morley ranks.

Recall that every ordinal can be written in the (Cantor normal) form $\sum_{i=0}^{k} \omega^{\alpha_i} n_i$ for some $k < \omega$, with $\alpha_0 > \cdots > \alpha_k$ and $n_i < \omega$ for all $i \leqslant k$. The commutative sum of two ordinals $\alpha = \sum_i \omega^{\alpha_i} n_i$ and $\alpha' = \sum_i \omega^{\alpha_i} n_i'$ is then defined as $\alpha \oplus \alpha' = \sum_i \omega^{\alpha_i}(n_i + n_i')$.

Lascar rank is well-behaved with respect to fibres:

THEOREM 2.3 (Lascar Inequalities).

(1) $U(a/bA) + U(b/A) \leqslant U(ab/A) \leqslant U(a/bA) \oplus U(b/A)$.

(2) *If a and b are independent over A, then $U(ab/A) = U(a/A) \oplus U(b/A)$.*

(3) *Suppose $U(a/Ab) < \infty$ and $U(a/A) \geqslant U(a/Ab) + \omega^\alpha \cdot n$. Then $U(b/A) \geqslant U(b/Aa) + \omega^\alpha \cdot n$.*

The following construction will allow us to talk about quotient structures.

DEFINITION 5. Let $\mathfrak{M}$ be a model of a theory T. For every $\emptyset$-definable equivalence relation $E(\bar{x}, \bar{y})$ we add a new unary predicate $P_E(x)$ to our language, together with a new function π_E. This will form the language $\mathcal{L}^{\mathrm{eq}}$; the model $\mathfrak{M}^{\mathrm{eq}}$ is the disjoint union of domains $\mathfrak{M}_E$ consisting of all the equivalence classes of tuples in $\mathfrak{M}$ modulo E, for all $\emptyset$-definable E. We shall identify the original model $\mathfrak{M}$ with the set of classes modulo equality, and call it the *home sort*; the original constants, functions and relations will live there. All other elements will be *imaginary sorts*. Finally, we shall interpret P_E as $\mathfrak{M}_E$, and π_E as the projection from a tuple in the home sort to its equivalence class modulo E. The theory T^{eq} is then the $\mathcal{L}^{\mathrm{eq}}$-theory of $\mathfrak{M}^{\mathrm{eq}}$.

It can be checked that T^{eq} does not depend on the choice of our original model $\mathfrak{M}$. In general, a model of T^{eq} will have elements satisfying no predicate P_E; these elements form the *superfluous sort* and carry no structure at all. In particular, the structure obtained by simply omitting superfluous elements will be an elementary substructure. Finally, any model of T can be uniquely expanded to a model of T^{eq} without superfluous elements. It follows that T is λ-stable iff T^{eq} is λ-stable.

Particular sorts of T^{eq} are the *permutation sorts* (classes of $\mathfrak{M}^j$ modulo E, where $(m_1, \ldots, n_j)$ and $(n_1, \ldots, n_j)$ are equivalent modulo E iff one tuple is a permutation of the other), and *tuple sorts* (classes modulo E, where $(m_1, \ldots, m_j)E(n_1, \ldots, n_j)$ holds iff the tuples are the same). If $E_{\overline{m}}(x, y)$ is an $\overline{m}$-definable equivalence relation, we may consider the $\emptyset$-definable equivalence relation $(\bar{x}, \bar{s})E(\bar{y}, \bar{t})$ iff $\bar{s} = \bar{t}$, and either $E_{\bar{s}}$ is an equivalence relation and $E_{\bar{s}}(\bar{x}, \bar{y})$, or else $E_{\bar{s}}$ is not an equivalence relation and $\bar{x} = \bar{y}$. Then the classes of $E_{\overline{m}}$ are just the same as the classes of E which come from pairs with second co-ordinate $\overline{m}$. For example, if T is a theory of groups and H is a subgroup defined by some formula $\varphi(x, \overline{m})$, the (right) cosets of H are imaginary elements, namely classes of the equivalence relation $E(x, \bar{s}; y, \bar{t})$ given by

$$\bar{s} = \bar{t} \wedge \left[x = y \vee \left[\varphi\left(xy^{-1}, \bar{s}\right) \wedge \forall u \forall v \left(\left(\varphi(u, \bar{s}) \wedge \varphi(v, \bar{s})\right) \rightarrow \varphi\left(uv^{-1}, \bar{s}\right)\right)\right]\right]$$

with second co-ordinate $\overline{m}$; left cosets are dealt with similarly.

Recall from the introduction that $\mathrm{acl}(A)$ is the set of elements lying in a finite A-definable set; we take this to include imaginary elements. We also define the *definable closure* $\mathrm{dcl}(A)$ of a subset A of $\mathfrak{M}^{\mathrm{eq}}$ as the collection of all elements of $\mathfrak{M}^{\mathrm{eq}}$ which lie in a 1-element A-definable subset of $\mathfrak{M}^{\mathrm{eq}}$.

DEFINITION 6. Let X be a definable set. The *canonical base* $\mathrm{Cb}(X)$ for X is an element $m \in \mathfrak{M}^{\mathrm{eq}}$ such that any automorphism of $\mathfrak{C}$ stabilizes X setwise if and only if it fixes m.

Let p be a type over a model. The *canonical base* $\mathrm{Cb}(p)$ is the smallest definably closed subset of the domain of p over which p does not fork.

These canonical bases exist and are unique (up to definable closure). If X is defined by the formula $\varphi(\bar{x}, \overline{m})$, then $\forall \bar{x}\, [\varphi(\bar{x}, \bar{y}) \leftrightarrow \varphi(\bar{x}, \bar{z})]$ defines an equivalence relation E, and $\mathrm{Cb}(X)$ is the class of $\overline{m}$ modulo E. If $d_p\bar{x}\, \varphi(\bar{x}, \bar{y})$ is the φ-definition of p, where p is a type over a model, then $\mathrm{Cb}(p) = \mathrm{dcl}(\mathrm{Cb}(d_p\bar{x}\, \varphi(\bar{x}, \bar{y}))\colon \varphi \in \mathcal{L})$.

3. Chain conditions

One of the first group-theoretic consequence of stability are chain conditions on definable and type-definable subgroups. In this section we shall first derive the chain conditions implied by the various forms of stability, and then look at the group theoretic consequences of those chain conditions, without further use of the stability assumption, as they may also hold in a wider context, e.g. groups definable in the field of real numbers (or any *o*-minimal field), or $\mathfrak{M}_c$-groups (satisfying the the chain condition on centralizers). In fact, many results which had first been proven in the stable context have subsequently been generalized to the context of groups with certain chain conditions.

DEFINITION 7. A subgroup H of a group G is $\bigwedge$*-definable* if it is the intersection of definable subgroups of G. A subgroup K of H is *definable relative to* H, or *relatively definable*, if it is the intersection of H with a definable subgroup of G.

A family $\mathfrak{H}$ of subgroups of G is *uniformly definable* if there is a single formula $\varphi(x, \bar{y})$ such that all $H \in \mathfrak{H}$ are definable by an instance $\varphi(x, \overline{m})$, for some $\overline{m}$.

So $\bigwedge$-definable subgroups are type-definable in a particular way.

DEFINITION 8. A group G satisfies the *icc* (uniform chain condition on intersections of uniformly definable subgroups) if for any formula $\varphi(x, \bar{y})$ there is some $n_\varphi < \omega$ such that any chain of intersections of subgroups defined by instances of $\varphi(x, \bar{y})$ of G has length at most n_φ.

PROPOSITION 3.1 [11]. *A stable group G satisfies the icc. In particular, G is an $\mathfrak{M}_c$-group, and centralizers of arbitrary sets are definable.*

PROOF. If $(\varphi(x, \overline{m}_i)\colon i < \omega)$ is a chain of uniformly definable subgroups, then

$$\forall x\, [\varphi(x, \bar{x}_1) \to \varphi(x, \bar{x}_2)]$$

orders the set $\{\overline{m}_i\colon i < \omega\}$, contradicting stability. Hence chains of uniformly definable subgroups are bounded. Now show that there is some $n < \omega$ such that any intersection of finitely many groups in a uniformly definable family is in fact a subintersection of size n, using the stability assumption. $\square$

COROLLARY 3.2. *If G is a stable group, then there is a finite bound n on the number of pairwise commuting non-Abelian normal subgroups of G. In particular, there are only finitely many minimal, or $\bigwedge$-definable minimal, normal non-Abelian subgroups.*

PROOF. If $\{N_i\colon i \in I\}$ is a family of pairwise commuting non-Abelian subgroups and $n_i \in N_i - Z(N_i)$ for $i \in I$, then $\bigcap_{i \in I} C_G(n_i)$ does not equal a proper subintersection. So $|I|$ must be bounded by some finite n, but any two distinct minimal or $\bigwedge$-definable minimal normal subgroups must commute. $\square$

DEFINITION 9. Let G be a group. Then G has the
- κcc if any chain of $\bigwedge$-definable subgroups has cardinality less than κ.
- ωdcc if any descending chain of definable subgroups has finite length.
- ωdcc^0 if any descending chain of definable subgroups, each of infinite index in its predecessor, has finite length.
- ωacc^0 if any ascending chain of definable subgroups, each of infinite index over its predecessor, has finite length.

LEMMA 3.3. *Let G be a type-definable group with $U(G) < \infty$, and H a relatively definable subgroup of G. Then $U(H) = U(G)$ if and only if H has finite index in G. An analogous assertion holds for D and RM. Moreover, if $RM(G) = RM(H) < \infty$, then $DM(G) = DM(H) \cdot |G : H|$.*

PROOF. The formula $gx \in H$ (i.e. the coset $g^{-1}H$) forks over $\emptyset$ if and only if $g \notin \mathrm{acl}(\emptyset)$, this proves the first assertion (for all three ranks). For the "moreover", note that translation by a coset representative defines a bijection between the types in H of Morley rank $RM(G)$ and the types in the coset of Morley rank $RM(G)$. □

PROPOSITION 3.4. *A stable group satisfies the $|T|^+cc$. A superstable group satisfies the ωdcc^0. A superstable group of finite U-rank satisfies the ωacc^0. An ω-stable group satisfies the ωdcc.*

PROOF. If $(H_i\colon i \in I)$ is a chain of $\bigwedge$-definable subgroups of G, let H_i^φ be the intersection of all supergroups of H_i which are defined by some instance of φ. Then $H_i = \bigcap_{\varphi\in\mathcal{L}} H_i^\varphi$. The H_i^φ are uniformly definable by the icc; as chains of the H_i^φ must be finite, $|I| < |T|^+$. The other assertions follow from Lemma 3.3. □

The chain conditions are clearly inherited by definable sections.

DEFINITION 10. Let G be a group with the icc, and $H \leqslant G$ a $\bigwedge$-definable subgroup. The *connected component* H^0 is the intersection of all relatively definable subgroups of finite index in H. We say that H is *connected* if $H = H^0$.

This definition seeks to mimic the connected component of an algebraic group. By the icc, the connected component is $\bigwedge$-definable. Note that $|H : H^0| \leqslant 2^{|T|}$, and $H^0 \trianglelefteq H$. A connected component is connected, as is any of its quotients by a $\bigwedge$-definable normal subgroup; conversely, if H/N and N are connected, so is H.

LEMMA 3.5. *The additive group of an infinite division ring D with icc is connected.*

PROOF. If H is a relatively definable additive subgroup of finite index in D, so is any translate dH, for $d \in D^\times$. By the icc, $\bigcap_{d\in D^\times} dH$ is a relatively definable ideal of finite index, i.e. D itself. □

(In the next section we shall also prove the multiplicative connectivity of a stable division ring. This, however, goes beyond chain conditions; for instance the multiplicative group of the non-zero real numbers has a definable subgroup of index 2, namely the positive reals.)

PROPOSITION 3.6. *A finite G-invariant subset A of a connected group G is central. A connected group with finitely many commutators is Abelian.*

PROOF. If $a \in A$, then $|G : C_G(a)| \leqslant |A|$, so $a \in Z(G)$ by connectivity. If the set of commutators of G is finite, so are $g^G g^{-1}$ and g^G for any $g \in G$, whence $g \in Z(G)$. □

Clearly, for a definable subgroup H of a group G the normalizer $N_G(H)$ and the groups in the ascending central series $Z_i(H)$ are definable. If H is only $\bigwedge$-definable, we need the icc:

LEMMA 3.7. *If G is a group with the icc and $H \leqslant G$ a $\bigwedge$-definable subgroup, then $N_G(H)$ is $\bigwedge$-definable, and $Z_i(H)$ is relatively definable for all $i < \omega$.*

PROOF. If $H = \bigwedge_i H_i$, let $\overline{H}_i$ be the intersection of all G-conjugates of H_i which contain H. Then $\overline{H}_i$ is definable, and $N_G(H) = \bigcap_i N_G(\overline{H}_i)$. The second assertion follows from the definability of centralizers. □

In general, neither the groups in the derived nor in the descending central series will be definable. For any group G and $n < \omega$, the subsets of elements of order dividing n is (relatively) definable; the subset of n-elements (of order dividing some power of n) is (relatively) definable iff there is some bound on the power needed.

COROLLARY 3.8. *Let G be a group with the icc, and $H \leqslant G$ be $\bigwedge$-definable. If H is soluble (nilpotent), then there is a finite series of relatively definable normal subgroups of H with Abelian (central) quotients, of length the derived length (nilpotency class) of H.*

PROOF. By induction on the derived length (nilpotency class), using Lemma 3.7. □

We shall now treat nilpotency in groups with the chain condition on centralizers ($\mathfrak{M}_c$). As Bryant [44] has constructed an $\mathfrak{M}_c$-group G such that $G/Z(G)$ is not $\mathfrak{M}_c$, this property is not preserved under quotients. Recall that the *Fitting subgroup* $F(G)$ of a group G is the subgroup generated by all normal nilpotent subgroups, and the *Hirsch–Plotkin radical* $HP(G)$ is the subgroup generated by all normal locally nilpotent subgroups. These are characteristic subgroups, and $HP(G)$ is locally nilpotent.

LEMMA 3.9. *A locally nilpotent $\mathfrak{M}_c$-group G is soluble, and the derived length is bounded by the maximal length k of a chain of centralizers (if there is such k).*

PROOF. It follows from local nilpotency and $\mathfrak{M}_c$ that unless G is Abelian, $Z_2(G) > Z(G)$. Then $G > C_G(g) \geqslant G'$ for any $g \in Z_2(G) - Z(G)$; the assertion follows inductively. □

If G acts on a group A and $X \subseteq G$, we put $C_A(X) = \{a \in A\colon \forall x \in X\ a^x = a\}$.

LEMMA 3.10. *Let G be an Abelian group acting on an Abelian group A. Suppose that there are finitely many elements $g_0, \ldots, g_k$ in G such that $C_A(G) = C_A(g_0, \ldots, g_k)$. Let a be an element in A, and suppose that for all $i = 0, \ldots, k$ there is some non-zero $m_i < \omega$ with $(g_i - 1)^{m_i} a = 0$. Then $a \in C_A^m(G)$, where $m = 1 + \sum_{i=0}^k (m_i - 1)$.*

PROOF. Consider the action of $\mathbb{Z}G$ on A. Then $(g_i - 1)^{m_i} a = 0$ for all $i \leqslant k$; by commutativity $(g_{i_1} - 1) \cdots (g_{i_m} - 1)a = 0$ for all $i_1, \ldots, i_m \leqslant k$. Hence all of $g_0, \ldots, g_k$ annihilates $(g_{i_2} - 1) \cdots (g_{i_m} - 1)a$, and so does G. Thus $(g_1' - 1)(g_{i_2} - 1) \cdots (g_{i_m} - 1)a = 0$ for any $g' \in G$; inductively we see $(g_1' - 1)(g_2' - 1) \cdots (g_m' - 1)a = 0$ for any $g_1', \ldots, g_m' \in G$. □

DEFINITION 11. A *(left normalized) commutator condition* on $x_0, \ldots, x_k$ is a word $w(x_0, \ldots, x_k)$ of the form $[x_0, x_{i_1}, \ldots, x_{i_n}]$, where $i_0, \ldots, i_n \in \{0, \ldots, k\}$ with $i_1 \neq 0$. A group G satisfies a commutator condition w if $w(\bar{g}) = 1$ for all tuples $\bar{g} \in G$.

A particular commutator condition is the *n-th Engel identity* $[x,_n y] = 1$ (where $[x,_0 y] = x$ and $[x,_{n+1} y] = [[x,_n y], y]$). An element $g \in G$ is *right Engel* if $[g,_n x] = 1$ for all x in G and some $n < \omega$; it is *left Engel* if $[x,_n g] = 1$ for all x in G, and *Engel* if it is right or left Engel. If n is bounded independently of x, we call g *bounded* Engel, or n-Engel. Engel elements in soluble groups have been studied by Gruenberg [74].

PROPOSITION 3.11. *Let G be a soluble $\mathfrak{M}_c$-group such that for any $g \in G$ there is a commutator condition $w(x, y)$ such that $w(gh, g) = 1$ for all $h \in G'$. Then G is nilpotent.*

PROOF. By induction on the derived length, using Lemma 3.10. □

COROLLARY 3.12. *A soluble $\mathfrak{M}_c$-group satisfying some commutator condition is nilpotent. In particular, a uniformly locally nilpotent $\mathfrak{M}_c$-group G is nilpotent.*

PROOF. A commutator condition $w(x, y_1, \ldots, y_k)$ gives rise to one in two variables, namely $w'(x, y) = w(x, y, \ldots, y)$; now use Proposition 3.11. For the second assertion, note that the group is soluble by Lemma 3.9. □

THEOREM 3.13 [63]. *The Fitting subgroup of an $\mathfrak{M}_c$-group G is nilpotent.*

PROOF. $F(G)$ is locally nilpotent, and soluble by Lemma 3.9. If $f \in F(G)$, it lies in a normal nilpotent subgroup K of class c, say. So $[h, f] \in K$ for any $h \in F(G)$, and $[fh,_{c+1} f] = [h,_{c+1} f] = 1$. Hence $F(G)$ is nilpotent by Proposition 3.11. □

THEOREM 3.14. *$HP(G)$ is nilpotent for any ω-saturated $\mathfrak{M}_c$-group G.*

PROOF. $HP(G)$ is soluble by Lemma 3.9; if $HP(G) > F(G)$, then there is a G-invariant subgroup $F \leqslant HP(G)$ with $F/F(G)$ non-trivial Abelian. If $F(G)$ has nilpotency class c, then $F(G)$ is definable as $\{g \in G\colon \forall x_1, \ldots, x_c\, [g, g^{x_1}, \ldots, g^{x_c}] = 1\}$, so by ω-saturation for any $f \in F$ there is $n < \omega$ such that $[f,_n g] = 1$ for all $g \in F(G)$. Hence F is nilpotent by Proposition 3.11, a contradiction. □

A more involved application [192] of Lemma 3.10 yields:

THEOREM 3.15. *The set of bounded left Engel elements in an $\mathfrak{M}_c$-group equals the Fitting subgroup.*

To characterize the unbounded left Engel elements, we need the icc:

THEOREM 3.16 [176]. *The set of unbounded left Engel elements in an icc group equals the Hirsch–Plotkin radical.*

PROOF. Suppose for a contradiction that the left Engel elements do not generate a locally nilpotent subgroup. Use the icc to find two distinct maximal locally nilpotent subgroups S and T generated by left Engel elements, such that the group I generated by the left

Engel elements in $S \cap T$ is maximal. Find two left Engel elements $x \in N_S(I) - I$ and $y \in N_T(I) - I$. Choose $k < \omega$ maximal with $u = [y,_k x] \notin S$, such k exists since x is left Engel. Then $x^{-u}x = [u, x] \in S$, that is $x^u \in S \cap S^u$. As $x^u \notin I$, it is a left Engel element in $S \cap S^u - I$, whence $u \in N_G(S) - S$ by maximality of I. But $\langle S, u\rangle$ is soluble and generated by left Engel elements, contradicting maximality of S.

Since every element in $HP(G)$ is left Engel, we are done. □

The *normalizer condition* states that $N_G(H) > H$ for all $H < G$. Recall that we can continue the ascending central series into the transfinite, by taking unions at the limit stages. A group G is *hypercentral* if $G = Z_\alpha(G)$ for some ordinal α; it is easy to see that a hypercentral group satisfies the normalizer condition. It follows from Bludov [32] that:

THEOREM 3.17. *If G is an $\mathfrak{M}_c$-group, the following are equivalent:*
(i) *G is locally nilpotent.*
(ii) *G is hypercentral.*
(iii) *G satisfies the normalizer condition.*

In the periodic case, Bryant [44] has shown:

THEOREM 3.18. *A locally nilpotent periodic $\mathfrak{M}_c$-group G is nilpotent-by-finite; if $G/Z_i(G)$ has finite exponent for some $i < \omega$, then G is nilpotent. For all d sufficiently large $F(G) = C_G(g^{d!}\colon g \in G)$, and this is also the maximal nilpotent subgroup of finite index.*

PROOF. G is soluble by Lemma 3.9; if $G/Z_i(G)$ has finite exponent, G is uniformly locally nilpotent and hence nilpotent by Corollary 3.12. Using periodicity, an induction shows that there are $n_i < \omega$ such that $g^{n_i} \in Z(G)$ for all $g \in Z_i(G)$.

For the general case, use induction on centralizer chains, together with the identity

$$\prod_{0 \leqslant i < j \leqslant p} (g_i - g_j) = \det(g_i^j) = 0$$

in a group of order p^2, for suitable p. □

A similar theorem to Theorem 3.18 holds for locally soluble groups [45]:

THEOREM 3.19. *A locally soluble periodic $\mathfrak{M}_c$-group is nilpotent-by-(Abelian of finite Prüfer rank)-by-finite. In particular, if the exponent of the group is finite, then the group is nilpotent-by-finite.*

Baudisch and Wilson have even shown [21]:

THEOREM 3.20. *Let G be a soluble icc group with a normal nilpotent subgroup N such that G/N is periodic. Then G is nilpotent-by-(Abelian of finite Prüfer rank)-by-finite.*

In fact, Kegel [101] proves (but this is considerably more difficult):

THEOREM 3.21. *A locally finite $\mathfrak{M}_c$-group of finite exponent is nilpotent-by-finite.*

In the ω-stable case, Baur, Cherlin and Macintyre [23] obtain virtual commutativity:

THEOREM 3.22. *An ω-categorical ω-stable group is Abelian-by-finite.*

PROOF. An ω-categorical stable group G is uniformly locally finite $\mathfrak{M}_c$, hence nilpotent-by-finite; we may assume inductively that the group is connected of nilpotency class 2. In a counter-example of minimal rank the Mal'cev correspondence [116] yields an infinite ring of definable endomorphisms of G' (modulo finite subgroups), which is a matrix ring over a definable field K. By uniform local finiteness, K is finite, as is R, a contradiction. □

We finish this section with a survey of the relevant Sylow theory. As we cannot use counting arguments, the proofs proceed by increasing the intersection of two hypothetical non-conjugate Sylow subgroups, using Sylow's Theorems for finite 2-generated subgroups [40,158,192,159].

THEOREM 3.23. *The maximal 2-subgroups of a periodic $\mathfrak{M}_c$-group are locally finite and conjugate. The same holds for the maximal p-subgroups for any prime p, provided any two p-elements generate a finite subgroup.*

In the non-periodic context, we have to exclude the infinite dihedral group (which has two conjugacy classes of Sylow 2-subgroups).

DEFINITION 12. A theory T is *small* if $S(\emptyset)$ is countable.

Clearly an ω-stable theory is small; $\mathbb{Z}$ is not small.

THEOREM 3.24. *The Sylow p-subgroups of an icc group G are locally finite and conjugate, provided*
(i) *$p = 2$ and G is small, or*
(ii) *G is soluble-by-finite and small.*

COROLLARY 3.25. *If G is an icc group and N a definable normal subgroup, then the Sylow p-subgroups of G/N are the images of the Sylow p-subgroups of G, provided*
(i) *$p = 2$ and G is periodic or small, or*
(ii) *G is soluble-by-finite and small.*

THEOREM 3.26. *If C is a G-invariant subset of an icc group G such that any two elements of C generate a finite p-group (of bounded exponent), then C generates a locally finite p-group (of bounded exponent).*

4. Generic types

The second important tool in the study of stable groups are generic types. From now on, let the ambient structure $\mathfrak{G}$ (or $\mathfrak{C}$) be stable.

DEFINITION 13. If G is a stable group and $A \subseteq G$, then $g \in \mathfrak{G}$ is *generic* over A if for all $h \in \mathfrak{G}$ with $h \mathop{\smile\!\!\!\!\!\!\mid}_A g$ we have $hg \mathop{\smile\!\!\!\!\!\!\mid} A, h$. A type $p \in S_1(A)$ is *generic* if it is realized by a generic element.

This defines *left generic*; there is an analogous notion on the right.

LEMMA 4.1.
(i) *If g is left generic over A and $h \in G \cap \operatorname{acl}(A)$, then hg is left generic over A.*
(ii) *If $g \mathop{\smile\!\!\!\!\!\!\mid}_A B$, then g is left generic over A if and only if it is left generic over AB.*
(iii) *If g is left generic over A, so is g^{-1}.*
(iv) *An element g is left generic over A if and only if it is right generic over A.*

PROOF. Unravel the definitions and use the properties of independence. □

A *stratified* formula is one of the form $\varphi(yx, \bar{y})$, where the type variable x is pre-multiplied by a parameter variable y. For $g \in G$ and $A \subseteq B$ we have

$$g \mathop{\smile\!\!\!\!\!\!\mid}_A B \quad \text{iff} \quad D(g/A, \varphi, \omega) = D(g/B, \varphi, \omega)$$

for all stratified formulas φ.

THEOREM 4.2 [150]. *A stable group has a generic type. More precisely:*
(i) *$p \in S_1(A)$ is generic iff $D(p, \varphi, \omega)$ is maximal possible for all stratified formulas φ.*
(ii) *if G is superstable, then $p \in S_1(G)$ is generic iff $U(p)$ is maximal possible, or equivalently iff $D(p)$ is maximal possible.*
(iii) *if G is ω-stable, then $p \in S_1(A)$ is generic iff $RM(p)$ is maximal possible.*

PROOF. A type with the properties in (i), (ii) or (iii) is clearly left generic. For the converse, take $h \in G$ of maximal rank (any of them) and independent of $g \models p$ over A; consider the rank of gh^{-1} and use that the rank is invariant under nonforking extensions and left translation (here we need stratification of the formula φ).

Existence follows from the definition for D and RM. In a stable theory, enumerate all stratified formulas as $(\varphi_i\colon i < \alpha)$, and choose p such that $(D(p, \varphi_i, \omega)\colon i < \alpha)$ is maximal in the lexicographic ordering; this exists by compactness. Then p is generic. □

REMARK 1.
(i) As for any $h \in G$ and generic $g \mathop{\smile\!\!\!\!\!\!\mid}_G h$ both g^{-1} and gh are generic, every element of G is the product of two generic elements.

(ii) Since a superstable group has types of maximal rank, $U(G)$ is well-defined as the Lascar rank of the generic types of G.

(iii) If G acts transitively on a set X, we may use the action to define generic elements of X; as $X \cong G/\mathrm{Fix}_G(x)$ for any $x \in X$, after naming x this amounts to studying the action of G on a (left) coset space G/H. If $g \in G$ is generic, then g is generic in the coset gH over the canonical parameter $\mathrm{Cb}(gH)$, and gH is generic in G/H. So Lascar's inequality becomes $U(H) + U(G/H) \leqslant U(G) \leqslant U(H) \oplus U(G/H)$.

We call a formula *generic* if it lies in some generic type.

LEMMA 4.3. *A formula $\varphi(x)$ is generic iff finitely many translates of $\varphi(x)$ cover G. For every formula $\varphi(x, \bar{y})$ there is some $n_\varphi < \omega$ such that if $\varphi(x, \bar{a})$ is generic, then n_φ (left or right) translates of it cover G.*

PROOF. If finitely many translates cover G, then a generic type must satisfy one of them and we can translate back. Conversely, if $g \models \varphi$ is generic and $p \in S_1(G)$, choose $h \models p$ independent of g over G. Then $gh^{-1} \perp_G h$ by genericity; as $\models \varphi((gh^{-1})h)$, we get $\models d_p x\, \varphi(gh^{-1}x)$ and $G \models \exists y\, d_p x\, \varphi(yx)$. So $\varphi(h'x) \in p$ for some $h' \in G$; compactness now implies that finitely many translates of φ cover G. Lastly, as $D(\varphi(x, \bar{a}), \psi, \omega) \geqslant n$ is a type-definable condition on $\bar{a}$ for all stratified formulas ψ and $n < \omega$, so is genericity of $\varphi(x, \bar{a})$; compactness then yields the bound n_φ. □

DEFINITION 14. The *φ-stabilizer* of $p \in S_1(G)$ is defined as

$$\mathrm{stab}(p, \varphi) := \{g \in G\colon\ \forall y\, \forall \bar{z}\, d_p x\, [\varphi(gyx, \bar{z}) \leftrightarrow \varphi(yx, \bar{z})]\};$$

the *stabilizer* as $\mathrm{stab}(p) = \bigcap_{\varphi \in \mathcal{L}} \mathrm{stab}(p, \varphi)$.

Clearly $\mathrm{stab}(p, \varphi)$ is a definable and $\mathrm{stab}(p)$ a type-definable subgroup of G. We defined *left* stabilizers; similarly there are right ones, which may be different. Note that $g \in \mathrm{stab}(p)$ iff there is $x \models p$ with $x \perp_G g$ and $gx \models p$; in this case

$$\begin{aligned} D(p, \varphi, \omega) &= D(x/G, \varphi, \omega) = D(x/G \cup \{g\}, \varphi, \omega) \\ &= D(gx/G \cup \{g\}, \varphi, \omega) \geqslant D(gx/G, \varphi, \omega) = D(p) \end{aligned}$$

for all stratified formulas φ. Hence equality holds, and $gx \perp_G g$ as well.

LEMMA 4.4. *A type $p \in S_1(G)$ is generic iff* $\mathrm{stab}(p) = G^0$.

PROOF. A generic type p does not fork over $\emptyset$, so there are at most $2^{|T|}$ generic types, and $\mathrm{stab}(p) \geqslant G^0$. As any type must specify its coset modulo any subgroup of finite index, we also get $\mathrm{stab}(p) \leqslant G^0$.

As $D(G^0, \varphi, \omega) = D(G, \varphi, \omega)$ for any stratified formula φ, there is a generic type $p_0 \in S_1(G)$ extending the partial type $x \in G^0$. If $\mathrm{stab}(p) = G^0$, $g \models p$ and $g_0 \models p_0$ with $g_0 \perp_G$

g, then $\mathrm{tp}(g/G, g_0) = \mathrm{tp}(g_0 g/G, g_0)$. Therefore $\mathrm{tp}(g/G) = \mathrm{tp}(g_0 g/G)$, but the latter is generic. □

In particular, $\mathrm{stab}(p, \varphi)$ has finite index in G for generic p, and there are only finitely many generic φ-types. G^0 has a unique generic type, the *principal* generic type of G, and G/G^0 acts transitively (by translation) on the set of generic types of G. Moreover, as

$$(G \times H)^0 = G^0 \times H^0,$$

the generic elements of $G \times H$ are exactly the pairs of a generic element in G and an independent generic element of H. If G is ω-stable, then $|G : G^0| = DM(G)$. Finally, if G is type-definable connected with principal generic type p, then $\mathrm{Cb}(H) = \mathrm{Cb}(p)$.

PROPOSITION 4.5. *A stable division ring D has a unique additive and multiplicative generic type.*

PROOF. D^+ is connected, so D has a unique additive generic type p. Hence $dp = p$ for any $d \in D^\times$, so p is also the unique multiplicative generic type. □

EXAMPLE 1.
- (i) $(\mathbb{Q}, +)$ is connected and has a unique non-algebraic type over any subset A, its generic type.
- (ii) $(\mathbb{Z}, +)$ has connected component $\bigcap n\mathbb{Z}$ (in a saturated model!); the principal generic type over any subset A is the unique non-algebraic divisible type.
- (iii) The generic type of an algebraically closed field over any subset A is the unique non-algebraic type over A.

REMARK 2. It is clear that generic types in a superstable group are mapped to generic types under maps with finite fibres. This also holds in the stable case for definable endomorphisms with finite kernel, but is not true for general maps. Consider $G = \mathbb{Z}^\omega$, an Abelian group with a definable subgroup $2G$ of infinite index. So $2G$ is a non-generic subset, but the map defined by

$$x \mapsto \begin{cases} 2x & \text{for } x \in G - 2G, \\ x/2 & \text{for } x \in 2G - 4G, \\ x & \text{for } x \in 4G \end{cases}$$

exchanges the generic set $G - 2G$ with the non-generic set $2G - 4G$.

REMARK 3. The above development works just as well if G is a type-definable group in a stable structure. In fact, if H is an arbitrary subgroup of G, for any definable set $X \subseteq G$, either finitely many translates of X or of $G - X$ by elements in H cover H; in the first case we call X generic. Call a type generic for H if it only contains generic formulas. Then generic types for H exist [172], and are exactly the generic types of the smallest type-definable supergroup of H. More generally, it can be shown that if H is an arbitrary subset of a stable structure $\mathfrak{C}$ on which a definable function defines a group law, then H embeds into a definable group [128].

THEOREM 4.6 [80]. *A type-definable group G in a stable structure $\mathfrak{C}$ is contained in a definable supergroup and is $\bigwedge$-definable; if $\mathfrak{C}$ is ω-stable, G is definable. A type-definable division ring D is $\bigwedge$-definable; if $\mathfrak{C}$ is superstable, D is definable.*

PROOF. By compactness there is a definable superset $X \supseteq G$ such that for all x, y and z in X we have $(x * y) * z = x * (y * z)$ and $x * 1 = x = 1 * x$. If G is type-defined as $\bigwedge_{i \in I} \varphi_i(x)$, where the φ_i are closed under finite conjunctions, put $\psi_i(x) = \bigwedge_j d_j y\, \varphi_i(x * y)$, where d_j runs through the finitely many different $\varphi_i(x * y)$-definitions for generic types of G. Then $\bigwedge_{i \in I} \psi_i(x)$ also type-defines G.

By compactness there is $i \in I$ such that $x * y \in X$ for all $x, y \models \psi_i$. Then $gx \models \psi_i$ for all $x \models \psi_i$ and $g \in G$; if $Y = \{x \models \psi_i\colon \forall z\,[\psi_i(z) \to \psi_i(xz)]\}$, then the set of invertible elements of Y is a definable supergroup of G. As ψ_i can be replaced by ψ_j for any $\psi_j \vdash \psi_i$, we get $\bigwedge$-definability of G; definability in the ω-stable case follows from the ωdcc.

If D is a type-definable division ring, we find a definable multiplicative supergroup M on which addition is defined (but may go outside), and where the distributive laws hold. There is a definable additive subgroup $A \subseteq M \cup \{0\}$ containing D; intersecting it with its $D^\times$-translates, we may assume $dA \leqslant A$ for all $d \in D$. Put $D_0 = \{m \in M \cup \{0\}\colon mA \leqslant A\}$, this is a definable division ring containing D; as above $\bigwedge$-definability follows.

Finally, if $D_0 > D_1 > \cdots$ is a sequence of definable division rings in a superstable theory, then the rank must go down at each step and the sequence is finite. □

We finish this section by stating a theorem of Hrushovski [80] which reconstructs a group from generically given data; it generalizes a theorem of Weil [194] on reconstructing an algebraic group from a dense open subset.

THEOREM 4.7. *Let π be a partial type over a model $\mathfrak{M}$, and $*$ a partial $\mathfrak{M}$-definable operation defined on pairs of independent realizations of types in π, such that there are only finitely many different φ-types for the elements of π, where $\varphi(x, a, b, c, d)$ is the formula $[a * (b * x) = c * (d * x)]$. Suppose $*$ satisfies*

(1) Generic Independence: *$a * b \models \pi$ independently of a and of b, for any independent realizations $a, b \models \pi$,*
(2) Generic Associativity: *$(a * b) * c = a * (b * c)$ for any three independent realizations $a, b, c \models \pi$, and*
(3) Surjectivity: *for any independent $a, b \models \pi$ there are $c, c' \models \pi$ independently of $\{a, b\}$, with $a = b * c$ and $a = c' * b$.*

Then there are a type-definable group G, and a definable isomorphism between π and the set of generic elements of G, such that generically $*$ is mapped to the group multiplication. G is unique up to definable isomorphism.

PROOF. Consider the relation $R(x_0, y_0; x_1, y_1)$ on π^2 defined by

$$\exists z_0, z_1 \models \pi\ \big[z_0 \mathop{\smile\!\!\!\!\!\mid}_{\mathfrak{M}} x_0, y_0, x_1, y_1 \wedge z_1 \mathop{\smile\!\!\!\!\!\mid}_{\mathfrak{M}} x_0, y_0, x_1, y_1 \wedge x_0 * z_0 = x_1 * z_1 \wedge y_0 * z_0 = y_1 * z_1\big].$$

This is a definable equivalence relation; denote $(x, y)/R$ by $[x, y]$. One shows that the multiplication defined by $[x, y] \circ [y, z] = [x, y]$ is total and turns π^2/R into a type-definable group G. The class $[x * y, y]$ for $y \models \pi$ with $y \mathop{\smile\!\!\!\!\!\!\mid}_{\mathfrak{M}} x$ only depends on x; the map $x \mapsto [x * y, y]$ is the required embedding of π into G. $\square$

5. Superstable groups

As the Lascar inequalities show, the existence of a rank places considerable restrictions on possible interactions between definable sets.

LEMMA 5.1. *If G is a stable group and $p \in S_1(G)$ such that $a^{-1}b \mathop{\smile\!\!\!\!\!\!\mid}_G a$ for independent $a, b \models p$, then a left translate of p is a generic type of* $\mathrm{stab}(p^{-1})$.

PROOF. As $a^{-1}b \mathop{\smile\!\!\!\!\!\!\mid}_G b^{-1}$ as well, by definition $a^{-1}b \in \mathrm{stab}(p^{-1})$. Now

$$\begin{aligned} D\big(a^{-1}b/G, \varphi, \omega\big) &= D\big(a^{-1}b/G \cup \{a\}, \varphi, \omega\big) = D\big(b/G \cup \{a\}, \varphi, \omega\big) \\ &= D(b/G, \varphi, \omega) \end{aligned}$$

for all stratified formulas φ. If $x \models \mathrm{stab}(p^{-1})$ with $x \mathop{\smile\!\!\!\!\!\!\mid}_G b$, then $xb^{-1} \mathop{\smile\!\!\!\!\!\!\mid}_G x$ and

$$\begin{aligned} D(b/G, \varphi, \omega) &\geqslant D\big(b/G \cup \{xb^{-1}\}, \varphi, \omega\big) = D\big(x/G \cup \{xb^{-1}\}, \varphi, \omega\big) \\ &= D(x/G, \varphi, \omega). \end{aligned}$$

Hence $a^{-1}b$ is generic for $\mathrm{stab}(p^{-1})$, and b is generic for $a\,\mathrm{stab}(p^{-1})$. $\square$

It follows from Lascar's inequality that the conditions of Lemma 5.1 are satisfied if $U(p) = \sum_{i=0}^{k} \omega^{\alpha_i} n_i = \beta$ and $U(a^{-1}b/\mathfrak{M}) < \beta + \omega^{\alpha_k}$ for any two independent $a, b \models p$.

THEOREM 5.2 [30]. *If G is a superstable group with $U(G) = \sum_{i=0}^{k} \omega^{\alpha_i} n_i$, then G has a normal definable subgroup of Lascar rank $\sum_{i=0}^{l} \omega^{\alpha_i} n_i$ for all $l \leqslant k$.*

PROOF. For any $l \leqslant k$ there is $p \in S_1(G)$ with $U(p) = \sum_{i=0}^{l} \omega^{\alpha_i} n_i$; this p must satisfy the assumptions of Lemma 5.1 by the preceding remark. So $\mathrm{stab}(p^{-1})$ is a subgroup of the required Lascar rank. It follows from the Lascar inequality that $\mathrm{stab}(p^{-1})$ is commensurable with its conjugates; by connectivity it is normal, and contained in a definable normal subgroup of the same Lascar rank by the $\omega\mathrm{dcc}^0$. $\square$

COROLLARY 5.3. *A superstable simple group, or division ring, has Lascar rank $\omega^\alpha n$ for some ordinal α and some $n < \omega$.*

PROOF. The assertion is obvious for a simple group. If D is an infinite superstable division ring with $U(D) = \omega^\alpha n + \beta$ for some $\beta < \omega^\alpha$, Theorem 5.2 yields a definable additive subgroup H with $U(H) = \omega^\alpha n$. By the Lascar inequality H is commensurable with its translates dH for any $d \in D^\times$, and we may choose it invariant under translation. Then H is a non-trivial ideal and hence $H = D$. $\square$

THEOREM 5.4 [161,31]. *An infinite superstable group contains an infinite Abelian subgroup. More precisely, if $U(G) \geqslant \omega^\alpha$, then there is a relatively definable Abelian subgroup A with $U(A) \geqslant \omega^\alpha$.*

PROOF. If H is a minimal type-definable subgroup of G with $U(H) \geqslant \omega^\alpha$, then H is connected and has a unique generic type p, with $U(p) = \omega^\alpha n$ for some non-zero $n < \omega$. If H is non-Abelian, then $U(C_H(h)) < \omega^\alpha$ for $h \in H - Z(H)$, so

$$U(h^H) = U(H/C_H(h)) = \omega^\alpha n.$$

Hence the conjugacy class of h is generic; as the generic type is unique, all non-central elements are conjugate and $\overline{H} = H/Z(H)$ has a unique non-trivial conjugacy class.

As $\overline{H}$ is non-Abelian, its exponent is not 2. For $g \neq 1$ in $\overline{H}$ there is $h \in \overline{H}$ with $g^h = g^{-1}$, whence $g \in C_{\overline{H}}(h^2) - C_{\overline{H}}(h)$. If $h^2 = h^k$ for some $k \in \overline{H}$, we get

$$C_{\overline{H}}(h) < C_{\overline{H}}(h^k) < C_{\overline{H}}(h^{k^2}) < \cdots,$$

contradicting stability. □

It is unknown whether a stable group must have an infinite Abelian subgroup (although a small stable group always has a large Abelian subgroup [173]). One can also show that a superstable group must have infinitely many conjugacy classes [1].

PROPOSITION 5.5. *If G is connected with $U(G) = \sum_{i=0}^k \omega^{\alpha_i} n_i$, and σ is a definable endomorphism of G with $U(\ker(\sigma)) < \omega^{\alpha_k}$, then σ is surjective.*

PROOF. $U(\mathrm{im}(\sigma)) = U(G/\ker(\sigma))$; use the Lascar inequality. □

COROLLARY 5.6 [115,58]. *An infinite superstable ring without zero-divisors is a commutative algebraically closed field.*

PROOF. Consider first an infinite superstable field K. Then the maps $x \mapsto x^n$ and $x \mapsto x^p - x$ (in characteristic $p > 0$) with finite kernel are surjective by connectedness of K; in particular K is perfect. If K is not algebraically closed, it has a finite extension L which has a Galois extension L' of minimal possible degree; this must be either a Kummer or an Artin–Schreier extension. But that is impossible, since L is again superstable, and contains its roots and pseudo-roots.

If D is an infinite superstable division ring of Lascar rank $\omega^\alpha n$, it has an Abelian subfield K of Lascar rank at least ω^α. Then $[D:K]$ is finite, as is $[D:Z(D)]$, and $K/Z(D)$ is a finite extension, contradicting algebraic closedness of $Z(D)$.

Finally, if R is an infinite superstable ring without zero-divisors, the map $x \mapsto rx$ is surjective on $(R^+)^0$ for any $r \in R - \{0\}$ by Proposition 5.5. Hence there is a unit $1 \in (R^+)^0$, and $(R^+)^0 = R$, so R is a division ring. □

THEOREM 5.7 [83]. *A definable non-trivial automorphism σ of an infinite superstable field K has infinite order and finite fixed field K^σ. A family of uniformly definable automorphisms of K is finite.*

PROOF. If $U(K) = \omega^\alpha n$, then $U(K^\sigma) < \omega^\alpha$, since $[K : K^\sigma]$ must be infinite. So $\rho(x) = \sigma(x)/x$ is an endomorphism of $K^\times$ with kernel $(K^\sigma)^\times$, and is surjective by Proposition 5.5. If $\mathrm{char}(K) \neq 2$, choose $k \in K^\times$ with $\rho(k) = -1$, so $\sigma(k) = -k$ and $\sigma(k^2) = k^2$, whence $k \notin K^\sigma$ but $k^2 \in K^\sigma$. If $\mathrm{char}(K) = 2$, take a third root of unity ζ, and k with $\rho(k) = \zeta$, whence $k \notin K^\sigma$ but $k^3 \in K^\sigma$ (unless $\zeta \notin K^\sigma$). So K^σ is not algebraically closed, whence finite by Corollary 5.6.

Next, suppose $\sigma^q = \mathrm{id}$; we may assume that q is prime. Put $\tau = \sum_{i=0}^{q-1} \sigma^i$; since $\mathrm{im}(\tau) \leqslant K^\sigma$ and K^+ is connected, $\tau = 0$. If $q = 2$, then $\sigma = \mathrm{id}$. Otherwise find $k \in K^\times$ with $\sigma(k)/k = -1$. Then $\tau(k) = k \neq 0$, a contradiction.

If Σ is an infinite family of uniformly definable automorphisms of K, let F be a maximal subfield such that infinitely many $\sigma \in \Sigma$ agree on F (which exists by the icc). Then F is algebraically closed, and is the fixed field of $\sigma^{-1}\sigma'$ for some $\sigma, \sigma' \in \Sigma$ which agree on F, contradicting the first part. □

In particular, a superstable field of characteristic zero has no definable automorphism. In non-zero characteristic, any automorphism of a superstable field K is $\mathrm{acl}(\emptyset)$-definable and determined by its restriction to the algebraic closure k of the prime field; the group of definable automorphisms of K embeds into the group of automorphisms of k, which is pro-cyclic.

The following important theorem states that under certain circumstances the group generated by a family of type-definable sets is again type-definable.

DEFINITION 15. A type-definable subset X of G is called *α-indecomposable* if for all definable subgroups H of G either X is contained in a single coset of H, or

$$U_P(XH/H) \geqslant \omega^\alpha.$$

We say *indecomposable* instead of 0-indecomposable.

INDECOMPOSABILITY THEOREM [200,29]. *If $\mathfrak{X}$ is a family of type-definable α-indecomposable subsets containing 1 in a superstable group G with $U(G) < \omega^{\alpha+1}$, then $\langle \mathfrak{X} \rangle = X_1^{\pm 1} \cdots X_m^{\pm 1}$ for some $X_1, \ldots, X_m \in \mathfrak{X}$.*

PROOF. Choose $k < \omega$ maximal possible and $X_1, \ldots, X_n \in \mathfrak{X}$ such that $X_1^{\pm 1} \cdots X_n^{\pm 1}$ contains a type p with $U(p) = \omega^\alpha k$. By Lemma 5.1 a left translate of p is generic for $\mathrm{stab}(p^{-1})$; moreover $\mathrm{stab}(p^{-1}) \subseteq p^{-1}p$, and $X \subseteq \mathrm{stab}(p^{-1})$ by α-indecomposability for all $X \in \mathfrak{X}$. Hence $\langle \mathfrak{X} \rangle = \mathrm{stab}(p^{-1}) = (X_1^{\pm 1} \cdots X_n^{\pm 1})^{-1}(X_1^{\pm 1} \cdots X_n^{\pm 1})$. □

LEMMA 5.8. *If G is superstable, $H \leqslant G$ is type-definable connected with $U(H) = \omega^\alpha n$, and $g \in G$, then g^H and $[g, H]$ are α-indecomposable.*

PROOF. Suppose g^H is non-trivial, with $U(g^H K/K) < \omega^\alpha$ for some definable subgroup $K < G$. As $N = \bigcap_{h \in H} K^h$ is equal to a finite subintersection, we get

$$\omega^\alpha > U(g^H/N) = U(H/C_H(g/N)).$$

Hence $C_H(g/N) = H$ by Lascar's inequalities and connectivity, and g^H is contained in a single coset of N, whence of K. Thus g^H is α-indecomposable, as is $g^{-1}g^H = [g, H]$. □

COROLLARY 5.9. *A superstable non-Abelian group G without definable normal subgroups is simple.*

PROOF. $U(G) = \omega^\alpha n$ for some ordinal α and $n < \omega$ by Theorem 5.2. Any non-trivial conjugacy class generates a definable normal subgroup by Lemma 5.8 and the Indecomposability Theorem. □

Baudisch [18] has used this result to show that any superstable connected group G has a normal series $G = G_0 \triangleright G_1 \triangleright G_2 \triangleright \cdots \triangleright G_n = \{1\}$ such that G_i/G_{i+1} is either simple or Abelian, for all $i < n$.

COROLLARY 5.10. *In a superstable connected group G with $U(G) = \omega^\alpha n$, the groups in the derived and descending central series are type-definable.*

PROOF. Use Lemma 5.8 and the Indecomposability Theorem, with $H = G$. □

We shall now study the structure of soluble groups.

DEFINITION 16. If a type-definable group M acts definably on an infinite definable group G, then G is *M-minimal* if it has no proper infinite type-definable M-invariant subgroup.

PROPOSITION 5.11 [203]. *Suppose a type-definable Abelian group M acts definably on an M-minimal type-definable Abelian group A in a superstable structure, such that*

$$U(A) < U(M/C_M(A)) \cdot \omega.$$

Then there are a definable field K, a definable isomorphism $K^+ \cong A$, and a definable embedding $M/C_M(A) \hookrightarrow K^\times$.

PROOF. By Theorem 5.2 and M-minimality, A is connected and $U(A) = \omega^\alpha n$ for some α and $n < \omega$. If $C_M(A) = C_M(a_0, \ldots, a_k)$, then

$$\omega^\alpha \leqslant U(M/C_M(A)) \leqslant \bigoplus_{i \leqslant k} U(M/C_M(a_i)),$$

so there is $a \in A$ with $U(Ma) \geqslant \omega^\alpha$. By M-minimality, Ma is α-indecomposable, so $\langle Ma \rangle$ is type-definable, and equals A. By commutativity of M, the ring R of endomorphisms of

A generated by the action of M is type-definable; it is an integral domain by M-minimality of A, and an algebraically closed field K by Corollary 5.6. The assertion follows. $\square$

THEOREM 5.12 [122]. *If an infinite type-definable connected group G acts definably on the Abelian G-minimal type-definable group A in a superstable theory, and G has a definable normal Abelian subgroup M with $U(A) < U(M/C_M(A)) \cdot \omega$, then there is an infinite definable field K such that A is definably isomorphic to a finite-dimensional vector space over K on which G acts K-linearly and M scalarly.*

PROOF. If B is an M-minimal subgroup of A, put $N = \bigcap_{g \in G} C_M(B)^g$. Then $C_A(N) \geqslant B$ is G-invariant and equals A by G-minimality, whence $N = C_M(A)$. Hence

$$\begin{aligned} U\big(M/C_M(B)\big) \cdot \omega &\geqslant U(M/N) \cdot \omega = U\big(M/C_M(A)\big) \cdot \omega > U(A) \\ &\geqslant U(B); \end{aligned}$$

by Proposition 5.11 there is a definable field K with $B \cong K^+$ and $M/C_M(B) \hookrightarrow K^\times$. As $\langle Mb \rangle = B$ for any $b \in B$, there are no proper non-trivial M-invariant subgroups in B. Hence any sum $\sum_i g_i B$ is direct; since

$$U(A) < U\big(M/C_M(A)\big) \cdot \omega = U(K) \cdot \omega = U(B) \cdot \omega,$$

there is a bound on the length of such a sum, and $A = \bigoplus_{i=0}^k g_i B$.

If R is the ring of endomorphisms of A generated by M, the annihilators $\operatorname{ann}(g_i B)$ are prime ideals of R, and G permutes them by conjugation; by connectivity of G the action is trivial and $\operatorname{ann}(B) = \operatorname{ann}(gB)$ for all $g \in G$. Hence $\operatorname{ann}(B) = (0)$ and R embeds into K. By Theorem 5.7 the action of G on K by conjugation is trivial, so A is isomorphic to a vector space over K on which G acts K-linearly and M scalarly. $\square$

PROPOSITION 5.13. *If G is connected with $U(G) = \omega^\alpha n$, then $Z_\omega(G) = Z_n(G)$ for some $n < \omega$.*

PROOF. Choose $k \leqslant n$ maximal with $U(Z_i(G)) \geqslant \omega^\alpha k$ for some $i < \omega$. Then

$$\omega^\alpha > U\big(z^G/Z_i(G)\big) = U\big(G/C_G\big(z/Z_i(G)\big)\big)$$

for all $z \in Z_\omega(G)$, whence $U(C_G(z/Z_i(G))) = \omega^\alpha n$ and $z \in Z_{i+1}(G)$. $\square$

THEOREM 5.14 [122]. *If G is connected with $U(G) = \omega^\alpha n$, then G' is nilpotent.*

PROOF. By Proposition 5.13 we may assume $Z(G) = \{1\}$. Let $H = G^{(n)}$ be the last non-nilpotent group in the derived series; note that H is type-definable and connected by Corollary 5.10. Assume $n > 0$ and choose G-minimal $A \leqslant Z(H')$. If $U(A) < \omega^\alpha$, then $U(C_G(A)) = U(G)$ and A is central, a contradiction. Otherwise apply Theorem 5.12 to see that G/H' acts linearly and H/H' scalarly on A. As $\det(h) = 1$ for any $h \in H/H'$ (this is a commutator), $h = \zeta \cdot I$ for a root of unity ζ. By connectivity $h = 1$, whence $H = H'$, a contradiction. $\square$

6. Groups of finite Morley rank

In the introduction, we stated Cherlin's Conjecture for simple almost strongly minimal groups. These are just the simple groups of finite Morley rank:

LEMMA 6.1. *A simple group G or field K of finite Morley rank is almost strongly minimal.*

PROOF. There is a definable subset X of G or K of Morley rank and degree 1, i.e. a strongly minimal set; we may choose it inside a minimal infinite definable (additive) subgroup, with $1 \in X$ (or $0 \in X$). Then X^g (or gX) is indecomposable for every (non-zero) g; the family of these sets generates a definable normal subgroup (or ideal) by the Indecomposability Theorem, which must equal G (or K). $\square$

In the finite Morley rank context, we can improve on Corollary 5.10:

COROLLARY 6.2. *In a group of finite Morley rank, the groups in the derived and descending central series are type-definable.*

PROOF. By ω-stability, connected components have finite index and are definable. Now $[\gamma_i(G), G^0]$ is definable for any $i < \omega$ by Lemma 5.8 and the Indecomposability Theorem, as is $[\gamma_i(G)^0, G]$. Since $[\gamma_i(G), G]$ is a finite extension of $[\gamma_i(G)^0, G][\gamma_i(G), G^0]$, the assertion follows. $\square$

PROPOSITION 6.3 [114,126]. *A nilpotent group G of finite Morley rank is the central product of a divisible group D with a group B of bounded exponent; D has only finitely many elements of any given finite order.*

PROOF. If G is Abelian, choose $n < \omega$ with $nG = D$ minimal. Then D is divisible; if $g \in G$, then $ng \in D$ and there is $d \in D$ with $nd = ng$. Put $B = G[n]$, then $G = D + B$. Since $x \mapsto nx$ is surjective on D, it must have finite kernel and $D[n]$ is finite. For nilpotent G consider $G/Z(G)^0$ and use induction. $\square$

Recall from the introduction that a *bad field* is a field of finite Morley rank with a definable multiplicative subgroup T. It seems likely that they exist, at least in characteristic zero [156]; if they exist in characteristic $p > 0$, the methods of [191] show that there must be a locally finite one. However, there are restrictions:

PROPOSITION 6.4. *An definable multiplicative subgroup M with $U(M) \geqslant \omega^\alpha$ of a superstable field K with $U(K) = \omega^\alpha n$ generates K additively.*

PROOF. If H is a definable subgroup of K^+ with $U(MH/H) < \omega^\alpha$, put $N = \bigcap_{m \in M} mH$ (which is a finite intersection), and $R = \{k \in K\colon\ kN \leqslant N\}$, a subring of K containing M which must equal K. But $U(MN/N) < \omega^\alpha$, so $N \neq 0$, whence $N = H = K$ and the Indecomposability Theorem applies. $\square$

REMARK 4. If K has characteristic 0 and finite Lascar rank, it has no non-trivial proper relatively definable additive subgroups, and any additive endomorphism $\sigma : K^n \to K^m$ is linear.

PROOF. If A is an additive subgroup of K^+, consider $R = \{k \in K\colon\ kA \leqslant A\}$. For the second assertion, put $R = \{k \in K\colon\ \forall \bar{x} \in K^n\ \sigma(k\bar{x}) = k\sigma(\bar{x})\}$. In either case, $R = K$. □

Bad groups are described by Nesin [121], Borovik and Poizat [39] and Corredor [59]:

THEOREM 6.5. *Let G be a non-soluble connected group of finite Morley rank such that every definable connected proper subgroup is nilpotent. Call a maximal proper connected subgroup a Borel subgroup. Then:*

(i) *all Borel subgroups are conjugate.*
(ii) *distinct Borel subgroups intersect trivially.*
(iii) *the Borel subgroups cover G.*
(iv) *every finite subgroup is contained in some Borel subgroup.*
(v) *G has no involutions.*
(vi) *Borel subgroups are self-normalizing.*
(vii) *if B is a Borel subgroup and $g \in G - B$, then $G = BgB \ldots gB$.*

In analogy to the revised proof of the classification of finite simple groups, call a group of finite Morley rank a *K-group* if its simple definable sections are algebraic groups over algebraically closed fields, and a *K^*-group* if all its proper simple definable sections are K-groups. We say that G is *tame* if it does not interpret a bad field or a bad group.

LEMMA 6.6. *A tame simple group G of finite Morley rank has infinite 2-Sylow subgroups.*

PROOF. By tameness, G has a soluble non-nilpotent section which is isomorphic to $K^+ \rtimes K^\times$ by Proposition 5.11. Either K^+ or $K^\times$ has an infinite 2-subgroup, which lifts to an infinite 2-subgroup of G by Corollary 3.25. □

As for finite groups, there should be two cases: $\mathrm{char}(K) = 2$ and $\mathrm{char}(K) \neq 2$.

DEFINITION 17. A group is of *even* type if its Sylow 2-subgroups are of bounded exponent; it is of *odd* type if its Sylow 2-subgroups are divisible-by-finite.

THEOREM 6.7 [100]. *A simple K^*-group of finite Morley rank has even or odd type.*

Note that Theorem 6.7 does not use tameness (in our definition, finite Sylow 2-subgroups are both even and odd). Its proof uses the fact that in a K-group a divisible connected nilpotent subgroup and a nilpotent connected subgroup of bounded exponent commute. Assuming that G is a counterexample, it derives the existence of a *weakly embedded* subgroup M (i.e. M has infinite Sylow 2-subgroups, but $M \cap M^g$ has finite Sylow 2-subgroups for all $g \in G - M$). But such a group cannot be simple.

Work on even type simple tame groups of finite Morley rank is proceeding well and the classification seems within reach of present methods [4]. In fact, there is even hope to eliminate tameness from the assumptions and prove:

EVEN TYPE CONJECTURE. A simple group G of finite Morley rank without infinite degenerate definable sections is an algebraic group over an algebraically closed field.

Here a group is *degenerate* if its Sylow 2-subgroups are finite.
By contrast, the groups of odd type are less well understood.

DEFINITION 18. The *2-rank* $m(G)$ of a group G is the maximal rank of an elementary Abelian 2-subgroup of G; the *normal 2-rank* $n(G)$ is the maximal rank of a normal elementary Abelian subgroup of a Sylow 2-subgroup of G. The *2-generated core* $\Gamma_{S,2}(G)$ of a Sylow 2-subgroup $S \leqslant G$ is the minimal definable subgroup of G containing $N_G(U)$ for all $U \leqslant S$ with $m(U) \geqslant 2$.

An involution i of G is *classical* if $C_G(i)$ has a subnormal subgroup A isomorphic to $SL_2(K)$ for some algebraically closed field K with $i \in Z(A)$; the component A is a *classical subgroup* of G.

G satisfies the *B-conjecture* if $C_G(i)^0 = F(C_G(i))^0 E(C_G(i))$ for all involutions $i \in G$, where $E(H)$ is the product of all quasi-simple subnormal subgroups of H.

In a group G of finite Morley rank of odd type, $n(H)$ and $m(H)$ are finite for any $H \leqslant G$ by Proposition 6.3. By definition $\Gamma_{S,2}(G)$ is definable; moreover, by Corollary 3.2 there is a bound on the number of quasi-simple subnormal subgroups, and $E(H)$ is definable as well [25,125,182].

THEOREM 6.8 [34]. *If G is a simple K^*-group of finite Morley rank of odd type which is tame or locally finite, then either*

(i) $n(G) \leqslant 2$,
(ii) *G has a proper 2-generated core, or*
(iii) *G satisfies the B-conjecture and contains a classical involution.*

CONJECTURE 1 [34]. If S is a Sylow 2-subgroup of a simple tame K^*-group G of finite Morley rank, then $\Gamma_{S,2}(G) = G$.

If K is a classical subgroup of G, let $\mathcal{D}_K$ be the graph with vertex set $\{K^g\colon g \in G\}$ and edges $\{(K^g, K^{g'})\colon [K^g, K^{g'}] = 1\}$.

THEOREM 6.9 [28]. *If G is a simple tame K^*-group of finite Morley rank and odd type which satisfies the B-conjecture and contains a classical involution with connected graph $\mathcal{D}_K$, then G is isomorphic to one of the following groups: $PSL_n(F)$ for $n \geqslant 5$, $PSp_{2n}(F)$ for $n \geqslant 3$, $PSO_n(F)$ for $n \geqslant 9$, $E_6(F)$, $E_7(F)$, $E_8(F)$ or $F_4(F)$, for some algebraically closed field F with* $\mathrm{char}(F) \neq 2$.

Groups not covered by this theorem are the small groups $PSL_2(F)$, $PSL_3(F)$, $G_2(F)$, $PSO_5(F) \cong PSp_4(F)$, $PSO_6(F) \cong PSL_4(F)$, $PSO_7(F)$ and $PSO_8(F)$.

CONJECTURE 2. If $\mathcal{D}_K$ is disconnected but has an edge, then $G \cong PSO_n(F)$; if $\mathcal{D}_K$ has no edge, then $G \cong PSL_3(F)$, for some algebraically closed field F.

In the absence of the tameness assumption (or just to study ω-stable groups in general), it would be helpful to have some substitute for Sylow's theorems. In the soluble case, the theory of formations looks like a promising candidate.

DEFINITION 19. The *Frattini subgroup* $\Phi(G)$ is the intersection of all proper maximal connected definable subgroups of G.

A family $\mathfrak{F}$ of connected soluble groups is a *connected formation* if

(i) it is closed under homomorphic images;
(ii) it is closed under direct products;
(iii) if $N \lhd G$ is finite and $G/N \in \mathfrak{F}$, then $G \in \mathfrak{F}$.

We call $\mathfrak{F}$ *saturated* if $G/\Phi(G) \in \mathfrak{F}$ implies $G \in \mathfrak{F}$. If G is a connected soluble group of finite Morley rank and $\mathfrak{F}$ a connected formation, let $G_{\mathfrak{F}}$ be the smallest definable normal subgroup of G with $G/G_{\mathfrak{F}} \in \mathfrak{F}$. A definable subgroup $H \leqslant G$ is *$\mathfrak{F}$-covering* if $K = K_{\mathfrak{F}}H$ for all definable $K \geqslant H$.

If π is a set of primes, a *π^*-group* is a connected soluble group whose definable connected Abelian sections are π-divisible.

LEMMA 6.10. *The class $\mathfrak{N}$ of connected nilpotent groups of finite Morley rank and the class $\mathfrak{F}_\pi$ of π^*-groups of finite Morley rank are saturated connected formations.*

THEOREM 6.11 [72]. *A connected formation $\mathfrak{F}$ is saturated iff every soluble connected group of finite Morley rank has an $\mathfrak{F}$-covering subgroup. If $\mathfrak{F}$ is saturated, any two $\mathfrak{F}$-covering subgroups are conjugate.*

$\mathfrak{N}$-covering groups are self-normalizing nilpotent subgroups (*Carter subgroups*), and $\mathfrak{F}_\pi$-covering subgroups are maximal π^*-subgroups. If G is soluble of finite Morley rank, then $G_{\mathfrak{N}} = \bigcap_{i<\omega} \gamma_i(G)$ (which equals $\gamma_k(G)$ for big k) and $G_{\mathfrak{F}_\pi}$ is the maximal normal definable connected π-subgroup of G of bounded exponent.

The properties of Carter subgroups generalize to the superstable context and beyond [179]; it is likely that this is true for the results about formations in general.

7. One-based groups

DEFINITION 20. Call $\mathfrak{M}$ *one-based* if $\bar{a} \perp_{\mathrm{acl}(\bar{a}) \cap \mathrm{acl}(\bar{b})} \bar{b}$ for all tuples $\bar{a}$ and $\bar{b}$.

Clearly, $\mathrm{acl}(\bar{a}) \cap \mathrm{acl}(\bar{b})$ is the smallest algebraically closed set over which $\bar{a}$ and $\bar{b}$ can be independent.

EXAMPLE 2. Let $\mathfrak{M}$ be a module over a ring R in the module language

$$\mathcal{L} = \{0, +, \lambda_r \colon r \in R\},$$

where λ_r is scalar multiplication by r. Then every definable subset of $\mathfrak{M}^n$ is a Boolean combination of cosets of $\mathrm{acl}(\emptyset)$-definable subgroups of $\mathfrak{M}^n$. It follows that $\mathrm{Cb}(\overline{m}/\mathfrak{M})$ is given by the canonical parameters for the $\mathfrak{M}$-definable cosets of $\mathrm{acl}(\emptyset)$-definable subgroups containing $\overline{m}$, which are all in $\mathrm{acl}(\overline{m})$. This implies that $\mathfrak{M}$ is one-based.

More generally, an *Abelian structure* is an Abelian group A together with a collection of predicates for subgroups of A^n, for various $n < \omega$. Again every definable set is equivalent to a Boolean combination of cosets of $\mathrm{acl}(\emptyset)$-definable subgroups, and any Abelian structure is stable and one-based. The aim of this section is to establish some kind of converse.

PROPOSITION 7.1. *A connected type-definable subgroup H of a stable one-based group G is type-definable over* $\mathrm{acl}(\emptyset)$.

PROOF. Put $u = \mathrm{Cb}(H)$. If h is generic for H over u, and g is principal generic for G over u, h, then hg is generic for Hg over u, g, and principal generic for G over u, h. Let $v = \mathrm{Cb}(Hg) = \mathrm{Cb}(hg/g, u)$, so $v \in \mathrm{dcl}(g, u)$, and $v \in \mathrm{acl}(hg)$ by one-basedness. As $H = (Hg)(Hg)^{-1}$, we get $u \in \mathrm{dcl}(v) \subseteq \mathrm{acl}(hg)$. But $hg \mathop{\smile\!\!\!\!\!\mid} u$, so $u \in \mathrm{acl}(\emptyset)$. □

THEOREM 7.2. *A stable one-based group is Abelian-by-finite.*

PROOF. Assume that G is connected, and let $H_g = \{(h, h^g) \in G^2\colon\ h \in G\}$, for $g \in G$. Then H_g is connected, and type-definable over $\mathrm{acl}(\emptyset)$ by Proposition 7.1. As $G/Z(G)$ is in bijection with $\{H_g\colon\ g \in G\}$, compactness yields that $G/Z(G)$ is finite. □

THEOREM 7.3 [91]. *If G is stable one-based and $p \in S_1(G)$, then p is a generic type of a G-definable coset of a connected type-definable subgroup $H \leqslant G$. Every definable subset of G is a Boolean combination of cosets of* $\mathrm{acl}(\emptyset)$*-definable subgroups.*

PROOF. If $\mathrm{tp}(h/G)$ is generic and $x \models p$ with $h \mathop{\smile\!\!\!\!\!\mid}_G x$, put

$$C = \mathrm{acl}\big(\mathrm{Cb}(hx/G, h), G\big)$$

and choose $h' \mathop{\smile\!\!\!\!\!\mid}_C h, x$ with $\mathrm{tp}(h'/C) = \mathrm{tp}(h/C)$. Then $hx = h'x'$ for some $x' \models p$. So $x \mathop{\smile\!\!\!\!\!\mid}_{C,h} h'^{-1}h$ and $x \mathop{\smile\!\!\!\!\!\mid}_G C, h$, whence $x \mathop{\smile\!\!\!\!\!\mid}_G h'^{-1}h$. Thus $h'^{-1}h \in \mathrm{stab}(p)$.

Put $u = \mathrm{Cb}(h\,\mathrm{stab}(p))$, then $u \in \mathrm{acl}(G, h) \cap \mathrm{acl}(G, h') = C$; by one-basedness

$$u \in \mathrm{acl}\big(\mathrm{Cb}(hx/G, h), G\big) \subseteq \mathrm{acl}(G, hx).$$

If $v = \mathrm{Cb}(h\,\mathrm{stab}(p)\,x)$, then hx is generic for $hx(x^{-1}\,\mathrm{stab}(p)\,x) = h\,\mathrm{stab}(p)\,x$ over G, v, so $v \in \mathrm{Cb}(hx/G, v)$. By one-basedness $v \in \mathrm{Cb}(hx/G, v) \subseteq \mathrm{acl}(hx)$.

If $w = \mathrm{Cb}(\mathrm{stab}(p)\,x)$, then

$$\mathrm{stab}(p)\,x = \big(h\,\mathrm{stab}(p)\big)^{-1}\big(h\,\mathrm{stab}(p)\,x\big)$$

implies

$$w \in \mathrm{dcl}(u, v) \subseteq \mathrm{acl}(G, hx).$$

By genericity $hx \mathop{\smile\!\!\!\!\!\!\mid}_G x$, whence $w \mathop{\smile\!\!\!\!\!\!\mid}_G x$. So $w \in \mathrm{dcl}(G, x)$ implies $w \in \mathrm{acl}(G)$, and $\mathrm{stab}(p)\,x$ is type-definable over G. It follows that p is generic for $\mathrm{stab}(p)\,x$.

Finally, as a connected group is type-definable over $\mathrm{acl}(\emptyset)$ by Proposition 7.1 and has a unique generic type, the last assertion follows. □

Call a type $p \in S(A)$ *regular* if whenever $a, b \models p$ and $a \mathop{\smile\!\!\!\!\!\!\mid}_A B$ but $b \not\mathop{\smile\!\!\!\!\!\!\mid}_A B$, then $a \mathop{\smile\!\!\!\!\!\!\mid}_A Bb$ (so non-forking realizations cannot fork with forking realizations). A group is *regular* if it has a regular generic type. For instance, the Lascar inequalities imply that a type of Lascar rank ω^α is regular. If $\mathrm{tp}(a/A)$ is regular and $b \mathop{\smile\!\!\!\!\!\!\mid}_A c$, then either $a \mathop{\smile\!\!\!\!\!\!\mid}_A b$ or $a \mathop{\smile\!\!\!\!\!\!\mid}_A c$. If X is the set of realizations of p in $\mathfrak{C}$ and $A \subset X$, the cardinality $\dim(A)$ of a maximal independent subset of A is uniquely determined; p is *locally modular* if $\dim(A \cup B) + \dim(A \cap B) = \dim(A) + \dim(B)$ for any closed (w.r.t. dependence) subsets A and B of X with $A \cap B \neq \emptyset$.

A theory of finite Lascar rank is one-based iff every type of rank 1 (which is regular) is locally modular. For infinite rank, local modularity is more general than one-basedness; nevertheless, the previous results in this section still hold (with algebraic closure replaced by a suitable p-closure). In particular, Theorem 7.3 yields a certain division ring which controls forking on X.

DEFINITION 21. Let G be a connected Abelian stable group with regular generic type, and G_0 the (undefinable) subgroup of all non-generic elements of G. A *p-endogeny* is an endomorphism of G/G_0 induced by a definable subgroup of G^2.

It is easy to see that the p-endogenies form a division ring.

THEOREM 7.4 [81]. *If R is the ring of p-endogenies of a connected Abelian group G with locally modular generic type p, then $g_0, g_1, \ldots, g_n \in G$ are dependent iff there are $r_i \in R$, not all zero, with $\sum_{i=0}^n r_i(g_i) = G_0$.*

PROOF. "If" is obvious, so suppose $\{g_0, \ldots, g_n\}$ is dependent, and any subtuple is independent. Then $\mathrm{tp}(g_0, \ldots, g_n/G_0)$ is the generic type of a coset of a connected G_0-definable group H; put

$$r_i = \{(x_i, x_n) \in G^2\colon\ (\bar{0}, x_i, \bar{0}, x_n) \in H\}$$

for $i < n$, and $r_n(x) = -x$. Then $r_i \in R$ for all $i \leqslant n$, and $\sum_{i \leqslant n} r_i(g_i) = G_0$. □

There is a generalization of one-basedness (and local modularity), which allows the forking geometry to be slightly more complicated [87].

DEFINITION 22. A stable structure $\mathfrak{M}$ is *CM-trivial* if $\mathrm{acl}(aA) \cap \mathrm{acl}(B) = \mathrm{acl}(A)$ (for some $A \subseteq B$) implies $\mathrm{Cb}(a/A) \subseteq \mathrm{acl}(\mathrm{Cb}(a/B))$.

Using Theorem 6.5, Pillay [138] has shown that a CM-trivial group of finite Morley rank is nilpotent-by-finite; this can be generalized to the superstable context [187]. We should note that Hrushovski's amalgamation constructions [87,181] yield CM-trivial structures. An infinite field is never CM-trivial.

I should finish this section with some remarks on the two theorems by Hrushovski [88, 90] mentioned in the introduction:

MORDELL–LANG CONJECTURE. Let A be an Abelian variety with zero trace over the algebraic closure k of the prime field, and X a subvariety of A. If Γ is a subgroup of A of finite type, then $X \cap \Gamma$ is a finite union of cosets of subgroups of A.

MANIN–MUMFORD CONJECTURE. If A is an Abelian variety and X a subvariety of A over a number field k, then the torsion subgroup Γ of A intersects X in a finite union of cosets of subgroups of A.

The conclusions of the two theorems are equivalent to the statement that $X \cap \bar{\Gamma}$ is one-based, for some supergroup $\bar{\Gamma}$ of Γ of finite Lascar rank which is type-definable in a suitable expansion of the pure field structure. This expansion is a differentially or a separably closed field for the Mordell–Lang Conjecture, and an existentially closed field with a named automorphism for the Manin–Mumford Conjecture; the first structure is superstable, the second stable, and the last one simple (see Section 9). In all three the Trichotomy Conjecture holds; one thus has to exclude the possibilities that $X \cap \bar{\Gamma}$ is degenerate or a field. (For more details, see [43,53,137].)

8. The group configuration and the binding group

In this section, we shall reconstruct definable groups from geometric data.

GROUP CONFIGURATION THEOREM. *Suppose $a_1, a_2, a_3, b_1, b_2, b_3$ are tuples in a stable structure satisfying:*

(i) $a_i \in \operatorname{acl}(a_j, a_k)$ *for* $\{i, j, k\} = \{1, 2, 3\}$.

(ii) $b_i \in \operatorname{acl}(a_j, b_k)$ *for* $\{i, j, k\} = \{1, 2, 3\}$.

(iii) *all other triples and all pairs are independent.*

Then there is a type-definable connected group G acting transitively on a set X, such that for a generic element $x \in X$ (with respect to that action) there is $b \models \operatorname{tp}(b_1)$ such that $\operatorname{acl}(x) = \operatorname{acl}(b)$ (possibly after adding some parameters).

The proof proceeds by first adding some independent parameters A and replacing the original sixtuple by a new sixtuple $(a'_1, a'_2, a'_3, b'_1, b'_2, b'_3)$, such that $\operatorname{acl}(a_i A) = \operatorname{acl}(a'_i A)$ and $\operatorname{acl}(b_i A) = \operatorname{acl}(b'_i A)$ for $i = 1, 2, 3$, but such that in addition $b_i \in \operatorname{dcl}(a_j, b_k)$ for $\{i, j, k\} = \{1, 2, 3\}$. So every $a \models \operatorname{tp}(a'_1/A)$ yields a definable function $f_a : \operatorname{tp}(b'_2/A) \to \operatorname{tp}(b'_3/A)$; two such functions have the same *germ* if they agree on some (equivalently: all) independent $b \models \operatorname{tp}(b'_2/A)$. Germs are quotients by a definable equivalence relation, so the family of germs $\bar{f}_a$ (for $a \models \operatorname{tp}(a'_1/A)$) is uniformly definable. Similarly, we get a family of germs $\bar{g}_{a'}$ from functions $g_{a'} : \mathrm{tp}(b'_3/A) \to \mathrm{tp}(b'_1/A)$ for any $a' \models \mathrm{tp}(a'_2/A)$; composition of functions induces composition of germs. It can then be shown that if $(a, a', a'') \models \mathrm{tp}(a'_1, a'_2, a'_3/A)$, then $\mathrm{Cb}(\bar{g}_{a'} \circ \bar{f}_a) \in \mathrm{acl}(a''A)$ and is thus independent over A of $\mathrm{Cb}(\bar{g}_{a'}) \in \mathrm{dcl}(a')$ and of $\mathrm{Cb}(\bar{f}_a) \in \mathrm{dcl}(a)$. This implies that $\mathrm{Cb}(\bar{f}_{a'}^{-1} \circ \bar{f}_a)$ is independent of $\mathrm{Cb}(\bar{f}_{a'})$ and $\mathrm{Cb}(\bar{f}_a)$ for independent $a, a' \models \mathrm{tp}(a'_1/A)$; we can now apply Theorem 4.7 to finish the proof. ([205,80]; see [42] for details.)

Call a type p *degenerate* (or *trivial*) if it has degenerate geometry (i.e. any pairwise independent n-tuple is independent, for all $n < \omega$). A generic type in a stable group cannot be trivial, as for two independent principal generic elements g and g' the triple (g, g', gg') is pairwise independent, but not independent. Surprisingly, there is some kind of converse for one-based theories.

COROLLARY 8.1. *If p is a non-trivial type in a stable one-based theory, then there is a type-definable connected Abelian group G with generic type q, such that there are dependent $a \models p$ and $g \models q$.*

PROOF. Adding parameters, by non-triviality we find $a_1, a_2, a_3 \models p$ which are pairwise independent, but dependent; replacing a_i by $\mathrm{acl}(a_i) \cap \mathrm{acl}(a_j, a_k)$ for $\{i, j, k\} = \{1, 2, 3\}$, we may assume $a_i \in \mathrm{acl}(a_j, a_k)$. Let $(b_2, b_3) \models \mathrm{tp}(a_2, a_3/a_1)$ independently of a_2a_3. Put $b_1 = \mathrm{acl}(a_2, b_3) \cap \mathrm{acl}(a_3, b_2)$; by one-basedness $a_2b_3 \mathop{\smile\!\!\!\!\!\mid}_{b_1} a_3b_2$. So $b_3 \mathop{\smile\!\!\!\!\!\mid}_{b_1, a_2} a_3b_2$; since $a_1 \in \mathrm{acl}(a_2, a_3)$ and $b_3 \in \mathrm{acl}(a_1, b_2)$, we have $b_3 \in \mathrm{acl}(b_1, a_2)$. The other dependencies for the group configuration follow similarly, and we can apply the Group Configuration Theorem. □

A similar theorem holds for a locally modular non-trivial regular type p in a stable theory [81]. So Theorem 7.4 serves to describe the forking geometry for any such type.

DEFINITION 23. A family Σ of partial types is *A-invariant* if it invariant under all automorphisms of $\mathfrak{C}$ fixing A pointwise.

DEFINITION 24. Let Σ be an A-invariant family of partial types in a stable theory. A partial type π over A is *Σ-internal* if for every $a \models \pi$ there is $B \mathop{\smile\!\!\!\!\!\mid}_A a$, types $\bar{\sigma} \in \Sigma$ based on B, and $\bar{c} \models \bar{\sigma}$ with $a \in \mathrm{dcl}(B\bar{c})$.

Internality is transitive, and preserved under non-forking extensions. Moreover, we can choose a fixed B in the definition which does not vary with a.

DEFINITION 25. A complete type $p \in S(A)$ in a stable theory is *foreign* to an A-invariant family Σ of partial types if $a \mathop{\smile\!\!\!\!\!\mid}_B \bar{c}$ for all $a \models p$, $B \mathop{\smile\!\!\!\!\!\mid}_A a$, and $\bar{c} \models \bar{\sigma}$ with $\bar{\sigma} \in \Sigma \cap S(B)$.

LEMMA 8.2. *If Σ is an A-invariant family of partial types in a stable theory and* $\mathrm{tp}(a/A)$ *is not foreign to Σ, then there is $a_0 \in \mathrm{dcl}(Aa) - \mathrm{acl}(A)$ such that* $\mathrm{tp}(a_0/A)$ *is Σ-internal.*

PROOF. We take $B \mathop{\smile\!\!\!\!\!\mid}_A a$ and $\bar{c}$ realizing types in Σ with $a \not\mathop{\smile\!\!\!\!\!\mid}_{AB} \bar{c}$. Choose

$$a_1 \in \mathrm{Cb}(B\bar{c}/Aa) - \mathrm{acl}(A),$$

and let a_0 be the finite set of Aa-conjugates of a_1 (an imaginary element). $\square$

If $\mathrm{tp}(a/A)$ is the generic type of a stable group, we can take $a_0 = aN$ for some definable normal subgroup N:

THEOREM 8.3 [80]. *If G is a stable group whose generic type is not foreign to some partial type π, then there is a relatively definable normal subgroup N of infinite index such that G/N is π-internal.*

PROOF. Possibly after increasing the model G, choose generic g over G and $a \models \pi$ with $g \not\perp_G a$. Put $C = \mathrm{Cb}(g, a/G)$ and

$$H = \{g' \in G\colon \mathrm{tp}(g'g, a/C) = \mathrm{tp}(g, a/C)\}.$$

Then H is a subgroup of G and type-definable over C; it has infinite index since $hg \perp_G a$ for generic $h \perp_G g, a$.

Take a principal generic h with $h \perp_G g, a$ and a Morley sequence $(g_i, a_i\colon i < \omega)$ in $\mathrm{tp}(g, a/C, h)$. As $(g_i\colon i < \omega)$ and $(hg_i\colon i < \omega)$ are both Morley sequences in $\mathrm{tp}(g/C, h)$ and thus have the same type, there are $(a'_i\colon i < \omega)$ realising π such that

$$\mathrm{tp}(g_i, a'_i\colon i < \omega/C, h) = \mathrm{tp}(hg_i, a_i\colon i < \omega/C, h).$$

The latter is a Morley sequence in $\mathrm{tp}(hg, a/C, h)$, so

$$\mathrm{Cb}(hg, a/C, h) \in \mathrm{dcl}(g_i, a'_i\colon i < \omega).$$

As $\mathrm{Cb}(hH) \in \mathrm{Cb}(hg, a/C, h)$ and $\mathrm{Cb}(hH) \perp_C (g_i\colon i < \omega)$, $\mathrm{tp}(\mathrm{Cb}(hH)/C)$ is π-internal. But hH is generic in G/H, so G/H is π-internal, as is G/H^g for every $g \in G$; the same holds for any definable supergroup $\overline{H}$ of H of infinite index in G. Since $N = \bigcap_{g\in G} \overline{H}^g$ equals a finite subintersection, G/N is π-internal. $\square$

Note that the converse is obvious: if N is a subgroup of G, then a generic type of G is foreign to G/N iff $N \geqslant G^0$. In general, Theorem 8.3 allows us to find particular type-definable sections of a stable group which are closely related to a given partial type.

DEFINITION 26. If $p, q \in S(A)$, the *binding group* of p over q is the group of automorphisms of $p^{\mathfrak{C}}$ which fix A and $q^{\mathfrak{C}}$.

BINDING GROUP THEOREM. *Suppose $p, q \in S(A)$ in a stable structure, with $A = \mathrm{acl}(A)$. If p is q-internal, then there are an A-definable function $f(\bar{x}, \bar{y})$ and a tuple $\bar{a}$ of realizations of p, such that for any $a \models p$ there is a tuple $\bar{b}$ of realizations of q with $a = f(\bar{a}, \bar{b})$. The binding group of p over q is type-definable.*

Such a tuple $\bar{a}$ is called a *fundamental system of solutions* of p relative to q.

PROOF. Choose $B \supseteq A$ such that for all $a \models p$ there is a tuple $\bar{b}$ realizing q with $a \in \mathrm{dcl}(B\bar{b})$. If $(a_i, \bar{b}_i\colon i < \omega)$ is a Morley sequence in $\mathrm{tp}(a, \bar{b}/B)$, then $a \in \mathrm{dcl}(\bar{b}, a_i, b_i\colon i < n)$ for some $n < \omega$. Put $\bar{a} = (a_i\colon i < n)$. Then $a \mathop{\smile\!\!\!\!\!\!\;|}_A \bar{a}$, and $a \in \mathrm{dcl}(A, \bar{a}, q^{\mathfrak{C}})$.

If $a' \models p$, choose $\bar{a}' \models \mathrm{tp}(\bar{a}/A)$ with $\bar{a}' \mathop{\smile\!\!\!\!\!\!\;|}_A a', \bar{a}$. Then $a'_i \in \mathrm{dcl}(A, \bar{a}, q^{\mathfrak{C}})$ for all $a'_i \in \bar{a}'$, and $a' \in \mathrm{dcl}(A, \bar{a}', q^{\mathfrak{C}})$, so $a \in \mathrm{dcl}(A, \bar{a}, q^{\mathfrak{C}})$.

The action on $p^{\mathfrak{C}}$ of any automorphism of $\mathfrak{C}$ fixing A and $q^{\mathfrak{C}}$ is determined completely by the image of $\bar{a}$, and vice versa. Put $\bar{p} = \mathrm{tp}(\bar{a}/A, q^{\mathfrak{C}})$. One can show that for any $\bar{a}' \models \bar{p}$ there is an automorphism of $\mathfrak{C}$ fixing A and $q^{\mathfrak{C}}$ and mapping $\bar{a}$ to $\bar{a}'$ (this is not obvious from saturation of $\mathfrak{C}$, since $|\bar{p}| = |\mathfrak{C}|$). Now composition of automorphisms induces a definable group operation on $\bar{p}$, turning it into the binding group. $\square$

9. Final remarks

Poizat noted that an equation which is satisfied generically by an algebraic group defines a closed generic subset, and is therefore satisfied by the entire group. He asked whether an analogous property holds in stable groups. This is only known for particular equations (nilpotency, solubility, small exponent [154,174,176]), or nilpotent-by-finite groups [98]; the results have been generalized to arbitrary groups (with a suitable notion of *generic*) in [99].

A recent important development is the extension from stable theories to *simple* theories (an unfortunate misnomer in the context of groups) [166,167,102–107,77,49,50]. Call a Theory *simple* if the non-forking relation is symmetric. In such theories independence satisfies properties (1)–(5) of Theorem 2.1; however (6) (boundedness) fails: a type over a model will in general have many non-forking extensions to any superset, and will not be definable. Objects which are definable in a stable theory will in general only be type-definable in a simple one; in particular the canonical base is only the class of an infinite tuple by a type-definable equivalence relation (a *hyperimaginary* element). This causes some technical problems, but on the whole the theory goes through: definable groups in simple theories satisfy the chain conditions and decomposition properties, up to commensurability. A type-definable group in a supersimple theory is $\bigwedge$-definable; analogues of Theorems 4.7 and 5.2, as well as the Indecomposability Theorem hold, and supersimple division rings are commutative perfect bounded. One-based groups with simple theory are finite-by-Abelian-by-finite [143,141,144,189,190]. However, some of the basic results about stable groups are still only conjectural in the simple case:

CONJECTURE 3. A supersimple group G with $SU(G) \geqslant \omega^\alpha$ contains an Abelian subgroup A with $SU(A) \geqslant \omega^\alpha$.

SU-rank is the analogue of Lascar rank for simple theories. One does not even know that an infinite supersimple group contains an infinite Abelian subgroup.

CONJECTURE 4. A supersimple field is perfect bounded pseudo-algebraically closed.

Hrushovski [85] has shown that perfect bounded pseudo-algebraically closed fields are supersimple of rank 1.

CONJECTURE 5. An ω-categorical supersimple theory has finite rank.

An ω-categorical superstable theory has finite rank [57]; it is shown in [70] that an ω-categorical supersimple group is finite-by-Abelian-by-finite, of finite SU-rank (compare with Theorem 3.22).

On aspect of stable groups which has only been touched on is the extensive model theory of modules (Abelian structures); for an introduction see [199,160]. As for other examples, as already mentioned in the introduction, certain theories of fields play an important rôle in the applications. Those include:

1. Algebraically closed fields in any characteristic; they are strongly minimal.
2. Differentially closed fields of characteristic 0; they are ω-stable of rank ω, the fixed field has rank 1 and is algebraically closed.
3. Separably closed fields of characteristic $p > 0$ and fixed Eršov invariant; they are stable non-superstable.
4. Pseudofinite fields; they are unstable supersimple of SU-rank 1.
5. Existentially closed difference fields of characteristic 0; they are unstable supersimple of SU-rank ω, the fixed field has SU-rank 1.

Good references are [117,162,164,146,140,94,198,8,92,93,85,43,53,54].

Another source of examples comes from Hrushovski's generalization of Fraïssé's universal-homogeneous model [73,86,109,87,181,13,15]; this has been adapted to the case of nilpotent goups of exponent p by Baudisch [20,19]. Fields of rank two are constructed by Poizat [157] and Baldwin and Holland, [10]. Recently, Hrushovski [89] has modified the construction to yield simple unstable structures.

Finally, there is an intriguing relationship between structures of finite Morley rank and certain ordered structures. An ordered structure $\mathfrak{M}$ is *o-minimal* if every definable subset is a finite union of open intervals (with endpoints in $\mathfrak{M} \cup \{\pm\infty\}$) and points; this is an ordered analogue of a strongly minimal set. A real closed field is o-minimal; more generally, the ordered field $\mathbb{R}$ expanded by all Pfaffian functions (including exponentiation) is o-minimal [170,195–197]. The properties of an o-minimal structures are similar to those of a structure of finite Morley rank, although in general the proofs are quite different. In particular, the Trichotomy Theorem and some version of the Indecomposability Theorem hold in any o-minimal structure [132]. A good reference is [169]; Groups definable in such a structure are classified in [136,129–131,133]. For attempts to find an ordered analogue to superstable structures see [26,27,149].

10. Bibliographical remarks

The classic text on model theory is Chang and Keisler [51]; a more modern approach (with a focus on stability theory) is Poizat [152] or Hodges [79]. Textbooks on Stability theory include [134,110,111,48]; [139] leads up to research level. Of course, one may also consult the original work [165]. An introduction to the extension of stability theory to simple structures will appear in [193]. Groups of finite Morley rank are treated in [37], stable groups (with a certain emphasis on ω-stable and superstable groups) in [153] and (in full generality) [186].

References

[1] I. Aguzarov, R.E. Farey and J.B. Goode, *An infinite superstable group has infinitely many conjugacy classes*, J. Symbolic Logic **56** (1991), 618–623.

[2] T. Altınel, *Groups of finite Morley rank with strongly embedded subgroups*, J. Algebra **180** (1996), 778–807.

[3] T. Altınel, A.V. Borovik and G. Cherlin, *Groups of mixed type*, J. Algebra **192** (1997), 524–571.

[4] T. Altınel, A.V. Borovik and G. Cherlin, *Groups of finite Morley rank and even type with strongly closed Abelian subgroups*, Preprint (1999).

[5] T. Altınel, A.V. Borovik and G. Cherlin, *On groups of finite Morley rank with weakly embedded subgroups*, J. Algebra **211** (1999), 409–456.

[6] T. Altınel and G. Cherlin, *On central extensions of algebraic groups*, J. Symbolic Logic **64** (1999), 68–74.

[7] T. Altınel, G. Cherlin, L.-J. Corredor and A. Nesin, *A Hall theorem for ω-stable groups*, J. London Math. Soc. **57** (1998), 385–397.

[8] J. Ax, *The elementary theory of finite fields*, Ann. Math. **88** (1968), 239–271.

[9] J. Baldwin, *Fundamentals of Stability Theory*, Springer, Berlin (1988).

[10] J. Baldwin and K. Holland, *Constructing ω-stable structures: Fields of rank* 2, J. Symbolic Logic (to appear).

[11] J. Baldwin and J. Saxl, *Logical stability in group theory*, J. Austral. Math. Soc. **21** (1976), 267–276.

[12] J.T. Baldwin, ed., *Classification Theory: Proceedings of the US–Israel Binational Workshop on Model Theory in Mathematical Logic, Chicago, '85*, Lecture Notes in Math., Vol. 1292, Springer, Berlin (1987).

[13] J.T. Baldwin, *An almost strongly minimal non-Desarguesian projective plane*, Trans. Amer. Math. Soc. **342** (1994), 695–711.

[14] J.T. Baldwin and A. Pillay, *Semisimple stable and superstable groups*, Ann. Pure Appl. Logic **45** (1989), 105–127.

[15] J.T. Baldwin and N. Shi, *Stable generic structures*, Ann. Pure Appl. Logic **79** (1996), 1–35.

[16] A. Baudisch, *Decidability and stability of free nilpotent Lie algebras and free nilpotent p-groups of finite exponent*, Ann. Math. Logic **23** (1982), 1–25.

[17] A. Baudisch, *On superstable groups*, Seminarbericht Nr. 104, Humboldt Universität, Berlin (1989), 18–42.

[18] A. Baudisch, *On superstable groups*, J. London Math. Soc. **42** (1990), 452–464.

[19] A. Baudisch, *Another stable group*, Ann. Pure Appl. Logic **80** (1996), 109–138.

[20] A. Baudisch, *A new uncountably categorical group*, Trans. Amer. Math. Soc. **348** (1996), 3889–3940.

[21] A. Baudisch and J.S. Wilson, *Stable actions of torsion groups and stable soluble groups*, J. Algebra **153** (1991), 453–457.

[22] W. Baur, *Elimination of quantifiers for modules*, Israel J. Math. **25** (1976), 64–70.

[23] W. Baur, G. Cherlin and A. Macintyre, *Totally categorical groups and rings*, J. Algebra **57** (1979), 407–440.

[24] O.V. Belegradek, *On almost categorical theories*, Sibirsk. Mat. Zh. **14** (1973), 277–288 (in Russian).

[25] O.V. Belegradek, *On groups of finite Morley rank, Abstracts of the Eighth International Congress of Logic, Methodology and Philosophy of Science (Moscow 1987)* (1987), 100–102.

[26] O.V. Belegradek, Y. Peterzil and F.O. Wagner, *Quasi-o-minimal structures*, J. Symbolic Logic (to appear).

[27] O.V. Belegradek, F. Point and F.O. Wagner, *A quasi-o-minimal group without the exchange property*, MSRI Preprint 98-051 (1998).

[28] A. Berkman, *The classical involution theorem for tame groups of finite Morley rank*, Preprint (1998).

[29] C. Berline, *Superstable groups; a partial answer to conjectures of Cherlin and Zil'ber*, Ann. Pure Appl. Logic **30** (1986), 45–61.

[30] C. Berline and D. Lascar, *Superstable groups*, Ann. Pure Appl. Logic **30** (1986), 1–43.

[31] C. Berline and D. Lascar, *Des gros sous-groupes Abéliens pour les groupes stables*, C. R. Acad. Sci. Paris **305** (1987), 639–641.

[32] V. Bludov, *On locally nilpotent groups*, Siberian Adv. Math. **8** (1998), 49–79.

[33] M. Boffa and F. Point, *Identités de Engel généralisées*, C. R. Acad. Sci. Paris Ser. I **313** (1991), 909–911.

[34] A.V. Borovik, *Simple locally finite groups of finite Morley rank and odd type*, Finite and Locally Finite Groups, B. Hartley, G.M. Seitz, A.V. Borovik and R.M. Bryant, eds, Kluwer Academic Publishers, Dordrecht (1995), 247–284.

[35] A.V. Borovik, G. Cherlin and L.-J. Corredor, *Standard components in groups of even type*, Preprint (1997).
[36] A.V. Borovik and A. Nesin, *On the Schur–Zassenhaus theorem for groups of finite Morley rank*, J. Symbolic Logic **57** (1992), 1469–1477.
[37] A.V. Borovik and A. Nesin, *Groups of Finite Morley Rank*, Oxford University Press, Oxford, UK (1994).
[38] A.V. Borovik and A. Nesin, *Schur–Zassenhaus theorem revisited*, J. Symbolic Logic **59** (1994), 283–291.
[39] A.V. Borovik and B.P. Poizat, *Simple groups of finite Morley rank without non-nilpotent connected subgroups* (in Russian), Preprint mimeographed by VINITI, N2062-B (1990).
[40] A.V. Borovik and B.P. Poizat, *Tores et p-groupes*, J. Symbolic Logic **55** (1991), 478–491.
[41] A.V. Borovik and S.R. Thomas, *On generic normal subgroups*, Automorphisms of First-Order Structures, R. Kaye and D. Macpherson, eds, Oxford University Press, Oxford, UK (1994), 319–324.
[42] E. Bouscaren, *The group configuration – after E. Hrushovski*, in: Nesin and Pillay [127], 199–209.
[43] E. Bouscaren, ed., *Model Theory and Algebraic Geometry*, Lecture Notes in Math., Vol. 1696, Springer, Berlin (1998).
[44] R.M. Bryant, *Groups with minimal condition on centralizers*, J. Algebra **60** (1979), 371–383.
[45] R.M. Bryant and B. Hartley, *Periodic locally soluble groups with the minimal condition on centralizers*, J. Algebra **61** (1979), 328–334.
[46] S. Buechler, *The geometry of weakly minimal types*, J. Symbolic Logic **50** (1985), 1044–1053.
[47] S. Buechler, *Locally modular theories of finite rank*, Ann. Pure Appl. Logic **30** (1986), 83–94.
[48] S. Buechler, *Essential Stability Theory*, Springer, Berlin (1996).
[49] S. Buechler, *Lascar strong type in some simple theories*, Preprint (1997).
[50] S. Buechler, A. Pillay and F.O. Wagner, *Elimination of hyperimaginaries for supersimple theories*, J. Amer. Math. Soc. (submitted).
[51] C.C. Chang and J.H. Keisler, *Model Theory*, North-Holland, Amsterdam (1973).
[52] O. Chapuis, *Universal theory of certain solvable groups and bounded Ore group-rings*, J. Algebra **176** (1995), 368–391.
[53] Z. Chatzidakis, *Groups definable in ACFA*, in: Hart et al. [78], 25–52.
[54] Z. Chatzidakis and E. Hrushovski, *The model theory of difference fields*, Trans. Amer. Math. Soc. **351** (8) (1999), 2997–3071.
[55] G. Cherlin, *Superstable division rings*, Logic Colloquium '77, A. Macintyre, L. Pacholski and J. Paris, eds, North-Holland, Amsterdam (1978), 99–111.
[56] G. Cherlin, *Groups of small Morley rank*, Ann. Math. Logic **17** (1979), 1–28.
[57] G. Cherlin, L. Harrington and A. Lachlan, *$\aleph_0$-categorical $\aleph_0$-stable structures*, Ann. Pure Appl. Logic **28** (1985), 103–135.
[58] G. Cherlin and S. Shelah, *Superstable fields and groups*, Ann. Math. Logic **18** (1980), 227–270.
[59] L.J. Corredor, *Bad groups of finite Morley rank*, J. Symbolic Logic **54** (1989), 768–773.
[60] M. Davis and A. Nesin, *Suzuki 2-groups of finite Morley rank (or over quadratically closed fields)*, Comm. Algebra (submitted).
[61] M. DeBonis and A. Nesin, *BN-pairs of finite Morley rank*, Comm. Algebra (submitted).
[62] F. Delon, *Separably closed fields*, in: Bouscaren [43].
[63] J. Derakhshan and F.O. Wagner, *Nilpotency in groups with chain conditions*, Oxford Quarterly J. Math. **48** (1997), 453–468.
[64] J.-L. Duret, *Instabilité des corps formellement réels*, Canad. Math. Bull. **20** (1977), 385–387.
[65] J.-L. Duret, *Les corps pseudo-finis ont la propriété d'indépendance*, C. R. Acad. Sci. Paris **290** (1980), A981–A983.
[66] J.-L. Duret, *Sur la théorie élémentaire des corps de fonctions*, J. Symbolic Logic **51** (1986), 948–956.
[67] J. Ershov, *Fields with solvable theory*, Dokl. Akad. Nauk SSSR **174** (1967), 19–20 (in Russian).
[68] D. Evans, ed., *Model Theory of Groups and Automorphism Groups*, LMS Lecture Notes, Vol. 244, Cambridge University Press, Cambridge, UK (1997).
[69] D. Evans and E. Hrushovski, *Projective planes in algebraically closed fields*, Proc. London Math. Soc. **62** (1991), 1–24.
[70] D. Evans and F.O. Wagner, *Supersimple ω-categorical groups*, J. Symbolic Logic (to appear).
[71] U. Felgner, *$\aleph_0$-categorical stable groups*, Math. Z. **160** (1978), 27–49.
[72] O. Frécon, *Sous-groupes anormaux dans les groupes de rang de Morley fini résolubles*, Preprint (1999).

[73] J. Goode, *Hrushovski's geometries*, Proceedings of 7th Easter Conference on Model Theory, B. Dahn and H. Wolter, eds (1989), 106–118.
[74] K.W. Gruenberg, *The Engel elements of a soluble group*, Illinois J. Math. **3** (1959), 151–168.
[75] C. Grünenwald and F Haug, *On stable groups in some soluble group classes*, Proceedings of the 10th Easter Conference on Model Theory, M. Weese and H. Wolter, eds, Humboldt Universität, Berlin (1993), 46–59.
[76] C. Grünenwald and F. Haug, *On stable torsion-free nilpotent groups*, Arch. Math. Logic **32** (1993), 451–462.
[77] B. Hart, B. Kim and A. Pillay, *Coordinatization and canonical bases in simple theories*, J. Symbolic Logic (to appear).
[78] B.T. Hart, A.H. Lachlan and M.A. Valeriote, eds, *Algebraic Model Theory*, Kluwer Academic Publishers, Dordrecht (1995).
[79] W. Hodges, *Model Theory*, Cambridge University Press, Cambridge, UK (1993).
[80] E. Hrushovski, *Contributions to Stable Model Theory*, PhD thesis, University of California at Berkeley, Berkeley (1986).
[81] E. Hrushovski, *Locally modular regular types*, in: Baldwin [12], 132–164.
[82] E. Hrushovski, *Almost orthogonal regular types*, Ann. Pure Appl. Logic **45** (1989), 139–155.
[83] E. Hrushovski, *On superstable fields with automorphisms*, in: Nesin and Pillay [127], 186–191.
[84] E. Hrushovski, *Unidimensional theories are superstable*, Ann. Pure Appl. Logic **50** (1990), 117–138.
[85] E. Hrushovski, *Pseudo-finite fields and related structures*, Preprint (1991).
[86] E. Hrushovski, *Strongly minimal expansions of algebraically closed fields*, Israel J. Math. **79** (1992), 129–151.
[87] E. Hrushovski, *A new strongly minimal set*, Ann. Pure Appl. Logic **62** (1993), 147–166.
[88] E. Hrushovski, *The Mordell–Lang conjecture for function fields*, J. Amer. Math. Soc. **9** (1996).
[89] E. Hrushovski, *Simplicity and the Lascar group*, Preprint (1997).
[90] E. Hrushovski, *The Manin–Mumford conjecture and the model theory of difference fields*, Ann. Pure Appl. Logic (submitted).
[91] E. Hrushovski and A. Pillay, *Weakly normal groups*, Logic Colloquium '85, Ch. Berline, E. Bouscaren, M Dickmann, J.-L. Krivine, D. Lascar, M. Parigot, E. Pelz and G. Sabbagh, eds, North-Holland, Amsterdam (1987), 233–244.
[92] E. Hrushovski and A. Pillay, *Groups definable in local fields and pseudo-finite fields*, Israel J. Math. **85** (1994), 203–262.
[93] E. Hrushovski and A. Pillay, *Definable subgroups of algebraic groups over finite fields*, J. Reine Angew. Math. **462** (1995), 69–91.
[94] E. Hrushovski and Ž. Sokolović, *Minimal subsets of differentially closed fields*, Trans. Amer. Math. Soc. (to appear).
[95] E. Hrushovski and B. Zil'ber, *Zariski geometries*, Bull. Amer. Math. Soc. **28** (1993), 315–323.
[96] E. Hrushovski and B. Zil'ber, *Zariski geometries*, J. Amer. Math. Soc. **9** (1996), 1–56.
[97] S.V. Ivanov and A.Yu. Ol'shanskii, *Some applications of graded diagrams in combinatorial group theory*, Groups St Andrews 1989, Vol. 2, C.M. Campbell and E.F. Robertson, eds, LMS Lecture Notes, Vol. 160, Cambridge University Press, Cambridge, UK (1992), 258–308.
[98] K. Jaber, *Équations génériques dans un groupe stable nilpotent*, J. Symbolic Logic **64** (1999), 761–768.
[99] K. Jaber and F.O. Wagner, *Largeur et nilpotence*, Comm. Algebra (to appear).
[100] E. Jaligot, *Groupes de type mixte*, J. Algebra **212** (1999), 753–768.
[101] O. Kegel, *Four lectures on Sylow theory in locally finite groups*, Group Theory (Singapore, 1987), K.N. Cheng and Y.K. Lerong, eds, Walter de Gruyter, Berlin (1989), 3–27.
[102] B. Kim, *Simple First Order Theories*, PhD thesis, University of Notre Dame, Notre Dame (1996).
[103] B. Kim, *Recent results on simple first-order theories*, in: Evans [68], 202–212.
[104] B. Kim, *A note on Lascar strong types in simple theories*, J. Symbolic Logic **63** (1998), 926–936.
[105] B. Kim, *Forking in simple unstable theories*, J. London Math. Soc. **57** (1998), 257–267.
[106] B. Kim and A. Pillay, *Simple theories*, Ann. Pure Appl. Logic **88** (1997), 149–164.
[107] B. Kim and A. Pillay, *From stability to simplicity*, Bull. Symbolic Logic **4** (1998), 17–26.
[108] L. Kramer, K. Tent and H. van Maldeghem, *BN-pairs of finite Morley rank*, Israel J. Math. (to appear).

[109] D.W. Kueker and C. Laskowski, *On generic structures*, Notre Dame J. Formal Logic **33** (1992), 175–183.
[110] D. Lascar, *Les groupes ω-stables de rang fini*, Trans. Amer. Math. Soc. **292** (1985), 451–462.
[111] D. Lascar, *Stability in Model Theory*, Longman, New York (1987).
[112] J.G. Loveys, *Abelian groups with modular generic*, J. Symbolic Logic **56** (1991), 250–259.
[113] J.G. Loveys and F.O. Wagner, *Le canada semi-dry*, Proc. Amer. Math. Soc. **118** (1993), 217–221.
[114] A. Macintyre, *On ω_1-categorical theories of Abelian groups*, Fund. Math. **70** (1971), 253–270.
[115] A. Macintyre, *On ω_1-categorical theories of fields*, Fund. Math. **71** (1971), 1–25.
[116] A.I. Mal'cev, *A correspondence between groups and rings*, The Metamathematics of Algebraic Systems. Collected Papers: 1936–1967, North-Holland, Amsterdam (1971), 124–137.
[117] D. Marker, M. Messmer and A. Pillay, *Model Theory of Fields*, Lecture Notes in Logic, Vol. 5, Springer, Berlin (1996).
[118] A. Mekler, *Stability of nilpotent groups of class* 2 *and prime exponent*, J. Symbolic Logic **46** (1981), 781–788.
[119] M. Messmer, *Groups and Fields Interpretable in Separably Closed Fields*, PhD thesis, University of Illinois at Chicago, Chicago (1992).
[120] M. Morley, *Categoricity in power*, Trans. Amer. Math. Soc. **4** (1965), 514–538.
[121] A. Nesin, *Nonsolvable groups of Morley rank* 3, J. Algebra **124** (1989), 199–218.
[122] A. Nesin, *Solvable groups of finite Morley rank*, J. Algebra **121** (1989), 26–39.
[123] A. Nesin, *On sharply n-transitive superstable groups*, J. Pure Appl. Algebra **69** (1990), 73–88.
[124] A. Nesin, *On solvable groups of finite Morley rank*, Trans. Amer. Math. Soc. **321** (1990), 659–690.
[125] A. Nesin, *Generalized Fitting subgroup of a group of finite Morley rank*, J. Symbolic Logic **56** (1991), 1391–1399.
[126] A. Nesin, *Poly-separated and ω-stable nilpotent groups*, J. Symbolic Logic **56** (1991), 694–699.
[127] A. Nesin and A. Pillay, eds, *The Model Theory of Groups*, University of Notre Dame Press, Notre Dame, USA (1989).
[128] L. Newelski, *On type-definable subgroups of a stable group*, Notre Dame J. Formal Logic **32** (1991), 173–187.
[129] M. Otero, Y. Peterzil and A. Pillay, *Groups and rings definable in o-minimal expansions of real closed fields*, Bull. London Math. Soc. **28** (1996), 7–14.
[130] Y. Peterzil, A. Pillay and S. Starchenko, *Definably simple groups in o-minimal structures*, Preprint (1999).
[131] Y. Peterzil, A. Pillay and S. Starchenko, *Simple algebraic groups and semialgebraic groups over real closed fields*, Preprint (1999).
[132] Y. Peterzil and S. Starchenko, *Definable homomorphisms of Abelian groups in o-minimal structures*, Preprint (1999).
[133] Y. Peterzil and C. Steinhorn, *Definable compactness and definable subgroups of o-minimal groups*, Preprint (1999).
[134] A. Pillay, *An Introduction to Stability Theory*, Clarendon Press, Oxford, UK (1983).
[135] A. Pillay, *Simple superstable theories*, in: Baldwin [12], 247–263.
[136] A. Pillay, *Groups and fields definable in an o-minimal structure*, J. Pure Appl. Algebra **53** (1988), 239–255.
[137] A. Pillay, *ACFA and the Manin–Mumford conjecture*, in: Hart et al. [78], 195–205.
[138] A. Pillay, *The geometry of forking and groups of finite Morley rank*, J. Symbolic Logic **60** (1995), 1251–1259.
[139] A. Pillay, *Geometric Stability Theory*, Oxford University Press, Oxford, UK (1996).
[140] A. Pillay, *Some foundational questions concerning differential algebraic groups*, Pacific J. Math. **179** (1997), 179–200.
[141] A. Pillay, *Definability and definable groups in simple theories*, J. Symbolic Logic **63** (1998), 788–796.
[142] A. Pillay and B. Poizat, *Pas d'imaginaires dans l'infini*, J. Symbolic Logic **52** (1987), 400–403.
[143] A. Pillay and B. Poizat, *Corps et chirurgie*, J. Symbolic Logic **60** (1995), 528–533.
[144] A. Pillay, T. Scanlon and F.O. Wagner, *Supersimple division rings*, Math. Res. Lett. (1998), 473–483.
[145] A. Pillay and Ž. Sokolović, *A remark on differential algebraic groups*, Comm. Algebra **92** (1992), 3015–3026.
[146] A. Pillay and Ž. Sokolović, *Superstable differential fields*, J. Symbolic Logic **57** (1992), 97–108.
[147] A. Pillay and C. Steinhorn, *Definable sets in ordered structures*, Trans. Amer. Math. Soc. **295** (1986), 565–592.

[148] F. Point, *Conditions of quasi-nilpotency in certain varieties of groups*, Comm. Algebra **22** (1994), 365–390.
[149] F. Point and F.O. Wagner, *Ultimately periodic groups*, Ann. Pure Appl. Logic (to appear).
[150] B.P. Poizat, *Sous-groupes définissables d'un groupe stable*, J. Symbolic Logic **46** (1981), 137–146.
[151] B.P. Poizat, *Groupes stables, avec types génériques réguliers*, J. Symbolic Logic **48** (1983), 339–355.
[152] B.P. Poizat, *Cours de théorie des modèles*, Nur Al-Mantiq Wal-Ma'rifah, Villeurbanne (1985).
[153] B.P. Poizat, *Groupes Stables*, Nur Al-Mantiq Wal-Ma'rifah, Villeurbanne (1987).
[154] B.P. Poizat, *Équations génériques*, Proc. of the 8th Easter Conference on Model Theory, B. Dahn and H. Wolter, eds, Humboldt Universität, Berlin (1990).
[155] B.P. Poizat, *Corps de rang deux*, The 3rd Barcelona Logic Meeting, Vol. 9, Centre de Recerca Matematica, Barcelona (1997), 31–44.
[156] B.P. Poizat, *L'égalité au cube*, J. Symbolic Logic (submitted).
[157] B.P. Poizat, *Le carré de l'égalité*, J. Symbolic Logic (to appear).
[158] B.P. Poizat and F.O. Wagner, *Sous-groupes périodiques d'un groupe stable*, J. Symbolic Logic **58** (1993), 385–400.
[159] B.P. Poizat and F.O. Wagner, *Liftez les Sylows! – Une suite à "Sous-groupes périodiques d'un groupe stable"*, J. Symbolic Logic (to appear).
[160] M. Prest, *Model Theory of Modules, LMS LN 130*, Cambridge University Press, Cambridge, UK (1988).
[161] J. Reineke, *Minimale Gruppen*, Z. Math. Logik **21** (1975), 357–359.
[162] A. Seidenberg, *An elimination theory for differential algebra*, Univ. California Math. Publ. **3** (1956), 31–65.
[163] S. Shelah, *Stable theories*, Israel J. Math. **7** (1969), 187–202.
[164] S. Shelah, *Differentially closed fields*, Israel J. Math. **16** (1973), 314–328.
[165] S. Shelah, *Classification Theory*, North-Holland, Amsterdam (1978).
[166] S. Shelah, *Simple unstable theories*, Ann. Pure Appl. Logic **19** (1980), 177–203.
[167] S. Shelah, *Toward classifying unstable theories*, Ann. Pure Appl. Logic **80** (1996), 229–255.
[168] L. van den Dries, *Definable groups in characteristic* 0 *are algebraic groups*, Abstr. Amer. Math. Soc. **3** (1982), 142.
[169] L. van den Dries, *Tame Topology and o-minimal Structures*, LMS Lecture Notes, Vol. 248, Cambridge University Press, Cambridge, UK (1998).
[170] L. van den Dries, A. Macintyre and D. Marker, *The elementary theory of restricted analytic fields with exponentiation*, Ann. Math. **140** (1994), 183–205.
[171] F.O. Wagner, *Stable Groups and Generic Types*, PhD thesis, University of Oxford, Oxford, UK (1990).
[172] F.O. Wagner, *Subgroups of stable groups*, J. Symbolic Logic **55** (1990), 151–156.
[173] F.O. Wagner, *Small stable groups and generics*, J. Symbolic Logic **56** (1991), 1026–1037.
[174] F.O. Wagner, *À propos d'équations génériques*, J. Symbolic Logic **57** (1992), 548–554.
[175] F.O. Wagner, *More on* $\mathfrak{R}$, Notre Dame J. Formal Logic **33** (1992), 159–174.
[176] F.O. Wagner, *Commutator conditions and splitting automorphisms for stable groups*, Arch. Math. Logic **32** (1993), 223–228.
[177] F.O. Wagner, *Quasi-endomorphisms in small stable groups*, J. Symbolic Logic **58** (1993), 1044–1051.
[178] F.O. Wagner, *Stable groups, mostly of finite exponent*, Notre Dame J. Formal Logic **34** (1993), 183–192.
[179] F.O. Wagner, *Nilpotent complements and Carter subgroups in stable* $\mathfrak{R}$*-groups*, Arch. Math. Logic **33** (1994), 23–34.
[180] F.O. Wagner, *A note on defining groups in stable structures*, J. Symbolic Logic **59** (1994), 575–578.
[181] F.O. Wagner, *Relational structures and dimensions*, Automorphism Groups of First-order Structures, R. Kaye and D. McPherson, eds, Oxford University Press, Oxford, UK (1994), 153–180.
[182] F.O. Wagner, *The Fitting subgroup of a stable group*, J. Algebra **174** (1995), 599–609.
[183] F.O. Wagner, *Hyperstable theories*, Logic: From Foundations to Applications (European Logic Colloquium 1993), W. Hodges, M. Hyland, C. Steinhorn and J. Truss, eds, Oxford University Press, Oxford, UK (1996), 483–514.
[184] F.O. Wagner, *Stable groups (a survey)*, Algebra. Proc. of the IIIrd International Conference on Algebra, Krasnoyarsk, Russia, 1993, Yu.L. Ershov, E.I. Khukhro, V.M. Levchuk and N.D. Podufalov, eds, Walter de Gruyter, Berlin (1996), 273–285.

[185] F.O. Wagner, *On the structure of stable groups*, Ann. Pure Appl. Logic **89** (1997), 85–92.
[186] F.O. Wagner, *Stable Groups*, LMS Lecture Notes, Vol. 240, Cambridge University Press, Cambridge, UK (1997).
[187] F.O. Wagner, *CM-triviality and stable groups*, J. Symbolic Logic **63** (1998), 1473–1495.
[188] F.O. Wagner, *Small fields*, J. Symbolic Logic **63** (1998), 995–1002.
[189] F.O. Wagner, *Groups in simple theories*, J. Symbolic Logic (submitted).
[190] F.O. Wagner, *Hyperdefinable groups in simple theories*, J. Math. Logic (submitted).
[191] F.O. Wagner, *Minimal fields*, J. Symbolic Logic (to appear).
[192] F.O. Wagner, *Nilpotency in groups with the minimal condition on centralizers*, J. Algebra **217** (1999), 448–460.
[193] F.O. Wagner, *Simple Theories*, Kluwer Academic Publishers, Dordrecht (to appear).
[194] A. Weil, *On algebraic groups of transformations*, Amer. J. Math. **77** (1955), 302–271.
[195] A. Wilkie, *Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function*, J. Amer. Math. Soc. **9** (1996), 1051–1094.
[196] A. Wilkie, *A general theorem of the complement and some new o-minimal structures*, Selecta Mathematica (to appear).
[197] A. Wilkie, *Model theory of analytic and smooth functions*, Logic Colloquium '97, D. McPherson and J. Truss, eds (to appear).
[198] C. Wood, *Notes on the stability of separably closed fields*, J. Symbolic Logic **44** (1979), 412–416.
[199] M. Ziegler, *Model theory of modules*, Ann. Pure Appl. Logic **26** (1984), 149–213.
[200] B. Zil'ber, *Groups and rings whose theory is categorical*, Fund. Math. **95** (1977), 173–188 (in Russian).
[201] B. Zil'ber, *Strongly minimal countably categorical theories*, Sibirsk. Mat. Zh. **21** (1980), 98–112 (in Russian).
[202] B. Zil'ber, *Totally categorical theories; structural properties and the non-finite axiomatizability*, Model Theory of Algebra and Arithmetic, L. Pacholski, J. Wierzejewski and A. Wilkie, eds, Lecture Notes in Math., Vol. 834, Springer, Berlin (1980), 381–410.
[203] B. Zil'ber, *Some model theory of simple algebraic groups over algebraically closed fields*, Colloquium Mathematicum (1984), 173–180.
[204] B. Zil'ber, *Strongly minimal countably categorical theories II*, Sibirsk. Mat. Zh. **25** (1984), 71–88 (in Russian).
[205] B. Zil'ber, *Strongly minimal countably categorical theories III*, Sibirsk. Mat. Zh. **25** (1984), 63–77 (in Russian).
[206] B. Zil'ber, *ω-stability of the field of complex numbers with a predicate distinguishing the roots of unity*, Preprint (1994).

Section 3A
Commutative Rings and Algebras

This Page Intentionally Left Blank

Artin Approximation

Dorin Popescu

Institute of Mathematics, University of Bucharest, P.O. Box 1-764, Bucharest 70700, Romania
E-mail:dorin@stailow.imar.ro

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

Artin approximation theory has many applications in algebraic geometry (for example to the algebraization of versal deformation and constructions of algebraic spaces (see [5–7])), in algebraic number theory, and in commutative algebra concerning questions of factoriality (see [5,23]). Here we present the main results of this theory with some applications in commutative algebra (see, e.g., 2.10, 2.13). Most of the proofs are based on the so-called General Néron Desingularization. Its proof is given here only for rings containing $\mathbb{Q}$ with the idea of avoiding the difficulties of the nonseparable case; the interested reader is invited to examine the general case in a very nice exposition of Swan [38]. Another consequence of General Néron Desingularization (we omit this here since it has not too much in common with Artin approximation theory) says that a regular local ring containing a field is a filtered inductive limit of regular local rings essentially of finite type over $\mathbb{Z}$. This is a partial positive answer to a conjecture by Swan and proves the Bass–Quillen Conjecture in the equicharacteristic case using Lindel's result [24] (see also [32,38]). We require here a lot of results from the theory of Henselian rings, étale maps and André–Quillen homology which we briefly present without proofs, since the reader can find them in some excellent books, such as, e.g., [34,21,1], or in the short exposition [38].

1. Henselian rings and the property of Artin approximation

Let (A, m, k) be a local ring. A is a *Henselian* ring if it satisfies the Hensel Lemma, that is given a monic polynomial f in a variable Z over A and a factorization $\bar{f} = f \bmod m = \bar{g}\bar{h}$ of $\bar{f}$ in two monic polynomials $\bar{g}, \bar{h} \in k[Z]$ having no common factors, there exist two monic polynomials $g, h \in A[Z]$ such that $f = gh$ and $\bar{g} = g \bmod m$, $\bar{h} = h \bmod m$. If A is a Henselian local ring then every finite A-algebra is a product of Henselian local rings. For details concerning this theory see [19,34,21,22].

THEOREM 1.1 (Implicit function theorem). *Let $f = (f_1, \ldots, f_r)$ be a system of polynomials in $Y = (Y_1, \ldots, Y_n)$ over A, $r \leqslant n$, and $\Delta_f \subset A[Y]$ the ideal generated by the $r \times r$-minors of the Jacobian matrix $(\partial f_i / \partial Y_j)$. Suppose A is Henselian. If $\tilde{y} \in A^n$ satisfies $f(\tilde{y}) \equiv 0 \bmod m$ and $\Delta_f(\tilde{y}) \not\subset m$ then there exists a solution $y \in A^n$ of f in A such that $y \equiv \tilde{y} \bmod m$ (y is unique if $r = n$).*

This theorem is the algebraic version of the well known Implicit Function Theorem from Differential or Analytic Geometry:

THEOREM 1.2. *Let $g_i(X_1, \ldots, X_s, Y_1, \ldots, Y_n) = 0$, $1 \leqslant i \leqslant n$, be some analytic equations, where $g_i : \mathbb{C}^{s+n} \to \mathbb{C}$ are analytic maps in a neighborhood of $(0, 0) \in \mathbb{C}^{s+n}$ with the property that $g_i(0, 0) = 0$, $1 \leqslant i \leqslant n$, and $\det(\partial g_i / \partial Y_j(0, 0)) \neq 0$. Then there exist some unique maps $y_j : \mathbb{C}^s \to \mathbb{C}$, $1 \leqslant j \leqslant n$, which are analytic in a neighborhood of $0 \in \mathbb{C}^s$ such that $y_j(0) = 0$ and $g(X, y) = 0$ in a neighborhood of 0.*

In particular Theorem 1.2 says clearly that in the local ring A of all analytic germs of maps defined in neighborhood of $0 \in \mathbb{C}^s$, Theorem 1.1 holds. It is well known that A is Henselian if and only if for every monic polynomial $h \in A[Z]$ in a variable Z such that $h(0) \in m$ and $\partial h/\partial Z(0) \notin m$ there exists an unique solution $z \in m$ of h in A (thus Theorem 1.1 also characterizes Henselian rings). The completion $\hat{A}$ of A in the m-adic topology is Henselian. There exists also another variant of Theorem 1.1, which says that if $f(\tilde{y}) \equiv 0 \bmod m^c$ for a $c \in \mathbb{N}$ then the solution y can be taken such that $y \equiv \tilde{y} \bmod m^c$.

LEMMA 1.3. *Let (A, m, k) be a Henselian local ring, $f = (f_1, \ldots, f_r)$ a system of polynomials in $Y = (Y_1, \ldots, Y_n)$ over A, $r \leqslant n$, and $\hat{y}$ a solution of f in $\hat{A}$ such that $\Delta_f(\hat{y}) \not\subset m\hat{A}$. Then f has a solution y in A such that $y \equiv \hat{y} \bmod m\hat{A}$.*

PROOF. Choose $\tilde{y} \in A^n$ such that $\tilde{y} \equiv \hat{y} \bmod m\hat{A}$ (Note that $\hat{A}/m\hat{A} \cong A/m$!). We have $f(\tilde{y}) \equiv f(\hat{y}) = 0 \bmod m\hat{A}$ and $\Delta_f(\tilde{y}) \equiv \Delta_f(\hat{y}) \not\equiv 0 \bmod m\hat{A}$. By Theorem 1.1 there exists $y \in A^n$ such that $f(y) = 0$ and $y \equiv \tilde{y} \bmod m$. □

REMARK 1.4. Lemma 1.3 says in particular that if a system of polynomial equations f over A has a special solution $\hat{y}$ in $\hat{A}$ (that is $\Delta_f(\hat{y}) \not\equiv 0 \bmod m\hat{A}$) then it has also a solution in A.

A Noetherian local ring (A, m) has the *Artin approximation property* (in brief, A has AP) if every finite system of polynomial equations over A has a solution in A if it has one in the completion $\hat{A}$ of A. The study of the Artin approximation started with the famous papers of M. Artin [4,5], which state that the convergent power series rings over a nontrivial valued field of characteristic zero, the Henselization of a local ring essentially of finite type over a field, and an excellent Dedekind ring all have AP. Note that Remark 1.4 shows a weak form of AP. Moreover, if A has AP then A is necessarily Henselian. Indeed, if f is a polynomial in Z over A such that $f(0) \in m$ and $(\partial f/\partial Z)(0) \notin m$ then f has an unique root in $m\hat{A}$ ($\hat{A}$ is Henselian!), which must be in m since A has AP.

LEMMA 1.5. *A Noetherian local ring (A, m) has AP if and only if every finite system of polynomial equations over A in $Y = (Y_1, \ldots, Y_n)$, $n \in \mathbb{N}$, has its set of solutions in A dense with respect to the m-adic topology in the set of its solutions in the completion $\hat{A}$ of A; that is, for every solution $\hat{y}$ of f in $\hat{A}$ and every positive integer c there exists a solution y of f in A such that $y \equiv \hat{y} \bmod m^c\hat{A}$.*

PROOF. The sufficiency is trivial. If $\hat{y}$ is a solution of f in $\hat{A}$ and $c \in \mathbb{N}$, choose a system of elements $\tilde{y}$ in A such that $\tilde{y} \equiv \hat{y} \bmod m^c\hat{A}$ ($A/m^c \cong \hat{A}/m^c\hat{A}$!). We have

$$\tilde{y} - \hat{y} = \sum_{i=1}^{s} \lambda_i \hat{z}_i$$

for some elements $\lambda_i \in m^c$ and $\hat{z}_i \in \hat{A}^n$, $1 \leqslant i \leqslant s$. Thus $(\hat{y}, (\hat{z}_i)_i)$ is a solution of

$$f = 0, \qquad g := \tilde{y} - Y - \sum_{i=1}^{s} \lambda_i Z_i = 0 \quad \text{in } \hat{A}.$$

If A has AP there exists a solution $(y, (z_i)_i)$ of $f = 0$, $g = 0$ in A. We have $f(y) = 0$, $y \equiv \tilde{y} \bmod m^c$ and so $y \equiv \hat{y} \bmod m^c \hat{A}$. □

PROPOSITION 1.6. *Let (A, m) be a Noetherian local ring which has AP and f a system of polynomial equations in $Y = (Y_1, \dots, Y_n)$ which has just a finite set of solutions in m^c (possible none) for a certain positive integer c. Then f has no other solutions in $m^c \hat{A}$, $\hat{A}$ being the completion of A.*

PROOF. Let $y^{(1)}, \dots, y^{(s)}$ be the solutions of f in m^c and suppose that there exists a solution $\hat{y}$ of f in $m^c \hat{A}$ which is different from the previous ones. Then there exists $t > c$ such that $y^{(i)} \not\equiv \hat{y} \bmod m^t \hat{A}$ for $1 \leqslant i \leqslant s$, the m-adic topology of A being separate. As A has AP we find a solution y of f in A such that $y \equiv \hat{y} \bmod m^t \hat{A}$ (see Lemma 1.5). Then $y \not\equiv y^{(i)} \bmod m^t \hat{A}$, $1 \leqslant i \leqslant s$, and so y is a solution of f in m^c different from all $y^{(i)}$. Contradiction! □

COROLLARY 1.7. *Let (A, m) be a Noetherian local ring which has AP and let $\hat{A}$ be its completion. Then*

(i) *A is reduced if and only if $\hat{A}$ is reduced,*

(ii) *If A is an integral domain then it is algebraically closed in $\hat{A}$.*

PROOF. (i) A is reduced if $\hat{A}$ is reduced, because the completion map $A \to \hat{A}$ is injective. If A is reduced then the polynomial $f = Z^n$ has in A only the solution $z = 0$. By the previous proposition f cannot have nonzero solutions in $\hat{A}$,that is $\hat{A}$ is also reduced.

(ii) If f is a polynomial in Z over A then it has at most deg f roots in A. By the previous proposition f cannot have other roots in $\hat{A}$. Thus A contains all roots of f in $\hat{A}$. □

PROPOSITION 1.8. *Let (A, m) be a Noetherian local ring which has AP and B a finite local A-algebra. Then B has AP too.*

PROOF. Let $w_1, \dots, w_s$ be a system of generators of B as A-module, $\phi : A^s \to B$, the map given by $(a_1, \dots, a_s) \to \sum_{i=1}^{s} a_i w_i$ and $u_j = (u_{j1}, \dots, u_{js}) \in A^s$, $1 \leqslant j \leqslant t$, a system of generators of $\operatorname{Ker} \phi$. Let $f = (f_1, \dots, f_r)$ be a system of polynomials in $Y = (Y_1, \dots, Y_n)$ over B, $\widehat{B}$ the completion of B and $\hat{y} = (\hat{y}_1, \dots, \hat{y}_n) \in \widehat{B}^n$. Then

$$\hat{y}_j = \sum_{\lambda=1}^{s} \hat{y}_{j\lambda} w_\lambda$$

for some $\hat{y}_{j\lambda} \in \hat{A}$ ($\hat{A} \otimes_A B \cong \widehat{B}$!) and $f_i(\hat{y})$ has the form

$$\sum_{\alpha=1}^{s} f_{i\alpha}((\hat{y}_{j\lambda})) w_\alpha$$

for some polynomials $f_{i\alpha}$ in $Y_{j\alpha}$ over A. Clearly $\hat{y}$ is a solution of f in $\widehat{B}$ if and only if there exist $(\hat{z}_{i\beta})$, $1 \leqslant i \leqslant r$, $1 \leqslant \beta \leqslant t$, in $\hat{A}$ such that

$$f_{i\alpha}\big((\hat{y}_{j\lambda})\big) = \sum_{\beta=1}^{t} \hat{z}_{i\beta} u_{\beta\alpha}, \quad 1 \leqslant i \leqslant r,\ 1 \leqslant \alpha \leqslant s$$

((u_i) generate also $\operatorname{Ker}(\hat{A} \otimes_A \phi)$!). Suppose $\hat{y}$ is a solution of f in $\widehat{B}$. Then $(\hat{y}_{j\lambda})$, $(\hat{z}_{i\beta})$ is a solution in $\hat{A}$ of the polynomial equations

$$f_{i\alpha}\big((Y_{j\lambda})\big) = \sum_{\beta=1}^{t} Z_{i\beta} u_{\beta\alpha}, \quad 1 \leqslant i \leqslant r, 1 \leqslant \alpha \leqslant s,$$

in $(Y_{j\lambda})$, $(Z_{i\beta})$ over A, which must have a solution $(y_{j\lambda})$, $(z_{i\beta})$ in A because A has AP. Clearly (y_j), $y_j = \sum_{\lambda=1}^{s} y_{j\lambda} w_\lambda$, forms a solution of f in B. $\square$

PROPOSITION 1.9. *Let (A, m) be a Noetherian local ring which has AP and $\hat{A}$ its completion. Then*
- (i) *A is an integral domain if and only if $\hat{A}$ is an integral domain,*
- (ii) *If $p \in \operatorname{Spec} A$ then $p\hat{A} \in \operatorname{Spec} \hat{A}$,*
- (iii) *For every field K which is a finite A-algebra, the ring $K \otimes_A \hat{A}$ is an integral domain.*

PROOF. (i) Suppose $\hat{A}$ is not an integral domain; that is, there exist two nonzero elements $\hat{x}, \hat{y} \in \hat{A}$ such that $\hat{x}\hat{y} = 0$. Choose a positive integer c such that $\hat{x}, \hat{y} \notin m^c \hat{A}$. By Lemma 1.3 there exist $x, y \in A$ such that $xy = 0$ and $x \equiv \hat{x}$, $y \equiv \hat{y} \bmod m^c$. It follows $x \notin m^c$, $y \notin m^c$ and so A is not an integral domain.

(ii) If $p \in \operatorname{Spec} A$ then A/p is an integral domain which has AP by the previous proposition. Thus $\hat{A}/p\hat{A} \cong \widehat{(A/p)}$ is an integral domain by (i) which is enough.

(iii) We may reduce to the case $A \subset K$ by (ii) and the previous proposition. Choose a basis $b = (b_1, \ldots, b_n)$ of K over the fraction field $Q(A)$ of A. Multiplying b with a nonzero element of A, we may suppose all b_i integral over A. Thus $B = A[b]$ is a finite A-algebra and so a finite product of local rings (A is Henselian!). Since B is an integral domain, it must be local. Then B has AP and its completion $\hat{B}$ is an integral domain by (i) and the previous proposition. Hence

$$K \otimes_A \hat{A} = Q(B) \otimes_A \hat{A} \cong Q(B) \otimes_B (B \otimes_A \hat{A}) \cong Q(B) \otimes_B \widehat{B}$$

is an integral domain. $\square$

REMARK 1.10. A convergent power series in $X = (X_1, \ldots, X_s)$ over $\mathbb{C}$ is irreducible in $\mathbb{C}[[X]]$ if it is irreducible in $A := \mathbb{C}\{X\}$. This is a consequence of Proposition 1.9 (ii) because A has AP by the Artin's result from [4].

Let $u : A \to A'$ be a morphism of Noetherian rings. The fibers of u are *geometrically reduced* (resp. *geometrically regular*) if for every field K, which is a finite A-algebra, the ring $K \otimes_A A'$ is reduced (resp. regular). The map u is called *regular* if it is flat and its fibers are geometrically regular. Let A be a Noetherian local ring and $\hat{A}$ the completion of A. The *formal fibers of A are geometrically reduced* (resp. *geometrically regular*) [25, Section 32] if the fibers of the completion map $A \to \hat{A}$ are geometrically reduced (resp. geometrically regular). A Noetherian local ring is *quasi-excellent* (resp. *universally Japanese* called also *Nagata*) if its formal fibers are geometrically regular (resp. reduced). An *excellent* local ring is an universally catenary, quasi-excellent local ring. A Henselian local ring is excellent if and only if it is quasi-excellent. For example Noetherian complete local rings and the convergent power series ring $\mathbb{C}\{X\}$ from 1.10 are all excellent (see [19]).

PROPOSITION 1.11. *A Noetherian local ring which has AP is Henselian and universally Japanese.*

PROOF. Let A be a Noetherian local ring which has AP. Note that the formal fibers of A are integral domains by Proposition 1.9(iii). In particular A is universally Japanese. $\square$

THEOREM 1.12 ([19,21,22]). *Let A be a Henselian local reduced ring which is universally Japanese, $\hat{A}$ its completion and $Q(A)$ the total fraction ring of A. Suppose that $Q(A) \otimes_A \hat{A}$ is normal. Then A is algebraically closed in $\hat{A}$.*

REMARK 1.13. This theorem says in particular that if A is an excellent Henselian local reduced ring then every polynomial equation over A in one variable which has a solution in $\hat{A}$ has also one (the same) in A.

This suggests the following result conjectured by M. Artin [6].

THEOREM 1.14 ([31,38]). *An excellent Henselian local ring has AP.*

The proof goes like in Lemma 1.3 using the following theorem, its proof being given in Sections 4–7 only for rings containing the rational numbers $\mathbb{Q}$ (for the complete proof see [30,31,33]; [3,27,38,37]):

THEOREM 1.15 (General Néron desingularization). *Let $u : A \to A'$ be a regular morphism of Noetherian rings, B a finite type A-algebra and $v : B \to A'$ a morphism of A-algebras. Then v factorises through an A-algebra $C = (A[Y]/(f))_g$, $Y = (Y_1, \dots, Y_n)$, where $f = (f_1, \dots, f_r)$, $r \leqslant n$, are polynomials in Y over A and g belongs to the ideal Δ_f generated by the $r \times r$-minors of the Jacobian matrix $(\partial f_i / \partial Y_j)$.*

Indeed, let (A, m) be an excellent Henselian local ring, h a system of polynomial equations in $Z = (Z_1, \dots, Z_s)$ over A and $\hat{z}$ a solution of h in the completion $\hat{A}$ of A. Then the A-morphism $v : B := A[Z]/(h) \to \hat{A}$, $Z \to \hat{z}$ factors through an A-algebra C as in Theorem 1.15 ($A \to \hat{A}$ is regular since A is excellent!), let us say $v = wq$, $w : C \to \hat{A}$, $q : B \to C$.

Now note that $\hat{y} = w(\hat{Y})$ is a solution of f such that $g(\hat{y}) = w(\hat{g}) \notin m\hat{A}$. Thus $\Delta_f(\hat{y}) \not\subset m\hat{A}$ and by Lemma 1.3 there exists a solution y of f in A such that $y \equiv \hat{y} \bmod m\hat{A}$. Then $g(y) \equiv g(\hat{y}) \not\equiv 0 \bmod m\hat{A}$ and so we get an A-morphism $u : C \to A$ by $Y \to y$. Clearly $z = uq(\hat{Z})$ is a solution of h in A.

REMARK 1.16. Artin's ideas to prove AP in [4,5] use mainly the Weierstrass Preparation Theorem and this is possible in the most important cases. These ideas do not work for the ring $\mathbb{C}\{X\}[[Z]]$, $X = (X_1, \ldots, X_s)$, $Z = (Z_1, \ldots, Z_t)$, which is excellent Henselian and so it has AP by Theorem 1.14.

2. Ultraproducts and the strong approximation property

Let D be a *filter* on $\mathbb{N}$, that is a family of subsets of $\mathbb{N}$ satisfying

(i) $\emptyset \notin D$,

(ii) if $s, t \in D$ then $s \cap t \in D$,

(iii) if $s \in D$ and $s \subset t \subset \mathbb{N}$ then $t \in D$.

An *ultrafilter* on $\mathbb{N}$ is a maximal in the set of filters on $\mathbb{N}$ with respect to the inclusion. A filter D on $\mathbb{N}$ is an ultrafilter if and only if $\mathbb{N}\backslash s \in D$ for each subset $s \subset \mathbb{N}$ which is not in D. An ultrafilter is nonprincipal if there exists no $r \in \mathbb{N}$ such that $D = \{s \mid r \in s \subset \mathbb{N}\}$. An ultrafilter on $\mathbb{N}$ is nonprincipal if and only if it contains the filter of all cofinite subsets of $\mathbb{N}$.

The *ultrapower* A^* of a ring A with respect to a nonprincipal ultrafilter D on $\mathbb{N}$ is the quotient ring of $A^{\mathbb{N}}$ by the ideal I_D of all $(x_n)_{n\in\mathbb{N}}$ such that the set $\{n \in \mathbb{N} \mid x_n = 0\} \in D$. Denote by $[(x_n)]$ the class modulo I_D of (x_n). Assigning to $a \in A$ the constant sequence $[(a, a, \ldots)]$ we get a ring morphism $\phi_A : A \to A^*$. Similarly we may speak about the ultrapower M^* of an A-module M which has also a structure of A^*-module. As above we have a canonical map $\phi_M : M \to M^*$. For details concerning this theory see [10,8,29, 31].

PROPOSITION 2.1. *Let A be a ring, D a nonprincipal ultrafilter on $\mathbb{N}$ and A^* the ultrapower of A with respect to D. Then*

(i) *A^* is an integral domain (resp. field) if A is an integral domain (resp. field),*

(ii) *If A is a domain, then $Q(A)^* \cong Q(A^*)$, where $Q(A)$, $Q(A^*)$ are the fraction fields of A, resp. A^*,*

(iii) *$(A^*)^q \cong (A^q)^*$, $q \in \mathbb{N}$,*

(iv) *If A is a field then A^* is a separable field extension of A,*

(v) *If $a \subset A$ is an ideal and a^* is the ultrapower of a with respect to D, then $a^* \subset A^*$ is an ideal and $(A/a)^* \cong A^*/a^*$,*

(vi) *If $p \subset A$ is a prime (resp. maximal) ideal then $p^* \subset A^*$ is a prime (resp. maximal) ideal,*

(vii) *If $a \subset A$ is a finitely generated ideal, then $a^* = \phi_A(a)A^*$.*

PROOF. (i) If $[(x_n)][(y_n)] = 0$ in A^* then $r = \{n \in \mathbb{N} \mid x_n y_n = 0\} \in D$. If A is an integral domain then $r = s \cup t$ for $s = \{n \in \mathbb{N} \mid x_n = 0\}$, $t = \{n \in \mathbb{N} \mid y_n = 0\}$. As D is an ultrafilter it follows either $s \in D$, or $t \in D$ ($s \cup t \in D$!) and so either $[(x_n)] = 0$, or $[(y_n)] = 0$. Now suppose that A is a field and let $0 \neq x = [(x_n)] \in A^*$. Then $c = \{n \in \mathbb{N} \mid x_n \neq 0\} \in D$ and $y_n = x_n^{-1}$ for $n \in c$ and 0 otherwise define an inverse $[(y_n)]$ of x.

(ii) If A is an integral domain then the inclusion $A \subset Q(A)$ induces the inclusion $A^* \subset Q(A)^*$. By (i) $Q(A)^*$ is a field. If $z = [(z_n)] \in Q(A)^*$ then $z_n = u_n/v_n$, $v_n \neq 0$, $u_n, v_n \in A$, and clearly $z = [(u_n)]/[(v_n)] \in Q(A^*)$. Hence $Q(A)^* = Q(A^*)$.

(iii) Let $(e_i)_{1 \leqslant i \leqslant q}$ be the canonical basis in A^q. We must show that $(\phi_{A^q}(e_i))_{1 \leqslant i \leqslant q}$ is a basis in $(A^q)^*$ over A^*. If $u = [(u_n)] \in (A^q)^*$ then we have

$$u_n = \sum_{i=1}^{q} u_{ni} e_i, \quad u_{ni} \in A.$$

Hence

$$u = \sum_{i=1}^{q} [(u_{ni})] \phi_{A^q}(e_i)$$

that is $(\phi_{A^q}(e_i))_i$ generate $(A^q)^*$. Now, if

$$\sum_{i=1}^{q} v_i \phi_{A^q}(e_i) = 0$$

for some $v_i = [(v_{ni})] \in A^*$ then

$$s = \left\{ n \in \mathbb{N} \mid \sum_{i=1}^{q} v_{ni} e_i = 0 \right\} \in D.$$

Thus $v_{ni} = 0$ for $n \in s$, $1 \leqslant i \leqslant q$, and so $v_i = 0$. Hence $(\phi_{A^q}(e_i))_i$ is a basis.

(iv) If A is a field then A^* is a field by (i). Let K be a finite field extension of A. Then $K \cong A^q$ for a certain q and we have $K^* \cong (A^q)^* \cong (A^*)^q \cong A^* \otimes_A K$ by (iii) as A^*-linear spaces. In particular $\dim_{A^*} K^* = \dim_{A^*} A^* \otimes_A K$. Hence the canonical ring surjection $A^* \otimes_A K \to K^*$ is an isomorphism.

(v) The surjection $A \to A/a$ induces the surjection $A^* \to (A/a)^*$ whose kernel is given by all $x = [(x_n)] \in A^*$ such that $s = \{n \in \mathbb{N} \mid x_n \in a\} \in D$. Put $y_n = x_n$ if $n \in s$ and $y_n = 0$ if $n \notin s$. Clearly $x = [(y_n)] \in a^*$.

(vi) follows from (i) and (v).

(vii) A surjection $A^q \to a$ induces a surjection $(A^q)^* \to a^*$ and it is enough to apply (iii). □

REMARK 2.2. It is easy to see that the ultrapower construction defines an exact functor $\operatorname{Mod} A \to \operatorname{Mod} A^*$, which coincides with $A^* \otimes_A -$ on finitely generated A-modules when A is Noetherian (this is already suggested by Proposition 2.1 (iii), (v), (vii)). In particular ϕ_A is flat when A is Noetherian.

PROPOSITION 2.3. *Let A be a ring, D a nonprincipal ultrafilter on $\mathbb{N}$ and A^* the ultrapower of A with respect to D. Then*

(i) *The Jacobson radical $J(A)$ of A is mapped by ϕ_A in $J(A^*)$.*

(ii) *If $m_1, \ldots, m_s$ are all maximal ideals of A then $m_1^*, \ldots, m_s^*$ are all maximal ideals of A^*. In particular A^* is local if A is local.*

(iii) *If A is an Artinian local ring then A^* is also Artinian and* $\operatorname{length}_A A = \operatorname{length}_{A^*} A^*$.

(iv) *If A is Henselian local ring then A^* is also Henselian local ring.*

PROOF. If $u \in J(A)$ and $x = [(x_n)] \in A^*$ then $1 + ux = [(1 + ux_n)]$ is invertible in A^* because $1 + ux_n$ is invertible in A.

(ii) By Proposition 1.1(v), (vii) we see that $A^*/J(A)^* \cong (A/J(A))^*$ is a product of fields $(A/m_i)^* \cong A^*/m_i^*$. Using (i) it follows that (m_i^*) are all maximal ideals of A^*.

(iii) If (A, m, k) is Artinian local then $m^s = 0$ for some $s \in \mathbb{N}$. Since m is finitely generated, $mA^* = m^*$ – the unique maximal ideal of A^* by (ii) is finitely generated and $(m^*)^s = 0$. If $p \in \operatorname{Spec} A^*$ then $(m^*)^s \subset p$ and so $m^* \subset p$. Thus m^* is the unique prime ideal of A^*. By the Cohen Theorem A^* is Noetherian and so even Artinian. If $0 = a_0 \subset a_1 \subset \cdots \subset a_t = A$ is a composition series of A then $0 = a_0^* \subset a_1^* \subset \cdots \subset a_t^* = A^*$ is a composition series of A^* because $k^* \cong (a_{i+1}/a_i)^* \cong a_{i+1}^*/a_i^*$ (as in 2.1(v), or 2.2), k^* being the residue field of A^* by Proposition 2.1(i) and (v). Thus $\operatorname{length}_A A = \operatorname{length}_{A^*} A^*$.

(iv) Suppose that (A, m) is a Henselian local ring. Let

$$f = \sum_{i=0}^{d} [(u_{ni})] Z^i$$

be a polynomial over A^* in a variable Z such that $f(0) \in m^*$ and $\partial f/\partial Z(0) \notin m^*$ and put

$$f_n = \sum_{i=0}^{d} u_{ni} Z^i \in A[Z].$$

Then $t = \{n \in \mathbb{N} \mid f_n(0) \in m, (\partial f_n/\partial Z)(0) \notin m\} \in D$. If $n \in t$ there exists $z_n \in m$ such that $f_n(z_n) = 0$, A being Henselian. Put $z_n = 0$ if $n \notin t$. Clearly $z = [(z_n)]$ is a solution of f in m^*. □

LEMMA 2.4. *Let (A, m) be a Noetherian local ring, $f = (f_e)_{e \in \mathbb{N}}$ a countable system of polynomials over A, D a nonprincipal ultrafilter on $\mathbb{N}$ and A^* the ultrapower of A with respect to D. The following statements are equivalent:*

(i) *f has a solution in $A_1 = A^*/m_\infty^*$, $m_\infty^* := \bigcap_{i \in \mathbb{N}} m^i A^*$,*

(ii) *for every $t \in \mathbb{N}$, $f^{(t)} = (f_1, \ldots, f_t)$ has a solution in $A^*/m^t A^*$.*

PROOF. We have to show only (ii) $\Rightarrow$ (i). For every $t \in \mathbb{N}$, let $\tilde{y}^{(t)} = [(\tilde{y}_n^{(t)})]$ be in A^* such that $f^{(t)}(\tilde{y}^{(t)}) \equiv 0 \bmod m^t A^*$. Then $s_t = \{n \in \mathbb{N} \mid f^{(t)}(\tilde{y}_n^{(t)}) \equiv 0 \bmod m^t\} \in D$. Note that

$$s'_t = \Bigl(\bigcap_{1 \leqslant r \leqslant t} s_r\Bigr) - \{0, \ldots, t\} \in D,$$

D being nonprincipal. We have $s'_1 \supset s'_2 \supset \cdots \supset s'_t \supset \cdots$ and $\bigcap_{r\in\mathbb{N}} s'_r = \emptyset$. Put $y_n = (\tilde{y}_n^{(t_n)})$ if $n \in s'_1$ and t_n is maximum such that $n \in s'_{t_n}$ and $y_n = 0$ if $n \notin s'_1$. Clearly $y = [(y_n)]$ satisfies $f(y) \equiv 0 \bmod m^*_\infty$. □

THEOREM 2.5 ([29,31]). *Let (A, m) be a Noetherian local ring, D a nonprincipal ultrafilter on $\mathbb{N}$, A^* the ultrapower of A with respect to D, and ψ_A the composite map*

$$A \to A^* \to A_1 := A^*/m^*_\infty, \quad m^*_\infty = \bigcap_{i\in\mathbb{N}} m^i A^*.$$

Then (A_1, mA_1) is a Noetherian complete local ring, $\dim A_1 = \dim A$ *and ψ_A is flat.*

PROOF. Let

$$B = \varprojlim A^*/m^i A^*$$

be the completion of A^* and $\rho : A^* \to B$ the canonical map. We claim that ρ is surjective. Indeed, if $z_i \in A^*$, $i \in \mathbb{N}$, has the property that $z_{i+1} \equiv z_i \bmod m^i A^*$ for all $i \in \mathbb{N}$, then by Lemma 2.4 the system of congruences $Z - z_i \equiv 0 \bmod m^i A^*$, $i \in \mathbb{N}$, has a solution $z \in A^*$ because it has solutions in $A^*/m^t A^*$ for all $t \in \mathbb{N}$. Clearly, ρ maps z on the system $(z_i)_i$. As $\operatorname{Ker} \rho = m^*_\infty$ we see that A_1 is complete local. A_1 is Noetherian because its maximal ideal mA_1 is finitely generated (see [25, 29.4]).

Now A/m^s is Artinian and by Proposition 2.3(iii) we have $A_1/m^s A_1 \cong A^*/m^s A^*$ Artinian and $\operatorname{length}_A A/m^s = \operatorname{length}_{A^*} A^*/m^s A^* = \operatorname{length}_{A_1} A_1/m^s A_1$. Thus the Hilbert–Samuel functions associated to A, resp. A_1, coincide. Hence $\dim A = \dim A_1$. As m generates the maximal ideal of A_1 we see that ψ_A must be flat by [25, Section 23]. □

COROLLARY 2.6. *In the notations and hypothesis of Theorem* 2.5 *suppose that A is quasi-excellent. Then ψ_A is regular.*

PROOF. Note that ψ_A is a flat morphism of Noetherian local rings by Theorem 2.5 and induces a separable residue field extension $k := A/m \to k^*$ by Proposition 2.1(iv). Since mA_1 is the maximal ideal of A_1 it is enough to apply the following Lemma, which is proved in Section 4. □

LEMMA 2.7 ([2,11,25]). *Let $\phi : A \to B$ be a flat morphism of Noetherian local rings. Suppose that*

(i) *A is quasi-excellent,*
(ii) *the residue field extension induced by ϕ is separable,*
(iii) *the maximal ideal of A generates the maximal ideal of B.*

Then ϕ is regular.

Let (A, m) be a Noetherian local ring. A has the *strong Artin approximation property* (in brief, A has SAP) if for every finite system of equations f in $Y = (Y_1, \ldots, Y_s)$ over A there exists a map $\nu : \mathbb{N} \to \mathbb{N}$ with the following property:

If $\tilde{y} \in A^s$ satisfies $f(\tilde{y}) \equiv 0 \bmod m^{\nu(c)}$, $c \in \mathbb{N}$, then there exists a solution $y \in A^s$ of f with $y \equiv \tilde{y} \bmod m^c$.

M. Greenberg [18] proved that excellent Henselian discrete valuation rings have SAP and M. Artin [5] showed that the Henselization (see Section 3) of a local ring which is essentially of finite type over a field has SAP. If (A, m) is a Noetherian local ring which has SAP then it has also AP. Indeed, let f be a finite system of polynomials over A in some variables Y, $\hat{y}$ a solution of f in the completion $\hat{A}$ of A and ν the SAP function associated to f. Choose $\tilde{y}$ in A such that $\tilde{y} \equiv \hat{y} \bmod m^{\nu(1)}\hat{A}$. We have $f(\tilde{y}) \equiv f(\hat{y}) = 0 \bmod m^{\nu(1)}\hat{A}$. Thus $f(\tilde{y}) \equiv 0 \bmod m^{\nu(1)}$ and so there exists a solution y of f in A such that $y \equiv \tilde{y} \bmod m$.

The following lemma shows that the SAP is more easy handled in the framework of ultrapowers.

LEMMA 2.8. *Let (A, m) be a Noetherian local ring, D a nonprincipal ultrafilter on $\mathbb{N}$, A^* the ultrapower of A with respect to D, and ψ_A the composite map*

$$A \to A^* \to A_1 := A^*/m^*_\infty, \quad m^*_\infty = \bigcap_{i \in \mathbb{N}} m^i A^*.$$

The following statements are equivalent:

(i) *A has SAP.*

(ii) *For every finite system of polynomials f over A, for every positive integer c and for every solution $\tilde{y}$ of f in A^* modulo m^*_∞ there exists a solution y of f in A^* such that $y \equiv \tilde{y} \bmod m^c A^*$.*

PROOF. (i) $\Rightarrow$ (ii) Let f, $\tilde{y} = [(\tilde{y}_n)]$ be like in (ii) and $\nu : \mathbb{N} \to \mathbb{N}$ the SAP function associated to f. We have in particular $f(\tilde{y}) \equiv 0 \bmod m^{\nu(c)} A^*$. Thus, the set $s = \{n \in \mathbb{N} \mid f(\tilde{y}_n) \equiv 0 \bmod m^{\nu(c)}\} \in D$. Then for every $n \in s$ there exists a solution y_n of f in A such that $y_n \equiv \tilde{y}_n \bmod m^c$, ν being the SAP function of f. Put $y_n = 0$ for $n \notin s$. Then $y = [(y_n)]$ is a solution of f in A^* such that $y \equiv \tilde{y} \bmod m^c A^*$.

(ii) $\Rightarrow$ (i) Suppose that there exists a finite system of polynomials f in some variables Y over A which has no SAP function; that is there exists a positive integer c such that

$(*)$ For every $n \in \mathbb{N}$ there exists $\tilde{y}_n$ in A such that $f(\tilde{y}_n) \equiv 0 \bmod m^n$ but there exists no solution y'_n of f in A such that $y'_n \equiv \tilde{y}_n \bmod m^c$.

Clearly, $\tilde{y} = [(\tilde{y}_n)]$ is a solution of f in A^* modulo $m^r A^*$ for all $r \in \mathbb{N}$. Thus $f(\tilde{y}) \equiv 0 \bmod m^*_\infty$. By (ii) there exists a solution $y = [(y_n)]$ of f in A^* such that $y \equiv \tilde{y} \bmod m^c A^*$. Then the set $s = \{n \in \mathbb{N} \mid f(y_n) = 0, y_n \equiv \tilde{y}_n \bmod m^c A^*\} \in D$ is nonempty. Take $n \in s$. Clearly, y_n contradicts $(*)$.

Part of the following theorem appeared in [28], extending some results from [40,41]. But the proof there has a gap in the nonseparable case, which was repaired in [23, Ch. 2]. Using ultrapowers easier proofs were given in [29] and [14]. The easiest one is given in [31, (4.5)] and it is presented here. □

THEOREM 2.9. *An excellent Henselian local ring has SAP. In particular a Noetherian local ring has AP if and only if it has SAP.*

PROOF. We try to imitate the proof of Theorem 1.14. Let (A, m) be an excellent Henselian local ring, D a nonprincipal ultrafilter on $\mathbb{N}$, A^* the ultrapower of A with respect to D, and ψ_A the composite map $A \to A^* \to A_1 := A^*/m^*_\infty$. By Lemma 2.8 it is enough to show that given a system of polynomials h in $Z = (Z_1, \ldots, Z_s)$ over A, a positive integer $c \in \mathbb{N}$ and $\tilde{z}$ a solution of h in A^* modulo m^*_∞ there exists a solution z of h in A^* such that $z \equiv \tilde{z} \bmod m^c A^*$. Then the A-morphism $v : B := A[Z]/(h) \to A_1$, $Z \to \tilde{z} \bmod m^*_\infty$, factorises through an A-algebra C as in Theorem 1.15 (ψ_A is regular by Corollary 2.6!), let us say $v = wq$, $w : C \to A_1$, $q : B \to C$. Thus, we have $C = (A[Y]/(f))_g$, $Y = (Y_1, \ldots, Y_n)$, $f = (f_1, \ldots, f_r)$, $r \leqslant n$, $g \in \Delta_f$ and $\hat{y} = w(\hat{Y})$ is a solution of f in A_1 such that $g(\hat{y}) = w(\hat{g}) \notin mA_1$. Hence $\Delta_f(\hat{y}) \not\subset mA_1$. Let $\tilde{y}$ be a lifting of $\hat{y}$ to A^*. In particular we have $f(\tilde{y}) \equiv 0 \bmod m^c A^*$, $\Delta_f(\tilde{y}) \not\subset mA^*$ and by a variant of the Implicit Function Theorem (A^* is Henselian by Proposition 2.3(iv)!) there exists a solution y of f in A^* such that $y \equiv \tilde{y} \bmod m^c A^*$. Then $g(y) \equiv g(\tilde{y}) \not\equiv 0 \bmod mA^*$ and so we get an A-morphism $u : C \to A^*$ by $Y \to y$. Clearly $z := uq(\hat{Z})$ is a solution of h in A^* such that $z \equiv wq(\hat{Z}) = v(\hat{Z}) = \tilde{z} \bmod m^c A^*$.

Now, if (A, m) is a Noetherian local ring, which has AP, then given a finite system of polynomials f over A, let ν be the SAP function associated to f considered over the completion $\hat{A}$ of A (the complete local rings are excellent Henselian and so they have SAP as above!). We claim that the same function ν works for f over A. Indeed, if $f(\tilde{y}) \equiv 0 \bmod m^{\nu(c)}$ for some elements $\tilde{y}$ of A and a $c \in \mathbb{N}$ then there exists a solution $\hat{y}$ of f in $\hat{A}$ such that $\hat{y} \equiv \tilde{y} \bmod m^c \hat{A}$. By Lemma 1.5 f has also a solution y in A such that $y \equiv \hat{y} \bmod m^c \hat{A}$. It follows $y \equiv \tilde{y} \bmod m^c \hat{A}$ and so $y \equiv \tilde{y} \bmod m^c$. □

PROPOSITION 2.10 ([23]). *Let (A, m) be an excellent Henselian local ring, which is an integral domain and $(x_n)_{n \in \mathbb{N}}$ a sequence of elements, which converges in the m-adic topology to an element $x \in A$. If x is irreducible in A then there exists a positive integer $t \gg 0$ such that x_n is irreducible for all $n \geqslant t$.*

PROOF. A has SAP by Theorem 2.9. Let ν be the SAP function associated to the polynomial $f := YZ - x$ over A. Then $t = \nu(1)$ works. Indeed, if x_n is reducible for a certain $n \geqslant t$ then there exist $\tilde{y}$, $\tilde{z} \in m$ such that $\tilde{y}\tilde{z} = x_n \equiv x \bmod m^t$. In particular, $f(\tilde{y}, \tilde{z}) \equiv 0 \bmod m^{\nu(1)}$ and so there exist y, $z \in A$ such that $f(y, z) = 0$ and $y \equiv \tilde{y}$, $z \equiv \tilde{z} \bmod m$. Thus $x = yz$ and $y, z \in m$ which is impossible since x is irreducible. □

LEMMA 2.11. *Let (A, m) be a Noetherian local ring which has AP, $\hat{A}$ its completion, f a system of polynomials over A in $Y = (Y_1, \ldots, Y_n)$ and $g_1, \ldots, g_r$ some systems of polynomials in Y and $Z = (Z_1, \ldots, Z_s)$. Then the following statements are equivalent:*

(i) *There exists a solution $\hat{y}$ of f in $\hat{A}$ such that all systems $g_i(\hat{y}, Z) = 0$, $1 \leqslant i \leqslant r$, have no solutions in $\hat{A}$,*

(ii) *There exists a solution y of f in A such that all systems $g_i(y, Z) = 0$, $1 \leqslant i \leqslant r$, have no solutions in A.*

PROOF. (i) $\Rightarrow$ (ii) Suppose there exists $\hat{y}$ in $\hat{A}$ such that $f(\hat{y}) = 0$ and $g_i(\hat{y}, Z) = 0$ has no solutions in $\hat{A}$ for all $1 \leqslant i \leqslant r$. Let ν_i be the SAP functions associated to $g_i(\hat{y}, Z)$, $1 \leqslant i \leqslant r$ ($\hat{A}$ has SAP by Theorem 2.9). Then $g_i(\hat{y}, Z)$ has no solutions in $A/m^c \cong \hat{A}/m^c\hat{A}$

for $c = \max(\nu_i(1))_i$. As A has AP we find a solution y of f in A such that $y \equiv \hat{y} \bmod m^c \hat{A}$ (see Lemma 1.5). Then $g_i(y, Z)$ has no solutions in A/m^c for all i and so they cannot have solutions in A.

(ii) $\Rightarrow$ (i) Suppose there exists y in A such that $f(y) = 0$ and $g_i(y, Z)$ has no solutions in A for all $1 \leqslant i \leqslant r$. Let ν_i' be the SAP functions associated to $g_i(y, Z)$, $1 \leqslant i \leqslant r$ (A has also SAP by Theorem 2.9!). Then $g_i(y, Z)$, $1 \leqslant i \leqslant r$ have no solutions in $A/m^t \cong \hat{A}/m^t\hat{A}$ for $t = \max(\nu_i'(1))_i$ and so they cannot have solutions in $\hat{A}$. □

LEMMA 2.12. *Let (A, m) be a Noetherian local integral domain, $a_1, \ldots, a_q$ a system of generators of m, $f = X_1X_2 - X_3X_4$, $g_1 = X_1Z_1 - X_3$, $g_2 = X_1Z_2 - X_4$ and $g_3 = \sum_{i,j=1}^{q} a_i a_j T_i V_j - x_1$ polynomials in some variables X, Z, T, V. Then A is not factorial if and only if there exists a solution $x = (x_1, x_2, x_3, x_4)$ of f such that each of the three polynomials $g_1(x, Z)$, $g_2(x, Z)$, $g_3(x, T, V)$ has no solutions in A.*

PROOF. A is not factorial if and only if there exists an irreducible element $x_1 \in A$, which is not prime; that is, there exists x_2, x_3, x_4 such that $x_1x_2 = x_3x_4$ and $x_3 \notin x_1A$, $x_4 \notin x_1A$ in other words x is a solution of f such that each polynomial $g_1(x, Z)$, $g_2(x, Z)$ has no solutions in A. Since x_1 is also irreducible, it is not a product of two elements from m, that is $g_3(x, T, V)$ has no solutions in A. Conversely, let $x = (x_1, x_2, x_3, x_4)$ be a solution of f such that each polynomial $g_1(x, Z)$, $g_2(x, Z)$, $g_3(x, T, V)$ has no solutions in A. Then $x_1 \neq 0$ (otherwise $T = V = 0$ is a solution of $g_3(x, T, V)$) and $x_1 \in m$ because otherwise $g_1(x, Z)$ has clearly a solution in A. As above note that x_1 is irreducible but not prime. □

THEOREM 2.13 ([23,8,31]). *Let (A, m) be an excellent Henselian local ring and $\hat{A}$ its completion. Then A is factorial if and only if $\hat{A}$ is factorial.*

The proof follows using Lemmas 2.11, 2.12, Proposition 1.9 and Theorem 1.14.

THEOREM 2.14 ([12,13,8,31]). *Let (A, m) be a Noetherian local ring, which has AP and $\hat{A}$ its completion. Then*

(i) *A is a normal domain if and only if $\hat{A}$ is a normal domain,*
(ii) *The formal fibers of A are geometrically normal domains (that is, for every field K which is a finite A-algebra, the ring $K \otimes_A \hat{A}$ is a normal domain).*

PROOF. If A is an integral domain, which is not normal then there exist $n \in \mathbb{N}$, $x_1, x_2, u_1, \ldots, u_n \in A$ with $x_2 \neq 0$ such that $x_1 \notin x_2A$ and $x_1^n + \sum_{i=1}^{n} u_i x_1^{n-i} x_2^i = 0$. Applying Lemma 2.11 for

$$f = X_1^n + \sum_{i=1}^{n} U_i X_1^{n-i} X_2^i, \qquad g_1 = X_2, \qquad g_2 = X_1 - X_2 Z$$

we get (i) ($g_1(x)$ has no solutions in A means that $x_2 \neq 0$!).

(ii) Let K be a field, which is a finite A-algebra and $\tilde{A}$ the integral closure of A in K. $\tilde{A}$ is finite over A because A is universally Japanese by Proposition 1.11 (see [25,33]).

Then $\tilde{A}$ is a product of Henselian local rings. Moreover $\tilde{A}$ is local because it is an integral domain and it has AP by Proposition 1.8. Then the completion of $\tilde{A}$ is $\tilde{A} \otimes_A \hat{A}$ and it must be a normal domain by (i). Thus $K \otimes_A \hat{A}$ is a normal domain because it is a fraction ring of $\tilde{A} \otimes_A \hat{A}$. □

REMARK 2.15. C. Rotthaus [35] proved that the formal fibers of a Noetherian local ring, which has AP are even geometrically regular. Thus the converse of Theorem 1.14 holds. On the other hand, we may speak about approximation properties for couples. The variant of Theorem 1.14 for couples goes through (see [31]), but the variant of Theorem 2.9 for couples does not hold as Spivakovsky noticed in [36].

3. Étale maps and approximation in nested subrings

A ring morphism $f : A \to B$ is called *quasi-smooth* after [38] if for any R-algebra C and ideal $I \subset C$ with $I^2 = 0$ any A-algebra morphism $B \to C/I$ lifts to an R algebra morphism $B \to C$. If such a lifting is unique then f is called *quasi-étale*. We call f *smooth* (resp. *étale*) if it is finitely presented and quasi-smooth (resp. quasi-étale). If f is essentially finitely presented (i.e. a localization of a finitely presented morphism) and quasi-smooth (resp. quasi-étale) then f is *essentially smooth* (resp. *essentially étale*).

A separable field extension and a polynomial extension $A \to A[X_1, \dots, X_n]$ are smooth. A localization $A \to S^{-1}A$ is essentially étale, S being a multiplicative system from A. Let $C = (A[Y]/(f))_g$ be the A-algebra from Theorem 1.15, $Y = (Y_1, \dots, Y_n)$, $f = (f_1, \dots, f_r)$. Then C is a smooth A-algebra. If $r = n$ then C is even étale over A. Composition of smooth (resp. étale) morphisms are still smooth (resp. étale) and tensoring a smooth (resp. étale) morphism $A \to B$ by an A-algebra we get still a smooth (resp. étale) morphism. Details can be found in [21,34,19].

THEOREM 3.1 (Grothendieck [20]). *If B is a local algebra essentially smooth over A, then $B = (A[Y, T]/(f))_P$, $Y = (Y_1, \dots, Y_n)$, where f is a monic polynomial in T over $A[Y]$, $P \subset A[Y, T]$ is a prime ideal containing f and $\partial f/\partial T \notin P$. Moreover if B is essentially étale over A then $n = 0$ above.*

It follows from this theorem that smooth algebras are flat. Moreover it is easy to see that smooth algebras over fields are regular rings and so smooth maps are regular morphisms. An étale algebra over a field k is a product of finite separable field extensions of k.

Let (A, m, k) be a local ring. An essentially étale A-algebra B must have mB as maximal ideal because B/mB is an essentially étale local k-algebra, that is a finite separable field extension of k. Hence mB is maximal. By [19,34] an essentially of finite presentation local A-algebra (B, b) is essentially étale if and only if it is flat, $b = mB$ and B/b is a finite separable extension of k. If (B, b) is an essentially of finite type flat-local A-algebra such that $b = mB$ and B/b is a separable extension of k then B is essentially smooth over A, providing A is Noetherian. Indeed, choose some elements $x = (x_1, \dots, x_r)$ in B lifting a separable transcendence basis of B/b over k. Then the map $f : A[X]_{mA[X]} \to B$ given by $X \to x$ is flat by the local flatness criterion [25]. Since B/b is separable over $k(x)$ we get by above that f is essentially étale, which is enough.

An essentially étale local A-algebra B is an *étale neighbourhood* of A if its structure morphism induces an isomorphism on the residue fields.

COROLLARY 3.2 ([34]). *A local A-algebra B is an étale neighbourhood of A if and only if $B \cong (A[T]/(f))_{(m,T)}$, f being a monic polynomial in T over A such that $f(0) \in m$, $\partial f/\partial T(0) \notin m$.*

PROOF. If $B = (A[T]/(f))_{(m,T)}$ with f as above then $\partial f/\partial T \notin (m, T)$ and so B is étale over A as we have seen. Since $k \otimes_A B \cong (k[T]/T \cdot \overline{g})_{(T)}$, where $\overline{g} \in k[T]$ is a polynomial with $\overline{g}(0) = \partial f/\partial T(0) \bmod m \neq 0$ we see that $k \otimes_A B \cong k[T]/(T) \cong k$. Hence B is an étale neighbourhood of A.

Conversely, if B is an étale neighbourhood of A then by Theorem 3.1 $B \cong (A[T]/(f))_P$, where f is a monic polynomial in T and $P \subset A[T]$ is a prime ideal such that $f \in P$ but $\partial f/\partial T \notin P$. Since $B/mB \cong A/m$ we see that $A + mB = B$ and so $T \bmod f = a + x$ for an $a \in A$ and $x \in mB$. Changing T by $T + a$ we may reduce to the case when $t = T \bmod f \in mB = PB$. Then P contains the maximal ideal (m, T) of $A[T]$ and so $P = (m, T)$. By Taylor's formula $0 = f(t) = f(0) + t(\partial f/\partial T(0) + tf_1(t))$ for a polynomial $f_1 \in A[T]$. It follows $f(0) \in (t) \subset mB$ and by faithful flatness $f(0) \in m$. Similarly $\partial f/\partial T(t) = \partial f/\partial T(0) + tf_2(t)$ for some $f_2 \in A[T]$ and so

$$\partial f/\partial T(0) \equiv \partial f/\partial T(t) = \partial f/\partial T \bmod f \notin mB.$$

Hence $\partial f/\partial T(0) \notin m$. □

Let (A, m) be a local ring. Then there exists a Henselian local A-algebra $\widetilde{A}$ such that for every Henselian local A-algebra B there exists an unique local A-morphism $\widetilde{A} \to B$ (for details see [34,21,19]). $\widetilde{A}$ is unique up to an isomorphism and it is called the *Henselization* of A. By construction $\widetilde{A}$ is the filtered inductive limit of all étale neighbourhoods of A. In particular $\widetilde{A}$ is a flat A-algebra.

LEMMA 3.3. *Let (A, m) be a local ring, and B an étale neighbourhood of A. Suppose (A, m) is the filtered inductive limit of an inductive system of local rings (C_i, m_i), $\varphi_{ij} : C_i \to C_j$. Then there exists $j \in I$ and an étale neighbourhood B_j of C_j such that $A \otimes_{C_j} B_j \cong B$.*

PROOF. By Corollary 3.2, $B \cong (A[T]/(f))_{(m,T)}$ for a monic polynomial f in T over A such that $f(0) \in m$, $\partial f/\partial T(0) \notin m$. Then

$$A[T] = \varinjlim_{j \in I} B_j[T]$$

and there exists $i \in I$, $f_i \in B_i(T)$ which is monic and such that f_i is mapped by $\varphi_i : C_i[T] \to A[T]$ in f. Since $f(0) \in m$ and $\partial f/\partial T(0) \notin m$ we may find $j \in I$, $j > i$, such that the polynomial $f_j = \varphi_{ij}(f_i)$ satisfies $f_j(0) \in m_j$, $\partial f_j/\partial T(0) \notin m_j$. Clearly $B_j = (C_j[T]/(f_j))_{(m_j,T)}$ works. □

LEMMA 3.4. *Let (A, m) be an excellent Henselian local ring, $\hat{A}$ its completion, $\mathcal{A}$ the category of local A-algebras and $F : \mathcal{A} \to$ Sets a covariant functor such that the canonical map*

$$\varinjlim_{i\in I} F(B_i) \to F(\varinjlim_{i\in I} B_i)$$

is surjective for every filtered inductive system of local A-algebras $(B_i, \varphi_{ij})_{i\in I}$, φ_{ij} being local A-morphisms. Then for every $\hat{z} \in F(\hat{A})$ and every positive integer c there exists a $z \in F(A)$ such that $z \equiv \hat{z} \bmod m^c$; that is, the canonical morphisms $F(\hat{A}) \to F(\hat{A}/m^c\hat{A}) \simeq F(A/m^c)$ and $F(A) \to F(A/m^c)$ map $\hat{z}$, resp. z, in the same element of $F(A/m^c)$.

PROOF. The proof follows an idea of M. Artin [5]. We may express $\hat{A}$ as a filtered inductive union of finite type sub-A-algebras of $\hat{A}$, let us say $\hat{A} = \bigcup_{i\in I} D_i$. We have $\hat{A} = \bigcup_{i\in I} B_i$ for $B_i := (D_i)_{m\hat{A}\cap D_i}$. Fix $\hat{z}$ and c. By hypothesis we may find $i \in I$ such that $\hat{z}$ is in the image of the canonical morphism $F(\varphi_i) : F(B_i) \to F(\hat{A})$, where φ_i denotes the inclusion $B_i \hookrightarrow \hat{A}$.

Let us say $\hat{z} = F(\varphi_i(z_i))$ for a $z_i \in F(B_i)$. Since A has AP by Theorem 1.14, we find an A-morphism $\psi_i : D_i \to A$ which coincides with $\varphi_i\,|_{D_i}$ modulo m^c, that is the following diagram commutes

$$\begin{tikzcd}
D_i \arrow[r] \arrow[d] & B_i \arrow[rr, "\varphi_i"] & & \hat{A} \arrow[d] \\
A \arrow[r] & A/m^c & \cong & \hat{A}/m^c\hat{A}
\end{tikzcd}$$

The left map is ψ_i which extends clearly to a local A-morphism $\alpha_i : B_i \to A$ and $z = F(\alpha_i)(z_i)$ works. □

PROPOSITION 3.5. *Let (A, m) be an excellent Henselian local ring, $\hat{A}$ its completion, $X = (X_1, \ldots, X_r)$ some indeterminates, $A\langle X\rangle$ (resp. $\hat{A}\langle X\rangle$) the Henselization of $A[X]_{(m,X)}$ (resp. $\hat{A}[X]_{(m\hat{A},X)}$), f a system of polynomials over $A[X]$ in $Y = (Y_1, \ldots, Y_n)$, c a positive integer and $\tilde{y} = (\tilde{y}_1, \ldots, \tilde{y}_n)$ a solution of f in $\hat{A}\langle X\rangle$ such that $\tilde{y}_i \in \hat{A}\langle X_1, \ldots, X_s\rangle$, $1 \leqslant i \leqslant n$, for some integers s_i, $0 \leqslant s_i \leqslant r_i$. Then there exists a solution $y = (y_1, \ldots, y_n)$ of f in $A\langle X\rangle$ such that $y_i \in A\langle X_1, \ldots, X_{s_i}\rangle$ and $y_i \equiv \tilde{y}_i \bmod m^c\hat{A}\langle X_s, \ldots, X_{s_i}\rangle$ for each $1 \leqslant i \leqslant n$.*

PROOF. Let (B, b) be a local A-algebra. Let $F(B)$ be the set of all tuples $(C_1, \ldots, C_n, y)$, where for each i, $1 \leqslant i \leqslant n$, $C_i \subset B\langle X\rangle$ is an étale neighbourhood of

$$B[X_1, \ldots, X_{s_i}]_{(b, X_1, \ldots, X_{s_i})}$$

such that $C_1 \subset C_2 \subset \cdots \subset C_n$ and $y = (y_1, \ldots, y_n)$ is a solution of f in C_n such that $y_i \in C_i$ for each $1 \leqslant i \leqslant n$. Let $\alpha : B \to B'$ be a local morphism of local A-algebras and

b' the maximal ideal of B'. Then $D_i = B'[X_1, \ldots, X_{s_i}]_{(b', X_1, \ldots, X_{s_i})} \otimes_{B[X_1, \ldots, X_{s_i}]} C_i$ is an essentially étale $B'[X_1, \ldots, X_{s_i}]_{(b', X_1, \ldots, X_{s_i})}$-algebra and

$$D_i/(b', X_1, \ldots X_{s_i})D_i \cong B'/b' \otimes_{B/b} B/b \cong B'/b'$$

is a field. Thus $C'_i = D_{(b', X_1, \ldots, X_{s_i})}$ is an étale neighborhood of $B'[X_1, \ldots, X_{s_i}]_{(b', X_1, \ldots, X_{s_i})}$ and clearly $C'_1 \subset C'_2 \subset \cdots \subset C'_n \subset B'\langle X\rangle$. Since the canonical morphism $C_n \to C'_n$ maps a solution y of f in C_n to a solution $y' = 1 \otimes y$ of f in C'_n we see that we may define a function $F(\alpha): F(B) \to F(B')$ by $(C_1, \ldots, C_n, y) \to (C'_1, \ldots, C'_n, y')$ and so a covariant functor $F: \mathcal{A} \to$ Sets, $\mathcal{A}$ being the category of local A-algebras.

Now suppose that B is the filtered inductive limit of a filtered inductive system $((B_j, b_j), \varphi_{jt})_{j \in J}$ of local A-algebras and let $(C_1, \ldots, C_n, y) \in F(B)$. We have

$$B[X]_{(b,X)} \cong \varinjlim B_j[X]_{(b_j,X)}.$$

By Lemma 3.3 there exists $j \in J$ and an étale neighborhood $C_i^{(j)}$ of

$$B_j[X_1, \ldots, X_{s_i}]_{(b_j, X_1, \ldots, X_{s_i})}$$

such that

$$B[X_1, \ldots, X_{s_i}]_{(b, X_1, \ldots, X_{s_i})} \otimes_{B_j[X_1, \ldots, X_{s_i}]} C_i^{(j)} \cong C_i.$$

Clearly for a $t \in J, t > j$, we may suppose that f has a solution $y^{(t)}$ in

$$C_n^{(t)} := \left(B_t[X_1, \ldots, X_{s_n}] \otimes_{B_j[X_1, \ldots, X_{s_n}]} C_n^{(j)}\right)_{(b_t, X_1, \ldots, X_{s_n})}$$

which is mapped to y by $C_n^{(t)} \to C_n$. Thus $(C_1^{(t)}, \ldots, C_n^{(t)}, y^{(t)})$ is mapped by $F(\varphi_t)$ in $(C_1, \ldots, C_n, y)$, $\varphi_t: B_t \to B$ being the limit map and so F satisfies the hypothesis of Lemma 3.4.

Next, let $\tilde{y}$ be as in the hypothesis. Then there exists an étale neighborhood C_1 of

$$\hat{A}[X_1, \ldots, X_{s_1}]_{(m\hat{A}, X_1, \ldots, X_{s_i})}$$

containing $\tilde{y}_1$. By recurrence we find in this way $(C_1, \ldots, C_n)$ such that $(C_1, \ldots, C_n, \tilde{y}) \in F(\hat{A})$. By Lemma 3.4. there exists $(D_1, \ldots, D_n, y) \in F(A)$ such that $(D_1, \ldots, D_n, y) \equiv (C_1, \ldots, C_n, \tilde{y}) \bmod m^c$. Thus $y_i \in D_i \subset A\langle X_1, \ldots, X_{s_i}\rangle$, $y_i \equiv \tilde{y}_i \bmod m^c \hat{A}\langle X_1, \ldots, X_{s_i}\rangle$ for all $1 \leqslant i \leqslant s_i$ and $f(y) = 0$. □

THEOREM 3.6. *Let (A, m) be an excellent Henselian local ring, $\hat{A}$ its completion, $X = (X_1, \ldots, X_r)$ some indeterminantes, $A\langle X\rangle$ the Henselization of $A[X]_{(m,X)}$, f a system of polynomials over $A[X]$ in $Y = (Y_1, \ldots, Y_n)$, c a positive integer and $\hat{y} = (\hat{y}_1, \ldots, \hat{y}_n)$ a solution of f in $\hat{A}[[X]]$ such that $\hat{y}_i \in \hat{A}[[X_1, \ldots, X_{s_i}]]$, $1 \leqslant i \leqslant n$, for some integers s_i, $0 \leqslant s_i \leqslant r$. Then there exists a solution $y = (y_1, \ldots, y_n)$ of f in $A\langle X\rangle$ such that $y_i \in A\langle X_1, \ldots, X_{s_i}\rangle$ and $y_i \equiv \hat{y}_i \bmod (m, X_1, \ldots, X_{s_i})^c \hat{A}[[X_1, \ldots, X_{s_i}]]$, $1 \leqslant i \leqslant n$.*

PROOF. First we see that it is enough to study the case when A is a Noetherian complete local ring. Indeed, if our statement holds for such rings, then there exists a solution $\widetilde{y} = (\widetilde{y}_1, \dots, \widetilde{y}_n)$ of f in $\hat{A}\langle X\rangle$ such that $\widetilde{y}_i \in \hat{A}\langle X_1, \dots, X_{s_i}\rangle$ and $\widetilde{y}_i \equiv \hat{y}_i \bmod (m\hat{A}, X_1, \dots, X_{s_i})^c \hat{A}[[X_1, \dots, X_{s_i}]]$, $1 \leqslant i \leqslant n$. Then it is enough to apply Proposition 3.5.

From now on we may suppose that A is a Noetherian complete local ring. After a renumbering we may suppose

$$0 \leqslant s_1 = \dots = s_{i_1} < s_{i_1+1} = \dots = s_{i_2} < \dots < s_{i_{t-1}+1} = \dots = s_{i_t} \leqslant r$$

and $i_t = n$. Set $r_j = s_{i_j}$, $1 \leqslant j \leqslant t$, $s_0 = r_0 = i_0 = 0$ and $D_j = A[[X_1, \dots, X_{r_{j-1}}]]$, for $1 \leqslant j \leqslant t$. Apply decreasing induction on j, $1 \leqslant j \leqslant t$, to find for each e, $i_{j-1} < e \leqslant n$, an element $y_e^{(j)}$ in $D_j\langle X_{r_{j-1}+1}, \dots, X_{s_e}\rangle$ such that

$$y_e^{(j)} \equiv \hat{y}_e \bmod (m, X_1, \dots, X_{s_e})^c A[[X_1, \dots, X_{s_e}]]$$

and

$$f\big(\hat{y}_1, \dots, \hat{y}_{i_{j-1}}, y_{i_{j-1}+1}^{(j)}, \dots, y_n^{(j)}\big) = 0.$$

If $j = t$ this follows because $D_t\langle X_{r_{t-1}+1}, \dots, X_{r_t}\rangle$ has AP, being excellent Henselian (see Theorem 1.14). Indeed, apply Lemma 1.5 to the system of polynomials

$$f(\hat{y}_1, \dots, \hat{y}_{i_{t-1}}, Y_{i_{t-1}+1}, \dots, Y_{i_t}).$$

Suppose $1 \leqslant j < t$. By induction hypothesis, we have $(y_e^{(j+1)})_{i_j < e \leqslant n}$ as above in $D_{j+1}\langle X_{r_j+1}, \dots, X_r\rangle$ such that

$$f\big(\hat{y}_1, \dots, \hat{y}_{i_j}, y_{i_j+1}^{(j+1)}, \dots, y_n^{(j+1)}\big) = 0.$$

If $j = 1$ and $s_{i_1} = 0$ then there exists nothing to show. Otherwise, apply Proposition 3.5 to the case $A = D_j\langle X_{r_{j-1}+1}, \dots, X_{r_j}\rangle$, $\hat{A} = D_{j+1}$ and to the system of equations

$$f\big(\hat{y}_1, \dots, \hat{y}_{i_{j-1}}, Y_{i_{j-1}+1}, \dots, Y_n\big) = 0.$$

We obtain the wanted solution $y^{(j)}$. □

COROLLARY 3.7. *Let K be a field, $X = (X_1, \dots, X_r)$ some indeterminantes, $K\langle X\rangle$ the algebraic power series in X over K (that is, the algebraic closure of $K[X]$ in $K[[X]]$ – the Henselization of $K[X]_{(X)}$), f a system of polynomials over $K\langle X\rangle$ in $Y = (Y_1, \dots, Y_n)$, c a positive integer and $\hat{y} = (\hat{y}_1, \dots, \hat{y}_n)$ a solution of f in $K[[X]]$ such that $\hat{y}_i \in K[[X_1, \dots X_{s_i}]]$, $1 \leqslant i \leqslant n$, for some positive integers $s_i \leqslant r$. Then there exists a solution $y = (y_1, \dots, y_n)$ of f in $K\langle X\rangle$ such that $y_i \in K\langle X_1, \dots, X_{s_i}\rangle$ and $y_i \equiv \hat{y}_i \bmod (X_1, \dots, X_{s_i})^c K[[X_1, \dots, X_{s_i}]]$, $1 \leqslant i \leqslant n$.*

PROOF. Using Theorem 3.6 it is enough to reduce the problem to the case when f is from $K[X, Y]$. Let p be the kernel of the map $K\langle X\rangle[Y] \to K[[X]]$ given by $Y \to \hat{y}$ (so p is prime!) and $q := p \cap K[[X, Y]]$. We may enlarge f in order to generate p. Since the map $K[X, Y] \to K\langle X\rangle[Y]$ is a filtered inductive limit of étale neighborhoods we see that $qK\langle X\rangle[Y]$ is a reduced ideal and p is a minimal ideal associated to $qK\langle X\rangle[Y]$. Thus there exists $g \in K\langle X\rangle[Y]\backslash p$ such that $gp \subset qK\langle X\rangle[Y]$. Then $g(\hat{y}) \neq 0$ and there exists $u \in \mathbb{N}$, $u \geqslant c$ such that $g(\hat{y}) \not\equiv 0 \bmod X^u$.

Apply Theorem 3.6 for a system of generators h of q and u. Then there exists a solution $y = (y_1, \dots, y_n)$ of h in $K\langle X\rangle$ such that $y_i \in K\langle X_1, \dots, X_{s_i}\rangle$ and $y_i \equiv \hat{y}_i \bmod (X_1, \dots, X_{s_i})^u K[[X_1, \dots, X_{s_i}]]$, $1 \leqslant i \leqslant n$. It remains to show that y is a solution of f too. But $g(y) \equiv g(\hat{y}) \not\equiv 0 \bmod X^u$ and so $g(y) \neq 0$. Thus

$$g \notin a := (Y_1 - y_1, \dots, Y_n - y_n)K\langle X\rangle[Y]$$

and $gp \subset qK\langle X\rangle[Y] \subset a$ because $h(y) = 0$. Hence $p \subset a$, that is $f(y) = 0$. □

REMARK 3.8. If the sets of variables X involved in the $\hat{y}_i$ are not "nested", that is they are not totally ordered by inclusion (as above!) then Corollary 3.7 does not hold as Becker [9] noticed. If in this corollary we consider the convergent power series ring over $K = \mathbb{C}$ instead of the algebraic power series ring over $\mathbb{C}$ then again the result does not hold as Gabrielov noticed in [16]. However a special approximation in nested subrings holds also in $\mathbb{C}\{X_1, \dots, X_n\}$ as Grauert showed in [17]. The above corollary appeared when char $K = 0$ in [23, Ch. III] with some nice applications in [23, Ch. IV], the proof being wrong. The idea of Proposition 3.5 comes from H. Kurke and G. Pfister, who actually noticed that the above corollary holds if rings of type $K[[X_1, \dots, X_s]]\langle X_{s+1}, \dots, X_r\rangle$ have AP. This follows from Theorem 1.14 and our presentation here follows essentially [31, (3.6), (3.7)]. Extensions of Theorem 3.6 are given in [39,37].

4. Cohen Algebras and General Néron Desingularization in Artinian local rings

LEMMA 4.1. *Let (A, m, k) be a Noetherian local ring and $L \supset k$ a simple field extension. Then there exists a flat local A-algebra (B, b), essentially of finite type such that $b = mB$ and $B/b \cong L$. If L/k is separable then B is essentially smooth.*

PROOF. Let $L = k(x)$. If x is transcendental over k then $B = A[X]_{mA[X]}$ is clearly essentially smooth. If x is algebraic over k then let $P \in A[X]$ be a monic polynomial lifting of $\overline{P} := \operatorname{Irr}(k, x)$ and $C = A[X]/(P)$. Then mC is prime in C because $C/mC \cong L$. Thus $B := C_{mC}$ works. If $L|k$ is separable then $\overline{P}$ is separable and so $\partial P/\partial X \notin mC$, that is B is essentially smooth. □

PROPOSITION 4.2. *Let (A, m, k) be a Noetherian local ring and $L \supset k$ a field extension. Then there exists a flat Noetherian local A-algebra (B, b) such that*

(i) *$B/b \cong L$, $b = mB$,*

(ii) (B, b) *is a filtered inductive limit of flat local* A*-algebras* $(C_i, c_i)_{i \in I}$ *essentially of finite type such that* $c_i = mC_i$, $i \in I$.

Moreover if L/k *is separable then* C_i *are essentially smooth over* A.

PROOF. By transfinite induction we construct a family of subfields $(K_i)_{1 \leqslant i \leqslant \theta}$ of L such that $K_0 = k$, $K_\theta = L$, $K_{i+1} = K_i(x_{i+1})$ for an element $x_{i+1} \in L \backslash K_i$, $i < \theta$, and $K_i = \bigcup_{j<i} K_j$ if i is a limit ordinal. Using again the transfinite induction we construct a family of local A-algebra $(C_i)_{1 \leqslant i \leqslant \theta}$ such that $B_0 = A$, B_{i+1} is a flat local B_i-algebra given by Lemma 4.1 and

$$B_i = \varinjlim_{j<i} B_j$$

if i is a limit ordinal. Denote $B = B_\theta$. By induction we see that all B_i are Noetherian. Indeed if i is a limit ordinal and $(B_j)_{j<i}$ are all Noetherian rings then B_i is Noetherian too since $b_j B_{j+1} = b_{j+1}$ and the map $B_j \to B_{j+1}$ are all flat. Thus (B, b) is Noetherian local and $B/b \cong L$ by construction.

Now note that if i is a limit ordinal then B_{i+1} is an extension of a flat, local B_j-algebra for a $j < i$ enough high. Thus (ii) holds. If L/K is separable then each flat A-algebra C_i from (ii) must be essentially smooth (see Section 3) because $c_i = mC_i$ and C_i/c_i is separable over k. □

Let (B, b) be a local algebra over a ring A. We call B a *formally smooth* A-algebra if given an A-algebra C, an ideal $I \subset C$ with $I^2 = 0$ and an A-algebra morphism $g : B \to C/I$ such that $g(b^s) = 0$ for a certain $s \in \mathbb{N}$, there exists a lifting of g to C. Clearly formal smoothness is weaker then quasi-smoothness and it is essentially a topological notion (with respect to the b-adic topology).

By [19,25, 19.7.1,28.10] a Noetherian local algebra (B, b) over a Noetherian local ring (A, m, k) is formally smooth if and only if B is flat over A and B/mB is a formally smooth k-algebra. This remind us the similar property of essentially smooth maps from Section 3 (in fact formally smooth, essentially of finite type, local A-algebra are essentially smooth).

PROOF OF LEMMA 2.7. Let (B, b, L) be a flat Noetherian local algebra over a Noetherian local ring (A, m, k) such that $b = mB$ and L/k is separable, that is B/mB is quasi smooth (so formally smooth!). Then B is formally smooth over A. But formally smooth morphisms $A \to B$ are regular if A is quasi-excellent (see [2,11]). □

THEOREM 4.3 (Grothendieck [19]). *Let* (A, m, k) *be a Noetherian local ring and* $L \supset k$ *a separable field extension. Then there exists an unique flat, Noetherian complete local* A*-algebra* (R, q) *such that* $q = mR$ *and* $R/mR \cong L$.

PROOF. By Proposition 4.2 there exists a flat Noetherian local A-algebra (B, b) such that $b = mB$ and $B/b \cong L$. Then the completion R of B works.

Now let (R, q), (R', q') be two flat Noetherian complete local A-algebras such that $q = mR$, $q' = mR'$ and $R/q \cong R'/q' \cong L$. Then R, R' are formally smooth A-algebras as we have seen. By recurrence the surjective map $R \to L$ can be lifted to an A-morphism $g_i : R \to R'/q'^i$ for all positive integers i. The maps (g_i) defines an A-morphism

$$g : R \to \varprojlim R'/q'^i \cong R'$$

such that $R/q \otimes_R g$ coincides with 1_L modulo some isomorphisms. Note that g is a surjection by [25, 8.4]. Similarly there exists a surjection $h : R' \to R$. Then gh and hj are isomorphisms since R, R' are Noetherian rings which is enough. □

The unique (up to an isomorphism) Noetherian complete local A-algebra (R, q) given by Theorem 4.3 is called the *Cohen A-algebra of residue field L.*

REMARK 4.4. The A-algebra A_1 defined in Section 2 is in fact the Cohen A-algebra of residue field k^*.

PROPOSITION 4.5. *General Néron Desingularization holds in Artinian local rings.*

PROOF. Let $u : A \to A'$ be a regular morphism of Artinian local rings and m the maximal ideal of A. Then A'/mA' must be a local regular ring which is also Artinian. Thus A'/mA' is a field k' and mA' is the maximal ideal of A'.

By Proposition 4.2 there exists a flat Noetherian local A-algebra (B, b) such that

(i) $B/b \cong k'$, $b = mB$,

(ii) (B, b) is a filtered inductive limit of essentially smooth A-algebras.

Since m is nilpotent, b is too and so B is Artinian too. Thus A', B are complete local rings and by Theorem 4.4 they are A-isomorphic. Hence A' is a filtered inductive limit of essentially smooth A-algebras. By Theorem 3.1 A' is a filtered inductive limit of A-algebras of type $(A[Y, T]/f)_g$, $Y = (Y_1, \ldots, Y_n)$, where f is a monic polynomial in T over $A[Y]$ and g is a multiple of $\partial f/\partial Y$. □

5. Jacobi–Zariski sequence and the smooth locus

Let B be an algebra over a ring A and $A[Y]/I$ a presentation of B over A, where $Y = (Y_i)_i$ is a set of indeterminantes not necessarily finite. Let $\bigoplus BdY_i$ be the free B-module on a basis $dY = (dY_i)_i$ in bijection with Y and $\tilde{d} : I \to \bigoplus BdY_i$ the map

$$f \to \sum (\partial f/\partial Y_i) dY_i.$$

Clearly $\tilde{d}(I^2) = 0$ and so $\tilde{d}$ induces a map $d : I/I^2 \to \bigoplus BdY_i$. Let $\Gamma_{B/A}$ and $\Omega_{B/A}$ be the kernel and cokernel of d. $\Gamma_{B/A}$ and $\Omega_{B/A}$ does not depend of the choice of the presentation and are functorial in B/A. If $B = A/a$ then $\Gamma_{B/A} \cong a/a^2$ and $\Omega_{B/A} = 0$. If $S \subset B$ is a multiplicative system then $\Gamma_{S^{-1}B/A} \cong S^{-1}(\Gamma_{B/A})$ and $\Omega_{S^{-1}B/A} \simeq S^{-1}\Omega_{B/A}$.

THEOREM 5.1 (Jacobi–Zariski sequence). *If $B \to C$ is a morphism of A-algebras then there exists a natural exact sequence of C-modules*

$$\Gamma_{C/A} \to \Gamma_{C/B} \to C \otimes_B \Omega_{B/A} \to \Omega_{C/A} \to \Omega_{C/B} \to 0$$

which can be extended on the left with the term $C \otimes_B \Gamma_{B/A}$ providing $\Omega_{B/A}$ is flat over B.

THEOREM 5.2. *Let $B = A[Y]/I$ be an algebra over a ring A. Then the following conditions are equivalent:*

(1) *B is quasi-smooth over A,*
(2) *$d: I/I^2 \to \bigoplus AdY_i$ is a split monomorphism,*
(3) *$\Gamma_{B/A} = 0$ and $\Omega_{B/A}$ is projective over A.*

The proofs of these theorems can be found, together with many other details, in [1] (an easy exposition of this subject is given in [38]). Let B be a finitely presented algebra over a ring A and $B = A[Y]/I$ such a finite presentation, $Y = (Y_1, \dots, Y_n)$. Let $H_{B/A}$ be the radical in B of $\sum \Delta_f((f):I)B$, the sum being taken over all finite systems of polynomials $f = (f_1, \dots, f_r)$ of I, Δ_f being the ideal generated by all $r \times r$-minors of the Jacobian matrix $(\partial f_i/\partial X_j)$. The ideal $H_{B/A}$ does not depend on the presentation as shows the following:

PROPOSITION 5.3. *Let $q \in \operatorname{Spec} B$. Then B_q is essentially smooth over A if and only if $q \not\supset H_{B/A}$.*

PROOF. Suppose that $q \not\supset H_{B/A}$ and let $p \subset A[Y]$ be the inverse image of q. Then there exists a finite system of polynomials f of I such that $I_p = (f)A[Y]_p$ and $\Delta_f \not\subset p$, let us say $M = \det(\partial f_i/\partial Y_j)_{1 \leqslant i,j \leqslant r} \notin p$. Let

$$\alpha: \bigoplus_{i=1}^{n} B_q dY_i \to \bigoplus_{i=1}^{r} B_q dY_i$$

be the projection on the first r summands and

$$\beta: \bigoplus_{i=1}^{r} B_q dY_i \to \left(I/I^2\right)_q$$

the surjective map given by $dY_i \to f_i \bmod I^2$. Then $\alpha d_q \beta$ is an isomorphism (its determinant is M!) and so $\beta, \alpha d_q$ are also isomorphisms. Thus d_q is a split monomorphism and so B_q is essentially smooth by Theorem 5.2.

Conversely, if d_q is a split monomorphism then $(I/I^2)_q$ is free of a certain rank $r \leqslant n$. Let f be a system of r polynomials from I which lifts a basis in $(I/I^2)_q$. Then

$$(f)A[Y]_q = qA[Y]_q$$

by Nakayama's Lemma and the matrix associated to d_q in the basis (f_i mod I^2) resp. dY_j is $(\partial f_i/\partial Y_j)$. As d_q splits we get $\Delta_f \not\subset q$. □

COROLLARY 5.4. *Let $s \in B$. Then B_s is smooth over A if and only if $s \in H_{B/A}$.*

PROOF. B_s is smooth over A if and only if $(\Gamma_{B/A})_s = 0$ and $(\Omega_{B/A})_s$ is projective by Theorem 5.2. The last conditions hold exactly when $(\Gamma_{B/A})_q = 0$ and $(\Omega_{B/A})_q$ is projective for all $q \in \operatorname{Spec} B$ with $s \notin q$, that is when B_q is essentially smooth (in other words $q \not\supset H_{B/A}$) for all $q \in \operatorname{Spec} B$ with $s \notin q$. Hence B_s is smooth over A if and only if $s \in \bigcap_{q \supset H_{B/A}} q = H_{B/A}$. □

LEMMA 5.5. *Let $v : B \to C$ be a morphism of finite presentation algebras over a ring A. Then $v(H_{B/A})C \cap H_{C/B} \subset H_{C/A}$.*

PROOF. If $q \in \operatorname{Spec} C$, $q \not\supset v(H_{B/A})C \cap H_{C/B}$ then $v^{-1}q \not\supset H_{B/A}$ and $q \not\supset H_{C/B}$ and so $B_{w^{-1}q}$ is essentially smooth over A and C_q is essentially smooth over B. Thus C_q is essentially smooth over A and so $q \not\supset H_{C/A}$. □

LEMMA 5.6. *Let A be a Noetherian ring $u : B \to A'$ a morphism of A-algebras and $x \in B$ an element. Suppose that B is of finite type over A and $u(x) \in \sqrt{u(H_{B/A})A'}$. Then there exists a finite type A-algebra C such that u factors through C, let us say $u = wv$, $v(x) \in H_{C/A}$ and $v(H_{B/A})C \subset H_{C/B}$ (in particular $v(H_{B/A})C \subset H_{C/A}$).*

PROOF. Let $t = (t_1, \ldots, t_s)$ be a system of generators of $H_{B/A}$. By hypothesis we have $u(x)^r = \sum_{i=1}^s u(t_i)z_i$ for some elements $z_i \in C$ and a certain $r \in \mathbb{N}$. Put

$$C = B[z]/\left(x^r - \sum_{i=1}^s t_i Z_i\right),$$

$Z = (Z_1, \ldots, Z_r)$, and let $w : C \to A'$ be the map given by $Z \to z$. Clearly C_{t_i} is a polynomial B-algebra and so $v(t_i) \in H_{C/B}$, v being the structure morphism of the B-algebra C. Thus $v(H_{B/A}) \subset H_{C/B}$ and $v(x) \in v(H_{B/A})C$. By Lemma 5.5 $v(H_{B/A}) \subset H_{C/A}$ and $v(x) \in H_{C/A}$.

Let B be a finitely presented A-algebra and $b \in H_{B/A}$ an element. We say that b is *standard* (resp. *strictly standard*) *with respect to the presentation* $B = A[Y]/I$ if b lies in $\sqrt{\Delta_f((f):I)B}$ (resp. $\Delta_f((f):I)B$) for a finite system of polynomials $f = (f_1, \ldots f_r)$ of I. □

LEMMA 5.7 (Elkik [15,38]). *If $b \in H_{B/A}$ and $(I/I^2)_b$ is free over B then b is standard for the above presentation.*

PROOF. Let $f_1, \ldots, f_r$ be a base for $(I/I^2)_d$ and $h \in A[Y]$ representing b. Then

$$I_h = (f)A[Y]_h + I_h^2$$

and we have $(1+\alpha)I_h \subset (f)A[Y]_h$ for a certain $\alpha = g/h^s$, $g \in I$. Thus

$$h^e(h^s + g) \in ((f):I)$$

for a certain $e \in \mathbb{N}$ and so $b^{s+e} \in ((f):I) \cdot B$. Since the map

$$d: I_h/I_h^2 \to \bigoplus_{i=1}^n B_b dY_i$$

is a split monomorphism the $r \times r$-minors of its matrix $(\partial f_i/\partial Y_j)$ generate whole B_b. Thus a power of b lies in Δ_f. □

LEMMA 5.8. *If $b \in H_{B/A}$ and $(\Omega_{B/A})_b$ is free then b is standard with respect to the presentation $B = A[Y, Z]/(I, Z)$ where $Z = (Z_1, \ldots, Z_n)$.*

PROOF. As B_b is smooth over A we have

$$\bigoplus_{i=1}^n B_b dY_i = \left(I/I^2\right)_b \oplus (\Omega_{B/A})_b$$

by Theorem 5.2. Since $\Omega_{B_b/A} = (\Omega_{B/A})_b$ is free of rank $\leqslant n$ we get $(I/I^2)_b \oplus B_b^n$ free. Now, let $J = (I, Z)$. We have

$$0 \to (J/J^2)_b \xrightarrow{d_J} \left(\bigoplus_{i=1}^n B_b dY_i\right) \oplus \left(\bigoplus_{j=1}^n B_b dZ_j\right) \to \Omega_{B_b/A} \to 0$$

and a surjective map

$$\varphi: \left(I/I^2\right)_b \oplus \left(\bigoplus_{j=1}^n B_b Z_j \bmod Z^2\right) \to \left(J/J^2\right)_b.$$

But d_J maps isomorphically $\varphi(\bigoplus_{j=1}^n B_b(Z_j \bmod Z^2))$ on $\bigoplus_{j=1}^n B_b dZ_j$. After splitting this off from $d_J\varphi$ it remains d_I which is injective since B_b is smooth over A. Thus φ must be an isomorphism and so $(J/J^2)_b \cong (I/I^2)_b \oplus B_b^n$ which is free. Hence b is standard for the presentation $B = A[Y, Z]/(I, Z)$ by Lemma 5.7. □

PROPOSITION 5.9. *Let $B = A[Y]/I$, $Y = (Y_1, \ldots, Y_n)$ be a finitely presented A-algebra and $C = S_B(I/I^2)$ be the symmetric algebra over A associated to I/I^2. Then $H_{B/A}C \subset H_{C/A}$ and $\Omega_{C_b/A}$ is free for each $b \in H_{B/A}$. Therefore C has a presentation such that the image in C of any $b \in H_{B/A}$ is standard.*

PROOF. Let $b \in H_{B/A}$ then

$$d_b : \left(I/I^2\right)_b \to \bigoplus_{j=1}^{n} B_b dY_i$$

is a split monomorphism and so $(I/I^2)_b$ is projective. Then C_b is locally isomorphic with a polynomial algebra over B_b and so it is smooth. Thus $b \in H_{C/B}$, even $b \in H_{C/A}$ by Lemma 5.5 and so $H_{B/A}C \subset H_{C/A}$. The Jacobi–Zariski sequence written for the A-morphism $B_b \to C_b$:

$$0 = \Gamma_{C_b/B_b} \to C_b \otimes_{B_b} \Omega_{B_b/A} \to \Omega_{C_b/A} \to \Omega_{C_b/B_b} \to 0$$

splits because $\Omega_{C_b/B_b} \cong C_b \oplus_{B_b} (I/I^2)_b$ is projective. Then

$$\Omega_{C_b/A} \cong C_b \otimes_{B_b} \left(\Omega_{B_b/A} \oplus \left(I/I^2\right)_b\right) \cong C_b \otimes_{B_b} \left(\bigoplus_{i=1}^{n} B_b dY_i\right)$$

which is free, the last isomorphism holds because d_b is a split monomorphism. Now it is enough to apply Lemma 5.8. □

COROLLARY 5.10. *Let A be a Noetherian ring, $u : B \to A'$ a morphism of A-algebras and $b \in B$ an element. Suppose that B is of finite type over A and $u(b) \in \sqrt{H_{B/A}A'}$. Then then exists a finite type A-algebra C such that*

(i) *u factors through C, let us say $u = wv$, $v : B \to C$, $w : C \to A'$,*
(ii) *$v(b)$ is standard for a certain presentation of C over A,*
(iii) *$v(H_{B/A}) \subset H_{C/A}$.*

PROOF. By Lemma 5.6 we may reduce to the case when $b \in H_{B/A}$. Applying Proposition 5.9 we get C as it is necessary except that we must show that there exists $w : C \to A'$ such that $u = wv$. Since C is a symmetric algebra over B, the inclusion $v : B \hookrightarrow C$ has a canonical retraction t given by $S_B(I/I^2) \to S_B(I/I^2)/(I/I^2) \cong B$. Thus $w = ut$ works. □

6. General Néron desingularization

THEOREM 6.1 (Néron [24]). *Let $R \subset R'$ be an unramified extension of discrete valuation rings which induces separable field extensions on the fraction and residue fields. Then R' is a filtered inductive union of its sub-R-algebras which are essentially smooth.*

Note that the hypothesis of Néron's desingularization says in fact that the inclusion map $R \to R'$ is regular. Thus Theorem 1.15 on discrete valuation rings follows from Theorem 6.1. The following theorem is a stronger variant of Theorem 1.15:

THEOREM 6.2. *Let $u: A \to A'$ be a morphism of Noetherian rings. Then u is regular if and only if A' is a filtered inductive limit of smooth A-algebras.*

The sufficiency is trivial because a filtered inductive limit of regular morphisms is regular, since a Noetherian ring, which is a filtered inductive limit of regular rings must be regular too. The necessity is equivalent with Theorem 1.15. Indeed, if Theorem 1.15 holds then for every finite A-algebra B and every morphism of A-algebras $v: B \to A'$ there exists a smooth A-algebra C such that v factors through C. Thus A' is a filtered inductive limit of smooth A-algebras by the following elementary lemma (a proof is given in [38]).

LEMMA 6.3. *Let S be a class of finitely presented algebras over a ring A and A' an A-algebra. The following statements are equivalent:*

(1) *A' is a filtered inductive limit of algebras in S,*
(2) *If B is a finitely presented A-algebra and $v: B \to A'$ is a morphism of A-algebras then v factors through some C in S.*

Conversely, if the necessity in Theorem 6.2 holds then for every finite type A-algebra B and every $v: B \to A'$ morphism of A-algebras there exists a smooth A-algebra D such that v factors through D, let us say $v = wt$, $w: D \to A'$, $t: B \to D$. Thus it is enough to apply the following

LEMMA 6.4. *Let $v: B \to A'$ be a morphism of algebras over a Noetherian ring A. Suppose that B is of finite type over A and $v(H_{B/A})A' = A'$. Then v factors through an A-algebra $C = (A[Y]/(f))_g$, $Y = (Y_1, \dots, Y_n)$, where $f = (f_1, \dots, f_r)$, $r \leqslant n$, are polynomials in Y over A and g belongs to the ideal Δ_f generated by the $r \times r$-minors of the Jacobian matrix $(\partial f_i / \partial Y_j)$.*

PROOF. By Corollary 5.10 there exists a finite type A-algebra C such that

(i) v factors through C, let us say $v = wt$, $t: B \to C$, $w: C \to A'$,
(ii) 1 is standard for a certain presentation of C over A, let us say $C \cong A[Y]/I$, $Y = (Y_1, \dots, Y_n)$.

Then there exists $f = (f_1, \dots, f_r)$ in I, $r \leqslant n$ such that $1 \in \Delta_f((f): I)C$. Thus there exists $\alpha \in I$ such that $g = 1 + \alpha \in \Delta_f \cap ((f): I)$. Hence $C \cong (A[Y]/(f))_g$ as it is required. □

From now on, let $u: A \to A'$ be a morphism of Noetherian rings, B a finite type A-algebra and $v: B \to A'$ an A-morphism. We can suppose that $h_B := \sqrt{v(H_{B/A})A'} \neq A'$ because otherwise we may apply Lemma 6.4. Let q be a minimal prime over ideal of h_B. After [38] we say that $A \to B \to A' \supset q$ is *resolvable* if there exists a finite type A-algebra C such that v factors through C, let us say $v = wt$, $t: B \to C$, $w: C \to A'$ and $h_B \subset h_C := \sqrt{w(H_{C/A})A'} \not\subset q$.

PROPOSITION 6.5. *Suppose that $p := u^{-1}q$ is a minimal over ideal of $u^{-1}(h_B)$, $A \to A'_q$ is flat and A'_q/pA'_q is a geometrically regular $k(p) := A_p/pA_p$-algebra. Then $A \to B \to A' \supset q$ is resolvable.*

For proving the necessity in Theorem 6.2 and so Theorem 1.15 it is enough to show that the above proposition holds. Indeed if $h_B \neq A'$ then we may choose a minimal prime over ideal q of h_B such that $p = u^{-1}q$ is a minimal prime over ideal of $a = u^{-1}(h_B)$ (if $h_B = \bigcap_{i=1}^{s} q_i$, q_i being minimal primes over ideals of h_B, then a minimal prime over ideal p of $a = \bigcap_{i=1}^{s} u^{-1}(q_i)$ must contain one of $u^{-1}q_i$ and so $p = u^{-1}q_i$). Thus the hypothesis of Proposition 6.5 are fulfilled (u is regular in Theorem 6.2!). Hence $A \to B \to A' \supset q$ is resolvable (roughly speaking we may increase h_B). By Noetherian induction we arrive to the case $h_B = A'$ which is solved by Lemma 6.4.

If A contains $\mathbb{Q}$ then Proposition 6.5 follows from the following two lemmas:

LEMMA 6.6. *Suppose ht* $q = 0$ *and* $A_p \to A_p \otimes_A B \to A'_q \supset qA'_q$, $p = u^{-1}(q)$ *is resolvable. Then* $A \to B \to A' \supset q$ *is resolvable too.*

LEMMA 6.7 (Main Lemma). *Let* $a \in A, \bar{A} := A/a^8A$, $\bar{B} := B/a^8B$, $\bar{A}' := A'/a^8A'$ *and* $\bar{q} = q/a^8A'$. *Suppose that*

(i) $\mathrm{Ann}_A(a^2) = \mathrm{Ann}_A(a), \mathrm{Ann}_{A'}(u(a)^2) = \mathrm{Ann}_{A'}(u(a))$,

(ii) *a is strictly standard for a certain presentation of B over A,*

(iii) $\bar{A} \to \bar{B} \to \bar{A}' \supset \bar{q}$ *is resolvable.*

Then $A \to B \to A' \supset q$ *is resolvable too.*

PROOF OF PROPOSITION 6.5 WHEN $A \supset \mathbb{Q}$. Apply induction on $ht\ q$. If $ht\ q = 0$ then by Lemma 6.6 it is enough to show that

$$A_p \to A_p \otimes_A B \to A'_q \supset qA'_q$$

is resolvable. It is easy to see that

$$H_{A_p \otimes_A B/A_p} \supset H_{B/A} \cdot (A_p \otimes_A B),$$

since the smoothness preserves by base change. If $qA'_q \supset H_{A_p \otimes_A B/A_p}$ (otherwise we have nothing to show) then we may apply the Proposition 4.6 (the Artinian local case) because the morphism $A_p \to A'_q$ is regular by hypothesis (it is flat and A'_q/pA'_q is regular, Artinian local $k(p)$-algebra thus a separable field extension of $k(p)$!).

Suppose now $ht\ q > 0$. If $ht\ p = 0$ then choose x in h_B which is mapped by v in a regular parameter of A'_q/pA'_q. Consider $u' : A[X] \to A'$ given by $X \to x$, $B' = B[X]$ and $v' : B' \to A'$ extending v by $X \to x$. The map $A[X]_{(p,x)} \to A'_q$ is flat by the local flatness criterion [25] since the map $k(p)[X]_{(X)} \to A'_q/pA'_q$ is flat by [25, Th. 23.1]. By construction $A'_q/(p,x)A'_q$ is still a regular local ring, even a geometrically regular $k(p)$-algebra since $\mathrm{char}\, k(p) = 0$ (*A contains* $\mathbb{Q}$!).

Clearly

$$H_{B[X]/A[X]} \supset H_{B/A} \cdot B[X]$$

as above and it is enough to show that

$$A[X] \to B[X] \to A' \supset q$$

is resolvable. Thus we may suppose $ht\ p > 0$ too. Choose a in $u^{-1}(h_B)$ such that a lies in no height 0 prime contained in p. We have $ht\ p/aA < ht\ p$ and so $ht\ q/aA' < ht\ q$ because $ht\ q = ht\ p + \dim A'_q/pA'_q$ by flatness of $A_p \to A'_q$ (see [25, 15.1]). By Corollary 5.10 there exists a finite type A-algebra C such that

(i) v factors through C, let us say $v = wt$, $t: B \to C$, $w: C \to A'$,
(ii) a is standard for a certain presentation of C over A,
(iii) $t(H_{B/A}) \subset H_{C/A}$.

Changing a to one of its powers we may suppose that a is strictly standard for the same presentation of C over A. Since the chain $\mathrm{Ann}_A(a) \subset \mathrm{Ann}_A(a^2) \subset \cdots$ stops by Noetherianity we may suppose $\mathrm{Ann}_A(a^2) = \mathrm{Ann}_A(a)$ and similarly $\mathrm{Ann}_{A'}(u(a))^2 = \mathrm{Ann}_{A'}(u(a))$, changing a to a power of a. Let $\bar{A} = A/(a^8)$, $\bar{A}' = A'/a^8A'$, $\bar{C} = C/a^8C$ and $\bar{q} = q/a^8A'$. By induction hypothesis ($ht\ \bar{q} < ht\ q$!) we may suppose that $\bar{A} \to \bar{C} \to \bar{A}' \supset \bar{q}$ is resolvable. Then by the Main Lemma $A \to C \to A' \supset q$ is resolvable too. □

PROOF OF LEMMA 6.6. Let D be a finite type A_p-algebra such that

$$v_p : B_p := A_p \otimes_A B \to A'_q$$

factors through D, let us say $v_p = \beta_p \alpha_p$, $\alpha_p : B_p \to D$, $\beta_p : D \to A'_q$ and $\beta_p(H_{D/A_p}) = A'_q$ (by hypothesis $A_p \to B_p \to A'_q \supset qA'_q$ is resolvable). Using Lemma 6.4 we may suppose even D smooth over A_p. We claim that we can change D such that there exists a finite type B-algebra C such that v factors through C, let us say $v = \beta\alpha$, $\beta : C \to A'$, α being canonically, $D \cong B_p \otimes_B C$ as B-algebras and v_p is induced by v.

Indeed, let $D \cong B_p[Z]/(g)$, $Z = (Z_1, \ldots, Z_s)$, $g = (g_1, \ldots, g_e)$, g_i being supposed with coefficients in B and β_p given by $Z \to y/t$ for $y \in A'^s$, $t \in A' \backslash q$. Indeed, for large integer $c \geqslant 1$ we get homogeneous polynomials $G_i(Y, T) = T^c g_i(Y/T)$, $Y = (Y_1, \ldots, Y_s)$, of positive degree such that $G_i(y, t) = 0$ in A'_q, that is $rG_i(y, t) = 0$ for a certain $r \in A' \backslash q$. Changing y to ry and t by rt we may suppose $G_i(y, t) = 0$ in A'. Let $C = B[Y, T]/(G)$ and $\beta : C \to A'$ be the extension of v by $Y \to y$, $T \to t$. The B_p-morphism $\gamma : D[T, T^{-1}] \to (B_p \otimes_B C)_T$ given by $Z \to Y/T$, $T \to T$ is an isomorphism and the map $\widetilde{\beta} : (B_p \otimes_B C)_T \to A'_q$ induced by β is such that $\widetilde{\beta}\gamma$ extends β_p. Thus we may change D to the smooth A_p-algebra $D[T, T^{-1}]$. Clearly $\beta(H_{C/A}) \not\subset q$ let us say $\beta(c) \notin q$ for a certain $c \in H_{C/A}$.

Suppose $C \cong A[X]/(f)$, $X = (X_1, \ldots, X_n)$, $f = (f_1, \ldots, f_m)$. If $h_B \subset h_C$ then C works. Otherwise, choose $w \in A' \backslash q$ and an integer $k > 0$ such that $w(h_B)^k = 0$ ($ht\ q = 0$!). Let $b_1, \ldots, b_\lambda$ be a system of generators of $H_{B/A}$,

$$E = C[X, U, V, W] \Big/ \left(f_j - \sum_{i=1}^{\lambda} U_i V_{ij}, WU_i \right)_{i,j},$$

where $U = (U_1, \ldots, U_\lambda)$, $V = (V_{ij})$, W are new indeterminantes and $\hat{\beta} : E \to A'$ the map extending β by $U_i \to v(b_i^k)$, $V \to 0$, $W \to w$. Since E_{U_i} is a polynomial ring over $B[U_i, U_i^{-1}]$ it is smooth over B. Then $E_{b_i U_i}$ is smooth over A, thus $b_i U_i \in H_{E/A}$ and so $h_B \subset h_E$. But $E_W \cong C[V, W, W^{-1}]$ is smooth over C and so E_{cW} is smooth over A. Hence $\beta(c)w \in h_E \backslash q$. $\square$

7. Proof of the Main Lemma

The main Lemma will follow from the following two lemmas:

LEMMA 7.1 (Lifting Lemma). *Let $A \to A'$ be a morphism of Noetherian rings, $d \in A$, $\bar{A} := A/(d^2)$, $\bar{A}' := A'/(d^2)$, $\widetilde{A} := A/(d)$, $\widetilde{A}' := A'/(d)$, $\bar{C}$ a finite type $\bar{A}$-algebra and $\bar{\beta} : \bar{C} \to \bar{A}'$ a morphism of $\bar{A}$-algebras. Suppose that* $\mathrm{Ann}_A(d^2) = \mathrm{Ann}_A(d)$, $\mathrm{Ann}_{A'}(d^2) = \mathrm{Ann}_{A'}(d)$. *Then there exist a finite type A-algebra D and a morphism $w : D \to A'$ of A-algebras such that*

(i) *$\widetilde{A} \otimes_{\bar{A}} \bar{\beta}$ factors through $\widetilde{A} \otimes_A w$,*

(ii) *$\pi^{-1}(h_{\bar{C}}) \subset h_D$, π being the surjection $A' \to \bar{A}'$ and $h_{\bar{C}} = \sqrt{\bar{\beta}(H_{\bar{C}/\bar{A}})\bar{A}'}$.*

LEMMA 7.2 (Desingularization Lemma). *Let $A \to A'$ be a morphism of Noetherian rings, B a finite type A-algebra, $v : B \to A'$ an A-morphism, $d \in A$, $\widetilde{A} := A/(d^4)$, $\widetilde{A}' := A'/(d^4)$ and $\widetilde{B} := B/d^4B$. Suppose that d is strictly standard for a presentation of B over A,* $\mathrm{Ann}_A(d^2) = \mathrm{Ann}_A(d)$ *and* $\mathrm{Ann}_{A'}(d^2) = \mathrm{Ann}_{A'}(d)$. *Let D be a finite type A-algebra and $w : D \to A'$ an A-morphism such that $\widetilde{A} \otimes_A v$ factors through $\widetilde{A} \otimes_A w$. Then there exist a finite type A-algebra E and an A-morphism $\gamma : E \to A'$ such that*

(i) *v and w factor through γ,*

(ii) *$H_{D/A}E \subset H_{E/A}$ (so $h_D \subset h_E$).*

Indeed, let $a \in A$, B be as in the Main Lemma and set $d = a^4$. Since

$$\bar{A} \to \bar{B} \to \bar{A}' \supset \bar{q} := q/dA'$$

is resolvable there exist a finite type $\bar{A}$-algebra $\bar{C}$ and a morphism $\bar{\beta} : \bar{C} \to \bar{A}'$ of $\bar{A}$-algebras such that $h_{\bar{B}} \subset h_{\bar{C}} \not\subset \bar{q}$ and $\bar{v} : \bar{B} \to \bar{A}'$ factors through $\bar{\beta}$. By Lemma 7.1 there exist a finite type A-algebra D and a morphism $w : D \to A'$ of A-algebras such that $\widetilde{A} \otimes_{\bar{A}} \bar{\beta}$ factors through $\widetilde{A} \otimes_A w$ and $\pi^{-1}(h_{\bar{C}}) \subset h_D$. Now apply Lemma 7.2 for $d = a$. Thus there exist a finite type A-algebra E and an A-morphism $\gamma : E \to A'$ such that v and w factor through γ and $h_D \subset h_E$. Since $H_{B/A}\bar{B} \subset H_{\bar{B}/\bar{A}}$ by base change we get $h_B\bar{A}' \subset h_{\bar{B}} \subset h_{\bar{C}}$ and so $h_D \supset h_B$. We have also $h_D \not\subset q$ because $h_{\bar{C}} \not\subset \bar{q}$.

PROOF OF THE LIFTING LEMMA. Let

$$H_{\bar{C}/\bar{A}} = \sqrt{\sum \bar{C}\bar{P}_i},$$

where $\bar{P}_i \in \Delta_{\bar{f}^{(i)}}((\bar{f}^{(i)}) : \bar{I})$ for some presentation $\bar{C} = \bar{A}[X]/\bar{I}$, $X = (X_1, \ldots, X_n)$, where each $\bar{f}^{(i)}$ is a finite subset of $\bar{I}$. Let $f_1, \ldots, f_s$ be polynomials in $A[X]$ such that $\bar{f}_j =$

f_j mod d^2, $1 \leqslant j \leqslant s$, generate $\bar{I}$ and $\{\bar{f}_1, \ldots, \bar{f}_s\}$ contains all elements from $\bar{f}^{(i)}$. Set $I = (f_1, \ldots, f_s, d^2)$ and let $\bar{\beta}$ be given by $X \to \bar{x} = x$ mod d^2 for a certain $x \in A'^n$. Then $f(x) = dz$, $z \in dA'^n$. Set $g_j(X, Z) = f_j - dZ_j$, $1 \leqslant j \leqslant s$, $Z = (Z_1, \ldots, Z_s)$ and $g = (g_j)_{1 \leqslant j \leqslant s}$.

To every $\bar{P}_i$ we associate a system of polynomials $F^{(i)}$ in $A[X, Z]$ in the following way: Let P_i be a lifting of $\bar{P}_i$ to $A[X]$. Since $P_i I \subset (d^2, f_1, \ldots, f_r)$ for a certain r in a certain ordering of (f_j) depending on i, it follows that

$$P_i f_j = \sum_{t=1}^{r} H_{jt}^{(i)} f_t + d^2 G_j^{(i)},$$

for some polynomials $H_{jt}^{(i)}$, $G_j^{(i)}$ from $A[X]$, $r < j \leqslant s$. Set

$$F_j^{(i)} = P_i Z_j - \sum_{t=1}^{r} H_{jt}^{(i)} Z_t - dG_j^{(i)}, \quad r < j \leqslant s, \ F^{(i)} = (F_j^{(i)})_{r<j\leqslant s}$$

and $D = A[X, Z]/J$, where $J = (g, F)$, $F = \bigcup_i F^{(i)}$. Note that

$$\begin{aligned} dF_j^{(i)} &= P_i(f_j - g_j) - \sum_{t=1}^{r} H_{jt}^{(i)}(f_t - g_t) - d^2 G_j^{(i)} \\ &= -P_i g_j + \sum_{t=1}^{r} H_{jt}^{(i)} g_t \in (g), \quad r < j \leqslant s. \end{aligned} \tag{1}$$

By construction $J \subset (g) + ((g):d) \cap (Z, d)$ (see (1)) and so $J(x, z) \subset \mathrm{Ann}_{A'} d \cap dA' = 0$. Thus we may define w by $(X, Z) \to (x, z)$, so (i) holds. Using again (1) we see that $D_d \cong A_d[X, Z]/(g) \cong A_d[X]$ is smooth over A. Hence $d \in H_{D/A}$.

It remains to show that $P_i \in H_{D/A}$. Let f stand for the column vector with entries f_j and let $H^{(i)}$ be the $(s \times s)$-matrix $(H_{jt}^{(i)})$, where we set $H_{jt}^{(i)} = 0$ if $t > r$ and $H_{jt}^{(i)} = P_i \delta_{jt}$ if $j \leqslant r$ for some fixed ordering of the f depending on i. Then we may write $P_i f \equiv H^{(i)} f$ mod d^2. Set $S^{(ik)} = P_k H^{(i)} - H^{(k)} H^{(i)}$, $i \neq k$. We have

$$S^{(ik)} f \equiv P_i P_k f - H^{(k)} P_i f \equiv P_i P_k f - P_i P_k f = 0 \bmod d^2.$$

Differentiating it follows that

$$S^{(ik)} \partial f / \partial X_e \in \left(d^2, f\right) \subset (d, g). \tag{2}$$

But P_i is a multiple of a $r \times r$-minor of $(\partial f_j / \partial X_e)_{1 \leqslant j \leqslant r, 1 \leqslant e \leqslant n}$, given by the columns, let us say $e_1, \ldots, e_r$ and $S^{(ik)}$ has the form $(S|0)$, where S denotes the first

r columns of $S^{(ik)}$. Thus there exists an $r \times r$-matrix $T^{(i)}$ over $A[X, Z]$ such that $(\partial f_j/\partial X_{e_q})_{1 \leqslant j,q \leqslant r}\, T^{(i)} = P_i\, \mathrm{Id}_r$. By (2) it follows that

$$P_i S^{(ik)} \equiv 0 \bmod (d, g) \tag{3}$$

and from the definition of $F^{(i)}$ we obtain $H^{(i)} Z \equiv P_i Z \bmod (d, F^{(i)})$. Thus

$$\begin{aligned} S^{(ik)} Z &= P_k H^{(i)} Z - H^{(k)} H^{(i)} Z \equiv P_k P_i Z - P_i H^{(k)} Z \\ &= P_i (P_k Z - H^{(k)} Z) \equiv P_i F^{(k)} \bmod (d, F^{(i)}). \end{aligned} \tag{4}$$

Set $E^{(i)} = (A[X, Z]/(g, F^{(i)}))_{P_i}$. By (3) $S^{(ik)} = 0$ in $E^{(i)}/dE^{(i)}$ and it follows that $F^{(k)} = 0$, $k \neq i$ in $E^{(i)}/dE^{(i)}$ by (4). Thus $F^{(k)} \in dE^{(i)}$. It is enough to show that $E^{(i)}$ is smooth over A. Indeed, if this holds then $E^{(i)}$ is flat over A and we have $\mathrm{Ann}_{E^{(i)}} d = \mathrm{Ann}_{E^{(i)}} d^2$. Then $F^{(k)} \in \mathrm{Ann}_{E^{(i)}} d \cap dE^{(i)} = 0$ (see (1)), that is $D_{P_i} \cong E^{(i)}$. Thus D_{P_i} is smooth over A and so $P_i \in H_{D/A}$. Hence $\pi^{-1}(h_{\bar{C}}) \subset h_D$.

It remains to show that $E^{(i)}$ is smooth over A. Note that by (1) $(g_j)_{r<j\leqslant s}$ are 0 in $N^{(i)} = (A[X, Z]/(g_1, \ldots, g_r, F^{(i)}))_{P_i}$ and so $N^{(i)} \cong E^{(i)}$. By definition of $F^{(i)}$ we have $N^{(i)} \cong (A[X, Z_1, \ldots, Z_r]/(g_1, \ldots, g_r))_{P_i}$. As a minor of $(\partial f/\partial X)$ (and so of $(\partial g/\partial X)$) divides P_i we see that $N^{(i)}$ is smooth over A which is enough. □

LEMMA 7.3. *Let $D \to A'$ be a morphism of Noetherian rings, C a finite type D-algebra, $\alpha : C \to A'$ an A-morphism of algebras, $d \in D$, $\tilde{D} = D/(d^4)$, $\tilde{A}' = A'/d^4 A'$ and $\tilde{C} = C/d^4 C$. Suppose that d is strictly standard for a presentation of C over D and there is a retraction of $\tilde{D}$-algebras $\tilde{\rho} : \tilde{C} \to \tilde{D}$ such that $\tilde{C} \otimes_C \alpha$ is the composite map*

$$\tilde{C} \xrightarrow{\tilde{\rho}} \tilde{D} \to \tilde{A}',$$

the last one being the structure map. Then there exists a finite type D-algebra E such that:

(i) *α factors through E,*
(ii) *the annihilator a in D of* $\mathrm{Ann}_D(d^2)/\mathrm{Ann}_D(d)$ *satisfies $aE \subset H_{E/D}$.*

PROOF. Let $C = D[X]/I$, $X = (X_1, \ldots, X_n)$ be a presentation of C over D for which d is strictly standard. Then there exists a finite system of polynomials $f = (f_i)_{1 \leqslant i \leqslant r}$ in I such that $d \equiv P \bmod I$ for a certain $P \in \Delta_f((f) : I)$ (in $D[X]$). Let $x \in D^n$ be an element such that $\tilde{\rho}$ is given by $X \to x \bmod d^4$. We have $I(x) \equiv 0 \bmod d^4$ and so $P(x) \equiv d \bmod d^4$ in D. Thus $P(x) = ds$ for a certain $s \in D$ with $s \equiv 1 \bmod d$.

Let $P = \sum T_\nu M_\nu$, where the M_ν are the $r \times r$ minors of $(\partial f/\partial X)$. If M_1 is given by the first r columns we have $M_1 = \det H_1$, where

$$H_1 = \begin{pmatrix} \partial f/\partial X \\ 0 \mid \mathrm{Id}_{n-r} \end{pmatrix}.$$

Similarly define the n square matrices H_ν such that $\det(H_\nu) = M_\nu$. Clearly the first r rows of all H_ν coincide. Let G'_ν be the adjoint matrix of H_ν and $G_\nu = T_\nu G'_\nu$. We have

$$\sum G_\nu H_\nu = \sum T_\nu M_\nu \,\mathrm{Id}_n = P\,\mathrm{Id}_n$$

and so

$$ds\,\mathrm{Id}_n = P(x)\,\mathrm{Id}_n = \sum G_\nu(x) H_\nu(x).$$

Let α be given by $X \to y \in A'^n$. By hypothesis we have $x' - y \in d^4 A'^n$, where x' is the image of x by $D \to A'$. Let us say $x' - y = d^3\varepsilon$ for $\varepsilon \in dA'^n$. Then $v^{(\nu)} := H_\nu(x)\varepsilon$ satisfies

$$\sum G_\nu(x) v^{(\nu)} = P(x)\varepsilon = ds\varepsilon$$

and so

$$s(x' - y) = d^2 \sum G_\nu(x) v^{(\nu)}.$$

As the first r rows of H_ν coincide we see that $v^{(\nu)} = (u_1, \ldots, u_r, v_{r+1}^{(\nu)}, \ldots, v_n^{(\nu)})$, where $u_1, \ldots, u_r$ are independent of ν.

Let

$$h = s(X - x) - d^3 W - d^2 \sum G_\nu(x) V^{(\nu)},$$

where $W = (W_1, \ldots, W_n)$ and $V^{(\nu)}$ are variables of the form $(U_1, \ldots, U_r, V_{r+1}^{(\nu)}, \ldots V_n^{(\nu)})$. Clearly h maps to 0 under the map $\phi : D[X, W, V] \to A'$ given by $X \to y$, $W \to 0$, $V^{(\nu)} \to v^{(\nu)}$. Since

$$f(X) - f(x) \equiv \sum_j \partial f/\partial X_j (X_j - x_j)$$

modulo higher order terms in $X_j - x_j$ and

$$s(X - x) \equiv d^3 W + d^2 \sum G_\nu(x) V^{(\nu)} \bmod h$$

we see that for $m = \max\deg f_i$ we have

$$\begin{aligned} s^m f(X) - s^m f(x) \equiv \; & \sum s^{m-1} \partial f/\partial X_j \left[d^3 W_j + d^2 \Big(\sum G_\nu(x) V^{(\nu)} \Big)_j \right] \\ & + d^4 Q' \bmod h \end{aligned}$$

with $Q' \in (W, V)^2 D[W, V]^n$. Thus

$$s^m f(X) - s^m f(x) \equiv s^{m-1} d^2 \sum \partial f/\partial X_j(x) \Big(\sum G_\nu(x) V^{(\nu)} \Big)_j + d^3 Q \bmod h,$$

where $Q \in (\sum DW_j + d(W, V)^2)D[W, V]^n$. But $\sum \partial f_i/\partial X_j(x)(G_\nu(x)V^{(\nu)})_j$ is the i-th entry of

$$H_\nu(x)G_\nu(x)V^{(\nu)} = T_\nu(x)M_\nu(x)V^{(\nu)},$$

that is $T_\nu(x)M_\nu(x)U_i$ because $i \leqslant r$. It follows that

$$s^m f(X) - s^m f(x) \equiv s^m d^3 U + d^3 Q \bmod h.$$

As $f(x) \equiv 0 \bmod d^4$ we have $f(x) = d^3 c$ for a $c \in dA'^r$. Take $g = s^m c + s^m U + Q$ and note that g lies in $D[W, V]$. We have

$$d^3 g \equiv s^m f \bmod h. \tag{$*$}$$

Set $E = D[X, W, V]/(I, g, h) \cong C[W, V]/(g, h)$. We have $d^3\phi(g) = 0$ from $(*)$ and $\phi(g) \in dA'^r$ because $c \in dA'^r$, $v^{(\nu)} \in dA'^n$ and $\phi(Q) \in dA'^r$. Since $\mathrm{Ann}_{A'}(d) = \mathrm{Ann}_{A'}(d^2)$ it follows $\phi(g) = 0$. Thus ϕ factors through E and so α factors too. Note that $E_d \cong C_d[W, V]/(h)$ because $(*)$ shows that g is 0 in E_d. Solving the equations $h = 0$ for the W we see that $E_d \cong C_d[V]$. Thus E_d is smooth over C and so over D because $d \in H_{C/D}$ by hypothesis. Hence $d \in H_{E/D}$.

Let $F = D[X, W, V]/(g, h)$. As $PI \subset (f)$ we get

$$P(x)IF_s \subset PIF_s + (X - x)IF_s \subset (f)F_s + d^2 IF_s$$

because $s(X - x) \in (d^2, h)$ as above. But $P(x) = ds$ and $(f)F_s = 0$ because of $(*)$. Thus $dIF_s = d^2 IF_s$. Then there exists $s' \in 1 + dD[X, W, V]$ such that $s'dIF_s = 0$ and so $dIF_{ss'} = 0$. It follows that $IF_{dss'} = 0$ and so $F_{dss'} \cong E_{dss'}$. As E_d is smooth over D we see that $dss' \in H_{F/D}$.

Next we show that $ss' \in H_{F/D}$. Since $F_s \cong D_s[W, V]/(g)$ solving the equations $h = 0$ in X we see that $\Delta_g \subset H_{F_s/D}$. But $(\partial g/\partial U) \equiv s^m \mathrm{Id}_r \bmod d$ and so $s^{mr} + de \in H_{F_s/D}$ for a certain $e \in F$. Then $ss'(s^{mr} + de) \in H_{F/D}$ and we get $s^{mr+1}s' \in H_{F/D}$ because $dss' \in H_{F/D}$. Thus $ss' \in H_{F/D}$. Since $h = 0$ in F_s we note that $X \equiv x \bmod d^2$ and so $IF_s \equiv I(x)F_s \equiv 0 \bmod d^2$. Thus $IF_s \subset d^2 F_s$, let us say $IF_s = db$ for some ideal b of F_s. As above $dIF_{ss'} = 0$ and so $d^2 bF_{ss'} = 0$. But $F_{ss'}$ is smooth over D, in particular flat. It follows that $aF_{ss'}$ is the annihilator in $F_{ss'}$ of $\mathrm{Ann}_{F_{ss'}}(d^2)/\mathrm{Ann}_{F_{ss'}}(d)$. Thus $dabF_{ss'} = 0$ and so $aIF_{ss'} = 0$. Hence $F_{ss't} = E_{ss't}$ for all $t \in a$. Since $F_{ss'}$ is smooth over D we get $ss't \in H_{E/D}$. As $ss' \equiv 1 \bmod d$ and $d \in H_{E/D}$ it follows $t \in H_{E/D}$ for all $t \in a$, that is $aE \subset H_{E/D}$. □

PROOF OF THE DESINGULARIZATION LEMMA. Let $C = B \otimes_A D$ and $\alpha : C \to A'$ being given by $b \otimes z \to v(b) \otimes w(z)$. By hypothesis there exists $\tilde{\tau} : \tilde{B} \to \tilde{D} := D/d^4 D$ such that $\tilde{A} \otimes_A w = (\tilde{A} \otimes_A w)\tilde{\tau}$. Then the map $\tilde{\rho} : \tilde{C} \to \tilde{D}$ given by $\tilde{b} \otimes \tilde{z} \to \tilde{\tau}(\tilde{b}) \otimes \tilde{z}$ is clearly a retraction of $\tilde{D}$-algebras such that $\tilde{C} \otimes \alpha$ is the composite map

$$\tilde{C} \xrightarrow{\tilde{\rho}} \tilde{D} \xrightarrow{\tilde{A}\otimes w} \tilde{A}'.$$

Applying Lemma 7.4 to the case $D \to C \xrightarrow{\alpha} A'$, $\tilde{\rho}$ we find a finite type D-algebra E such that

(1) α factors through E (so v and w factor too),

(2) the annihilator a in D of $\mathrm{Ann}_D(d^2)/\mathrm{Ann}_D(d)$ satisfies $aE \subset H_{E/D}$.

Let $t \in H_{D/A}$. Then D_t is smooth over A, in particular flat. Thus $\mathrm{Ann}_{D_t}(d^2) = \mathrm{Ann}_{D_t}(d)$ and so $aD_t = D_t$. By 2) it follows E_t smooth over D_t. Hence E_t is smooth over A and so $H_{D/A}E \subset H_{E/A}$. □

References

[1] M. André, *Homologie des Algèbres Commutatives*, Springer, Berlin (1974).

[2] M. André, *Localisation de la lissité formelle*, Manuscripta Math. **13** (1974), 297–307.

[3] M. André, *Cinq exposés sur la desingularization*, Handwritten manuscript, École Polytechnique Fédérale de Lausanne (1991).

[4] M. Artin, *On the solution of analytic equations*, Invent. Math. **5** (1968), 277–291.

[5] M. Artin, *Algebraic approximation of structures over complete local rings*, Publ. Math. IHES **36** (1969), 23–58.

[6] M. Artin, *Construction techniques for algebraic spaces*, Actes Congres Intern. Math. **1** (1970), 419–423.

[7] M. Artin, *Versal deformations and algebraic stacks*, Invent. Math. **27** (1974), 165–189.

[8] S. Basarab, V. Nica and D. Popescu, *Approximation properties and existential completeness for ring morphisms*, Manuscripta Math. **33** (1981), 227–282.

[9] J. Becker, *A counterexample to Artin approximation with respect to subrings*, Math. Ann. **230** (1977), 195–196.

[10] J. Becker, J. Denef, L. Lipshitz and L. van den Dries, *Ultraproducts and approximation in local rings I*, Invent. Math. **51** (1979), 189–203.

[11] A. Brezuleanu and N. Radu, *Sur la localisation de la lissité formelle*, C. R. Acad. Sci. Paris, Sér. A **276** (1973), 439–441.

[12] M.L. Brown, *Artin approximation property*, Thesis, Cambridge Peterhouse (August 1979).

[13] M. Cipu and D. Popescu, *Some extensions of Néron's p-desingularization and approximation*, Rev. Roum. Math. Pures Appl. **24** (10) (1981), 1299–1304.

[14] J. Denef and L. Lipshitz, *Ultraproducts and approximation in local rings II*, Math. Ann. **253** (1980), 1–28.

[15] R. Elkik, *Solutions d'equations à coefficients dans une anneau (!!) henselien*, Ann. Sci. École Norm. Sup. 4^e **6** (1973), 533–604.

[16] A.M. Gabrielov, *The formal relations between analytic functions*, Funktsional. Anal. i Prilozhen. **5** (4) (1971), 64–65 (in Russian).

[17] H. Grauert, *Über die Deformation isolierter Singularitäten analytischer Mengen*, Invent. Math. **15** (3) (1972), 171–198.

[18] M. Greenberg, *Rational points in Henselian discrete valuation rings*, Publ. Math. IHES **31** (1966), 59–64.

[19] A. Grothendieck and J. Dieudonné, *Éléments de Géometrie Algébrique*, IV, Parts 1 and 4, Publ. Math. IHES (1964).

[20] A. Grothendieck, *Technique de descente II*, Sém. Bourbaki **195** (1959–1960).

[21] B. Iversen, *Generic Local Structure in Commutative Algebra*, Lecture Notes in Math., Vol. 310, Springer, Berlin (1973).

[22] H. Kurke, G. Pfister and M. Roczen, *Henselsche Ringe and algebraische Geometry*, VEB, Berlin (1975).

[23] H. Kurke, T. Mostowski, G. Pfister, D. Popescu and M. Roczen, *Die Approximationseigenschaft lokaler Ringe*, Lecture Notes in Math., Vol. 634, Springer, Berlin (1978).

[24] H. Lindel, *On the Bass–Quillen conjecture concerning projective modules over polynomial rings*, Invent. Math. **65** (1981), 319–323.

[25] H. Matsumura, *Commutative Ring Theory*, Cambridge University Press (1986).

[26] A. Néron, *Modèles Minimaux des Variétés Abéliennes sur les Corps Locaux et Globaux*, Publ. Math. IHES **21** (1964).

[27] T. Ogoma, *General Néron desingularization based on the idea of Popescu*, J. Algebra **167** (1994), 57–84.
[28] G. Pfister and D. Popescu, *Die strenge Approximationseigenschaft lokaler Ringe*, Invent. Math. **30** (1975), 145–174.
[29] D. Popescu, *Algebraically pure morphisms*, Rev. Roum. Math. Pures Appl. **26** (6) (1979), 947–977.
[30] D. Popescu, *General Néron desingularization*, Nagoya Math. J. **100** (1985), 97–126.
[31] D. Popescu, *General Néron desingularization and approximation*, Nagoya Math. J. **104** (1986), 85–115.
[32] D. Popescu, *Polynomial rings and their projective modules*, Nagoya Math. J. **113** (1989), 121–128.
[33] D. Popescu, *Letter to the Editor. General Néron desingularization and approximation*, Nagoya Math. J. **118** (1990), 45–53.
[34] M. Raynaud, *Anneaux Locaux Henséliens*, Lecture Notes in Math., Vol. 169, Springer, Berlin (1970).
[35] C. Rotthaus, *Rings with approximation property*, Math. Ann. **287** (1990), 455–466.
[36] M. Spivakovsky, *Non-existence of the Artin function for Henselian pairs*, Math. Ann. **299** (1994), 727–729.
[37] M. Spivakovsky, *A new proof of D. Popescu's theorem on smoothing of ring homomorphisms*, J. Amer. Math. Soc. **294** (1999), 381–444.
[38] R. Swan, *Néron–Popescu desingularization*, Algebra and Geometry, National Taiwan University, 26–30 December 1995, Ming-chang Kang, ed., International Press, Cambridge (1998), 135–192.
[39] B. Teissier, *Résultats récents sur l'approximation des morphisms en algèbre commutative [d'après Artin, Popescu, André, Spivakovsky]*, Séminaire Bourbaki **784** (1994), 1–15.
[40] M. Van der Put, *A problem on coefficient fields and equations over local rings*, Compositio Math. **30** (3) (1975), 235–258.
[41] J.J. Wavrik, *A theorem on solutions of analytic equations with applications to deformations of complex structures*, Math. Ann. **216** (2) (1975), 127–142.

Section 3B
Associative Rings and Algebras

This Page Intentionally Left Blank

Fixed Rings and Noncommutative Invariant Theory

V.K. Kharchenko*

Universidad Nacional Autónoma de México, FES Cuautitlán, and Sobolev Institute of Mathematics, Novosibirsk, Russia

E-mail: vlad@servidor.unam.mx

Contents

*Supported by SNI, exp 18740 and CONACyT, 32130-E, México.

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

1. Existence of fixed elements

The theory of fixed rings deal with the relationships between a given ring R acted on by a finite group G of its automorphisms and the fixed ring $R^G = \{r \in R \mid \forall g \in G,\ r^g = r\}$. The simplest (and very important) relation is given by the *trace function* $t(x) = \sum_{g \in G} x^g$. It is easy to see that t is an R^G-bimodule homomorphism from R to R^G. The first question that arises is whether t is nonzero or, more generally, does there exists a nonzero fixed element at all?

The modern development of the theory was inspired by a small paper by C. Faith [Fa72], where, using non-commutative Galois theory for skew fields, he had proved that the fixed ring of an Ore domain is an Ore domain, and by G. Bergman's paper [Be71] where the question answered by C. Faith was put. A fundamental result was obtained in the Bergman and Isaacs paper [BI73]. In this paper the existence problem was formulated as follows.

Let R be a ring (without 1) and suppose that G is a finite group acting by automorphisms on R, such that $R^G = \{0\}$. If $R \neq \{0\}$, must then R have zero-divisors?

G. Bergman and I. Isaacs answered this question in the case when R has no additive $|G|$-torsion. Their answer was much more than was expected.

THEOREM 1.1. *Let G be a finite group acting on the ring R without $|G|$-torsion. If $R^G = \{0\}$ then $R^{f(|G|)} = \{0\}$. If R^G (or, more generally, $t(R)$) is nilpotent, then R is nilpotent. Here*

$$f(n) = \prod_{1 \leqslant m \leqslant n} \left(C_n^m + 1\right).$$

where C_n^m is the binomial coefficient $\binom{n}{m}$.

This result remained the basic tool of the growing theory for a long time. For instance it allows one to show that *in a semiprime ring without additive $|G|$-torsion every invariant nonzero one-sided ideal I has nonzero intersection with the fixed ring.* (For the notion of prime and semiprime ring see the paper "Simple, prime and semiprime rings" in Vol. 1 of the Handbook of Algebra, [Kh95]).

Indeed, if the intersection is zero then the ideal considered as the ring would be nilpotent $I^n = \{0\}$. Therefore the two-sided ideal generated by I is nilpotent of degree $n + 1$. This is impossible in a semiprime ring.

Another important corollary says that *the fixed ring of semiprime ring without $|G|$-torsion is semiprime.*

Indeed, let I be a nonzero nilpotent ideal of the fixed ring. The left ideal RI is invariant with respect to G. For arbitrary fixed element $a = \sum r_k i_k \in RI$ we have

$$|G|a = t(a) = \sum_{g \in G} a^g = \sum_{g \in G} \left(\sum r_k i_k\right)^g = \sum t(r_k) i_k \in R^G I \subseteq I.$$

This implies that the fixed ring of the ring RI is nilpotent. By the Bergman–Isaacs theorem RI is nilpotent, which is impossible in a semiprime ring.

A third important corollary uses the bound of nilpotency in the Bergman–Isaacs theorem.

THEOREM 1.2. *Let G be a finite group acting on a prime ring R without $|G|$-torsion. The fixed ring is a subdirect product of not more then $f(|G|)$ prime rings. In particular, the fixed ring has no direct sums of more than $f(|G|)$ nonzero two-sided ideals.*

Let us prove the second part of the theorem. If $I_1 \oplus \cdots \oplus I_{f(n)+1}$ is a direct sum of $f(n)+1$ nonzero ideals of the fixed ring then the set $S = \sum_{k<m} I_k R I_m$ is an invariant subring of R. Note that $I_k I_m = 0$ provided $k \neq m$ since the sum is direct. It is clear that $S^{f(n)+1} = 0$ while $S^{f(n)} \neq 0$ because in the prime ring R we have

$$S^{f(n)} \supseteq I_1 R I_2^2 R I_3^2 \dots I_{f(n)}^2 R I_{f(n)+1} \neq 0.$$

This inequality contradicts the Bergman–Isaacs theorem. Indeed, each fixed element of the ring S has the form

$$a = \sum_{k<m} \sum_{s} i_k^{(s)} r^{(s)} i_m^{(s)};$$

while the inclusions $t(r^{(s)}) \in R^G$ show that

$$na = t(a) = \sum_{k<m} \sum_{s} i_k^{(s)} t(r^{(s)}) i_m^{(s)} = 0.$$

Consequently $S^G = \{0\}$, and $S^{f(n)} = 0$. □

This theorem was proved first in a little more general form [Kh74]: *if the ring R (without $|G|$-torsion) is a subdirect product of m prime rings than R^G is a subdirect product of not more then $f(|G|)m$ prime rings.* The proof of this fact is based on the Bergman–Isaacs theorem in the way showed above. Later in James Osterburg paper [Os79] it was shown that the bound $f(|G|)m$ can be improved to $|G|m$. This fact has been proved without using the Bergman–Isaacs theorem but with the help of Galois theory concept of the *algebra* $\mathbf{B}(G)$ *of a group* G (see Definition 3.5.1 [Kh95]). If the ring R has no additive $|G|$-torsion and is a subdirect product of finite number of prime rings than the algebra $\mathbf{B}(G)$ of the group G is finite-dimensional semisimple algebra with an action of the group G. In particular it has a unique presentation as a direct sum of its minimal invariant two-sided ideals

$$\mathbf{B}(G) = \mathbf{B}_1 \oplus \mathbf{B}_2 \oplus \cdots \oplus \mathbf{B}_s.$$

The minimal number of prime subdirect cofactors (the *prime dimension*) of the fixed ring is equal to the number s in this decomposition (see also [Kh91], Theorem 3.6.7 for prime case and Theorem 5.5.10 in general).

The question itself that Bergman and Isaacs started with, has been answered later in a little bit more general form by the author [Kh75].

THEOREM 1.3. *Let R be a ring (without 1) and suppose that G is a finite group acting by automorphisms of R, such that $R^G = \{0\}$. If $R \neq \{0\}$, then R has nonzero nilpotent elements.*

For some time it looked natural to think that the ring in previous theorem should have a nonzero nilpotent ideal. The author's attempt to prove it [Kh75a] was killed by the following example of G. Bergman:

EXAMPLE 1.4. Let F be a field of characteristic $p \neq 0$, with an element $w \neq 0, 1$ of finite multiplicative order, say $w^n = 1$. Let S be the algebra of two by two matrices over the free algebra $F\langle x, y\rangle$ in two variables x, y. Let the group G be generated by the inner automorphisms induced by matrices $A = E + xe_{12}$, $B = E + ye_{12}$, $C = we_{11} + e_{22}$; where the e_{ij} are the matrix units (standard elementary matrices) and $E = e_{11} + e_{22}$ is the unit element of S. Then $|G| = np^2$, since A and B generate an Abelian normal subgroup H of order p^2. Now let R be the subring of S of those matrices whose entries have zero constant term. R is certainly prime (in particular it has no nonzero nilpotent ideals), and $R^G = \{0\}$. This last fact can be seen as follows: check that the centralizer of the matrices A, B is equal to $F\langle x, y\rangle e_{12}$ and note that the only matrices commuting with C must be diagonal matrices.

Since rings with no nilpotent elements always have fixed elements, we turn to the question of when the trace function is nontrivial. That this is not always the case is seen by the following example, due to R. Snider:

EXAMPLE 1.5. Let D be any division algebra of characteristic 2 which is four-dimensional over its center Z. Choose $x \in D$, $x \notin Z$ with $x^2 \in Z$. Then $(1 + x)^2 = 1 + x^2 \in Z$ also. Now let g be conjugation by x, and let h be conjugation by $1 + x$. Then if $G = \langle g, h\rangle$, G is a group of automorphisms of D, and $G = Z_2 \times Z_2$. However, it is not difficult to check that $t \equiv 0$.

Even when the trace is trivial, it is still possible to construct "partial trace" functions. For subset $\Lambda \subset G$, we let $t_\Lambda(x) = \sum_{g\in\Lambda} x^g$; when $\Lambda = G$, $t_\Lambda(x)$ is just the trace. We say that $t_\Lambda(x)$ is *non-trivial partial trace function* iff $\{0\} \neq t_\Lambda(R) \subseteq R^G$.

Such functions were shown to exist for division rings by D. Farcas and R. Snider in [FS77] and for domains (i.e. for rings without zero divisors) by S. Montgomery in [Mo79].

It seems natural at this point to ask if partial trace functions exist for every ring with no nilpotent elements. That is indeed the case if G is a solvable group (the result is due to M. Cohen and S. Montgomery). The first counter example was due to John Wilson, who has shown that for each finite non-Abelian simple group G, there exists a commutative ring R, with no nilpotent elements, such that no non-trivial partial trace functions exist for G on R. More generally, Isaacs and Passman have shown that if G is non-solvable group, there exists a finite sum of copies of Z_2 on which G acts with no non-trivial partial trace functions (for details see the Lecture Notes by S. Montgomery [Mo80]).

Even when no partial trace functions exist, there still arises the question: is it always possible to find a mapping $\tau : R \to R^G$ which is a non-trivial R^G-bimodule homomorphism?

This question has been considered in a number of papers [Kh75a,Kh77,Kh81,Pc82, CM82,GPV82,MP84] dealing with the Galois theory of prime and semiprime rings. The answer essentially depends on the structure of the algebra $\mathbf{B}(G)$ of the group G. Nevertheless if the ring R has no nilpotent elements then there exists such a mapping τ defined on an essential (i.e. with zero annihilator) two-sided ideal of the ring R.

The same mapping τ exists also for a semiprime ring R provided that the algebra $\mathbf{B}(G)$ is semiprime [Kh75a,Kh77]. It was noted firstly by Goursaud, Pascaud and Valette in [GPV82] that in this case the module of all such mappings given by forms

$$\tau(x) = \sum_i b_i x^g a_i, \quad a_i, b_i \in \mathbf{B}(G),$$

is principal. This means that there exists a *principal trace form* Tr such that every trace form has a presentation $\tau(x) = \mathrm{Tr}(ebxf)$, where e, f are idempotents from $\mathbf{B}(G)$ and b is an invertible element from $\mathbf{B}(G)$.

If the algebra $\mathbf{B}(G)$ is not semiprime then for the construction of trace forms it is necessary to find left and right conjugated ideals of $\mathbf{B}(G)$. The existence problem for trace forms was investigated in detail in the author's paper [Kh81] (see also [Kh91, Ch. 3, pp. 157–163]).

2. Schelter integrality

Let R be a ring (possibly without 1) and let T be a subring. If $r_1, r_2, \dots, r_n \in R$, then a *T-monomial* in the r_i's is a product of these elements in some order with elements of T such that at least one factor from T occurs. For example, if $t_1, t_2 \in T$, then $r_1 t_1 r_2 t_2^3 r_1 r_3$ is a T-monomial but $r_1^2 r_2 r_1 r_3$ is not unless $1 \in T$. The *degree* of such a monomial is the total number of the r_i's. We say that R is *fully integral* over T of degree m if for arbitrary $r_1, r_2, \dots, r_m \in R$ we have

$$r_1 r_2 \dots r_m = \phi(r_1, r_2, \dots, r_m),$$

where ϕ is a sum of T-monomials in the r_i's of degree less than m. In particular, setting $r_1 = r_2 = \cdots = r_m = r \in R$, we see that $r^m = \xi(r)$ where ξ is a sum of T-monomials in the r's of degree less than m. This says, by definition, that R is *Schelter integral* over T of degree m.

The last definition was inspired by the classical notion of integrality in the commutative theory and is equivalent to it for commutative rings. If the ring R is commutative and G is a finite automorphisms group then the polynomial in one variable $\xi(t) = \prod_{g \in G}(t - r^g)$ has fixed coefficients – they are symmetric functions in the r^g, $g \in G$. Evidently $\xi(r) = 0$. Therefore R is integral over the fixed ring R^G.

If R is fully integral (non-commutative) over R^G of degree m and additionally $R^G = 0$, then $R^m = 0$. Thus this type of integrality over fixed ring would be an appropriate generalization of the Bergman–Isaacs theorem.

As a simple example in a noncommutative situation, let $G = \{1, x\}$ have order 2 and act on R. If $r \in R$, then $t = r + r^x = \mathrm{tr}_G(r) \in R^G$ and $t - r = r^x$. Thus

$$r(t-r) + (t-r)r = rr^x + r^x r = s = \mathrm{tr}_G\bigl(rr^x\bigr) \in R^G.$$

In other words, r satisfies $2r^2 - rt - tr + s = 0$, and $2r$ is Schelter integral over R^G. The following theorem of D. Quinn [Qu89] shows that this is the case in the general situation also.

THEOREM 2.1. *Let G be a finite group acting on the ring R (without 1) and suppose that $|G| \cdot R = R$. Then R is fully integral of degree $m(|G|)$ over the fixed ring R^G.*

The proof of this theorem is based on the following modification by D. Passman [Pm81] of the Paré–Schelter result [PS78].

THEOREM 2.2. *Let A be a ring with 1 and let $S \supseteq T$ be subrings (without 1) of the $n \times n$ matrices ring over A. Assume that for all $k = 1, 2, \dots, n$ the matrix units $e_{k,k}$ satisfy*

(1) $(\sum_1^k e_{i,i})S(\sum_1^k e_{i,i}) \subseteq S$,

(2) T consists of diagonal matrices and $e_{k,k}Se_{k,k} = e_{k,k}Te_{k,k}$.

Then there exists an integer $m = m(n)$ such that R is fully integral over T of degree m.

This theorem implies immediately that if a group $G = \{g_1, \dots, g_n\}$ of order n acts on the ring R then the ring S of $n \times n$ matrices over R is fully integral over the subring of diagonal matrices $T = \{\mathrm{diag}(a^{g_1}, \dots, a^{g_n}) \mid a \in R\} \simeq R$. This fact allows one to obtain the Quinn theorem in the following way.

Let us suppose for the simplicity that the ring R has no n-torsion. In this case the multiplication of matrices from S by the formal matrix e with all entries equal to $1/n$ is defined. We have $e^2 = e$ both as matrix and as an operator. Let $s_1, \dots, s_n$ be arbitrary matrices from S. The integer dependence of the elements $es_1e, \dots, es_ne$ over T has the form

$$es_1e \dots es_ne = \sum t_{i1}es_{j_1}e \dots es_{j_{m-1}}et_{im}.$$

Multiplying this equality from the right and from the left by e, we see that the ring eSe is fully integral over the subring eTe. The ring eSe consists of the matrices, all coefficients of which are the same, while the ring eTe consists of the matrices, all the coefficients of which are fixed relative to G and equal to each other. Bearing in mind the fact that the product of the matrices $\|s\| \cdot \|v\|$ has the form $\|nsv\|$, and the possibility to cancel by n, we obtain the conclusion that R is fully integer over the fixed ring R^G.

The detailed proofs can be found in the book [Pm89] by D. Passman, pp. 254–264.

Theorem 2.1 has been proved first for Abelian groups by D. Passman [Pm81] and by S. Montgomery and L. Small [MS84] for PI-rings R.

3. Shirshov finiteness

It is well known in the classical Galois theory that a (commutative) field F is always a finite-dimensional space over fixed subfields F^G (for finite G). This fact is still valid for skew fields and, more generally, for simple Artinian rings A (e.g. for matrix rings over skew fields) under the condition that there is no additive $|G|$-torsion. If A is a skew field then both the left and right dimensions are finite. For simple Artinian rings finite dimension means that both the right, A_{A^G}, and the left, ${}_{A^G}A$, modules over R^G are finitely generated [Az46,FS77]. However if the ring R is not simple, this need not be the case.

EXAMPLE 3.1. Let $R = F\langle x, y\rangle$ be the free algebra on two generators over a field F of characteristic unequal to 2, and let ξ be the automorphism of order 2 determined by $x^\xi = -x$ and $y^\xi = y$. The group G generated by this automorphism has two elements $1, \xi$. The fixed ring is spanned by all monomials containing x an even number of times. Suppose R is finitely generated as a left R^G-module. Then for some n we have $R = \sum_w R^G w$ where w runs over all monomials of degree $\leqslant n$. But xy^n cannot belong to the right-hand side, thus we have a contradiction.

Nevertheless some kind of finiteness relations of the ring R with the fixed ring R^G still exists. This kind of relations is the second important tool in the modern theory of fixed rings (the first one consists of the Bergman–Isaacs theorem and was mentioned above). The following definitions are motivated by the Shirshov definition of finiteness over a central subring in the theory of PI-rings [Sh57].

DEFINITION 3.2. Let S, T be subrings of the ring R. The subring S is called *left Shirshov finite* over T in R if there exist elements $r_1, r_2, \ldots, r_m \in R$, such that $S \subseteq Tr_1 + Tr_2 + \cdots + Tr_m$.

DEFINITION 3.3. An element $a \in R$ is said to be *left finite* over the subring T in R if the principal left ideal of R generated by a is left Shirshov finite over T in R.

DEFINITION 3.4. Let T be a subring of the ring R. A two-sided ideal I of R is called *left locally finite* over T in R if every finitely generated left ideal J of R contained in I is left Shirshov finite over T in R, that is $J \subseteq Tr_1 + Tr_2 + \cdots + Tr_m$.

It is easy to see that the set of all elements in R that are left finite over T is a two-sided left locally finite over T ideal. Indeed, if a, b are left finite over T elements, say $Za + Ra \subseteq Tr_1 + Tr_2 + \cdots + Tr_m$ and $Zb + Rb \subseteq Tr_{m+1} + Tr_{m+2} + \cdots + Tr_s$, then the left ideal generated by a and b is contained in $\sum_{i \leqslant s} Tr_i$ and as well for each $r \in R$ the left ideal generated by ar is contained in $\sum_{i \leqslant m} Tr_i r$. This implies that each element of the two-sided ideal generated by a and b is left finite over T in R.

Thus, there exists the *largest* left locally finite over T in R ideal which we will denote by $Sh_R(T)$.

THEOREM 3.5. *If the non-zero ring R either is semiprime and has no additive $|G|$-torsion or has no non-zero nilpotent elements then $Sh_R(R^G) \neq 0$. Moreover, in these cases both the left and right annihilators of $Sh_R(R^G)$ in R are zero.*

The proof of this theorem is based on the construction of the trace forms. Starting with any trace form

$$\sum_i b_i x^g a_i = \tau(x), \quad a_i, b_i \in \mathbf{B}(G),$$

one can multiply this equality on the right by an arbitrary element s_j and substitute $x = zq_j$. Summation over j and using similar tricks with other trace forms makes it possible (of course not without serious effort) to find a formula of the type

$$za = \sum_k \tau_k(zd_k)s_k \tag{1}$$

with non-zero a and some (different or not) trace forms τ_k, such that $\tau_k(Rd_k) \subseteq R^G$. Now it is easy to see that the element a is finite over R^G in R. The proof shows in fact that the set of all the elements a in formulae of the type (1) is a two-sided ideal with zero annihilators.

The two cases considered in this theorem can be united by the Galois theory concept of an *M-group*. This is a finite (or, more generally, reduced-finite) group with semiprime algebra $\mathbf{B}(G)$. For M-groups Theorem 3.5 is still valid as well as for the general class of (reduced-)finite groups with quasi-Frobenius algebra (see [Kh91], Theorems 3.5.3, 5.10.1).

COROLLARY 3.6. *Let R be a finite direct sum of simple rings with 1 and G a (reduced-) finite group with semisimple (or, more generally, quasi-Frobenius) algebra $\mathbf{B}(G)$. Then R is a projective left R^G-module of finite rank.*

Indeed, by Theorem 3.5 the set of all a's for which a formula (1) is valid constitutes a two-sided ideal of R with zero annihilator. This is possible in a finite sum of simple rings only if this ideal is equal to R. For $a = 1$ we have

$$z = \sum_k \tau_k(zd_k)s_k, \tag{2}$$

and therefore R is generated by the s_k. In this formula $f_k(-) = \tau_k(-d_k) \in \mathrm{Hom}_{R^G}(R, R^G)$, and thus this implies the projectivity of R over R^G. □

Another interesting corollary of Shirshov finiteness is concerned with so called *biregular* rings. This is a ring R, such that every principle two-sided ideal (a) is generated by a central idempotent, $(a) = (e)$. Of course every direct sum as well as infinite direct product of simple rings with 1 is biregular. A biregular ring has a large enough center $C = C(R)$. Therefore essential is the structure of its ring over C. For instance, in the previous corollary the center is a direct sum of fields, in particular, we see that R is an injective C-module. Now we can formulate an "infinite product" variant of the latter corollary.

THEOREM 3.7. *Let R be an unitary biregular ring injective over its center. If G is a finite group with semisimple algebra $\mathbf{B}(G)$ then R is a projective left R^G-module of finite rank.*

The proof of this theorem can be obtained as an application of a metatheorem (see [Kh95] for the definitions). First of all the injectivity of R over its center warrants that R is a ring of global sections of the canonical (and also invariant) sheaf, while the biregularity implies that almost all the stalks of the invariant sheaf are finite direct sums of simple rings (see [Kh95], Theorem 3.5.11). Then, by Corollary 3.6, formula (2) is valid on these stalks. This formulae can be written in elementary logical language as follows

$$\exists d_1, \dots, d_n, s_1, \dots, s_n \forall z \quad z = \sum_k \tau_k(z d_k) s_k. \tag{3}$$

It can be seen that the number n in this formula depends only on the (reduced)-order of the group G. If we consider the trace forms in formula (3) like a sheaf operations than the metatheorem ([Kh95], Theorem 3.4.5) can be applied to it. This yields that formulae (3) is valid on the ring R, which gives that R is a projective R^G-module of finite rank. □

This theorem under the additional restriction that R is a self-injective ring as well as Corollary 3.6 was first obtained by D. Handelman and G. Renault [HR80] for a finite group of invertible order with the help of the skew group ring method. A little earlier in [Kh77, Theorem 10], the corollary had been partially proved (without any information about projectivity) for M-groups by using Shirshov finitness in the way presented here. In [Re79] the same result has been obtained by G. Renault for a ring R of type $R = \prod_{i=1}^{s}(A_i)_{n_i \times n_i}$ where the A_i are self-injective rings with no nilpotent elements (it is easy to see that all rings of this type are biregular and self-injective).

The notion of Shirshov finiteness was introduced in [Kh76] and plays important role in the Galois theory of semiprime rings [Kh77]. It is also appeared in a paper by J.M. Goursaud, J.L. Pascaud and J. Valette [GPV83, Prop. 22].

In the general case, Theorem 3.5 allows one to find a more or less convenient local construction of the ring R with the help of matrix rings over the fixed ring. This can be done in two steps.

First, every finitely generated left ideal $J \subseteq Sh_R(R^G)$ is a homomorphic image of a certain subalgebra S of the matrix ring $(R^G)_{m\times m}$. Indeed, if $J \subseteq R^G r_1 + R^G r_2 + \cdots + R^G r_m = M$, then each right multiplication $\rho(j) : m \to mj$ is a homomorphism of the left R^G-module $M + R^G$. Therefore ρ is an embedding of J into a ring of endomorphisms of the finitely generated R^G-module $M + R^G$. The endomorphism ring of finitely generated left module is a homomorphic image of a subring of the endomorphism ring of a free module. Namely, an endomorphism ξ of the module with generators $m_1, \dots, m_k$, say $\xi(m_i) = \sum r_i m_i$, can be extended to the endomorphism of a covering free module $\sum R^G f_i$ by the formula $\xi(f_i) = \sum r_i f_i$. Finally it is enough to remember that the endomorphism ring of a free R^G-module is isomorphic to a matrix ring over R^G.

In this way the ideal $Sh_R(R^G)$ is locally constructed from the fixed ring. The second step is to use the symmetric Martindale ring of quotients. Since the annihilators of the ideal $Sh_R(R^G)$ in R are zero, we have $Q(Sh_R(R^G)) = Q(R) \supseteq R$.

At the end of this section let us consider so called G-Galois extensions. Recall that a unitary ring R is said to be a *G-Galois extension* of the fixed ring if there exist elements $x_1, \dots, x_n; y_1, \dots, y_n$ such that $\sum x_i y_i^g = \delta_1^g$, where δ_1^g is the Kronecker delta.

In this case we have $r = \sum x_i t(y_i r)$, where $t(a) = \sum a^g$ is the trace function. Thus any G-Galois extension is a projective module of finite rank.

Note that the definition of G-Galois extension depended only on the existence of a finite number of elements. This implies that if a ring R is a subring of another one, S, with an action of the same group then S is a G-Galois extension also. It is well known that every Galois extension of field is G-Galois (see [Kh95, formula (11), page 777]). Therefore, we have a corollary:

COROLLARY 3.8. *If an unitary ring R is acted on by a finite group G and it has a subfield F such that the restriction of G to F is isomorphic to G then R is projective R^G-module of finite rank.*

4. Goldie rings

Recall that a ring R is called a *left Goldie ring* if it satisfies the ascending chain condition for left annihilators, and has no infinite direct sums of nonzero left ideals. By the Johnson theorem for semiprime rings the ascending chain condition in this definition can be replaced with nonsingularity: each essential left ideal should have zero right annihilator. Recall that a left ideal is called *essential* if it has a non-zero intersection with each non-zero left ideal.

In the Goldie theory the notion of a *regular element* plays important part. That is an element $r \in R$, such that $sr \neq 0$, $rs \neq 0$ for every non-zero $s \in R$. Important are the facts that every *essential* left ideal of a semiprime left Goldie ring has a regular element and every left ideal having a regular element is essential. The Goldie theorem claims that a ring R has semisimple Artinian classical ring of quotients $Q_{\mathrm{ct}}(R)$ if and only if it is a semiprime left Goldie ring.

The relation of Goldie rings with fixed ring theory is given by the following theorem.

THEOREM 4.1. *Let G be an M-group of automorphisms of a semiprime ring R. The fixed ring R^G is a left Goldie ring iff R is a left Goldie ring. In this case $Q_{\mathrm{cl}}(R)^G = Q_{\mathrm{cl}}(R^G)$. Moreover for the calculation of the ring of quotients it is enough to inverse only the set T of fixed regular elements $Q_{\mathrm{cl}}(R) = T^{-1}R$.*

One side of this theorem for Ore domains has been proved by C. Faith [Fa72]. For rings R with no additive $|G|$-torsion it has been obtained independently by M. Cohen [Co75], J. Fisher and J. Osterburg [FO78], and the author [Kh74]. The case of rings with no non-zero nilpotent elements has been considered in the author's paper [Kh75]. For the case of so called X-outer groups (that is when the algebra of the group is equal to the extended centroid of the ring R) it was proved by S. Montgomery [Mo78]. In the general form presented here the theorem was proved in the author paper [Kh77] by means of M. Cohen ideas. Y. Miyashita [Mi67] has a version involving an infinite group G. See also Tominaga [To73].

We are going to present here a very short proof of the first part of the theorem for the case when R has no $|G|$-torsion which is based on the Bergman–Isaacs theorem and on Shirshov finiteness.

If R is semiprime Goldie ring then the fixed ring (as well as any other subring) satisfies the ascending chain conditions on annihilators and it is enough to show that the Goldie dimension of R^G is finite.

If A, B are left ideals of R^G and $A \cap B = \{0\}$, then $t(RA \cap RB) \subseteq t(R)A \cap t(R)B = 0$. By the Bergman–Isaacs theorem we get $RA \cap RB = 0$. This remark shows that if $A_1 \oplus \cdots \oplus A_k \oplus \cdots$ is a direct sum of left ideals in R^G, then $RA_1 \oplus \cdots \oplus RA_k \oplus \cdots$ is a direct sum in R.

The inverse statement is based on the following useful lemma proved by efforts of D. Farkas and R. Snider [FS77], S. Montgomery [Mo78] and J.-L. Pascaud [Pc81].

LEMMA 4.2. *Let G be an M-group of automorphisms of a semiprime ring R. If the fixed ring R^G is a left Goldie ring then the left R^G-module R is isomorphic to a submodule of a free module of a finite rank.*

We will prove here this fact for the case when R has no $|G|$-torsion. By Theorem 3.5 the set Sh of all elements a for which formula (1) holds is a two-sided ideal with zero annihilator. The intersection $S = \bigcap_{g\in G} Sh^g$ contains the product $\prod_{g\in G} Sh^g$, therefore S is an invariant ideal with zero annihilators. By the Bergman–Isaacs theorem the intersection $S^G = S \cap R^G$ is a non-zero two-sided ideal of the fixed ring. Moreover S^G has zero annihilator in R^G. Indeed, if $S^G b = 0$ then $t(Sb) \subseteq t(S)b = 0$, contrary to the Bergman–Isaacs theorem. Thus S^G is an essential left ideal of R^G and we can find a regular element s in it. This element will be regular in the ring R also: if $bs = 0$, $a \neq 0$ then $b^g s = 0$ for all $g \in G$ and we have $\sum_{g\in G} Rb^g s = 0$, which implies $t(\sum_{g\in G} Rb^g)s = 0$ and $t(\sum_{g\in G} Rb^g) = 0$ again a contradiction with the Bergman–Isaacs theorem.

Thus, we have found a regular element $a = s$ in formula (1). The regularity means at least that Ra is isomorphic to R as a left R^G-module, and formula (1) shows that the mapping

$$x \to \sum \bigoplus \tau_k(xb_k)$$

is the sought for embedding. □

The theorem follows easily from the just proved lemma. If R^G is semiprime Goldie ring then it has finite Goldie dimension. Finiteness of Goldie dimension is preserved under direct sums. Thus every free left R^G-module of finite rank has finite left Goldie dimension as well as any of its submodules, among which one can find R, has.

Finally, by the Johnson theorem it is enough to show that each essential left ideal L of R has zero right annihilator. The intersection $A = \bigcap L^g$ is again essential and, additionally, an invariant ideal. Thus A^G is a nonzero left ideal of the fixed ring. Moreover it is an essential ideal: if $A^G \cap \rho = 0$ than $t(A \cap R\rho) \subseteq A^G \cap t(R)\rho = 0$ and by the Bergman–Isaacs theorem $A \cap R\rho = 0$ – a contradiction. Therefore A^G contains a regular in R^G element s. As we have seen above it will be regular in R as well. Since $s \in L$, we get that the right annihilator of L is zero. The theorem is proved.

5. Noetherian rings

Recall that a ring is called *left Noetherian* if it satisfies the ascending chain condition for left ideals. One can see that this condition is equivalent to the fact that each left ideal is generated by a finite number of elements. Well known is the fact that a finitely-generated module over a Noetherian ring is also Noetherian, i.e. it satisfies the ascending chain condition for submodules.

The following result was first obtained by D. Farkas and R. Snider [FS77] for the case of finite G with no $|G|$-torsion, and by J.-L. Pascaud [Pc81] in this general form. Both proofs were based on Lemma 4.2 in a way as set forth below.

THEOREM 5.1. *Let G be an M-group of automorphisms of a semiprime ring. If the fixed ring is left Noetherian, then R is left Noetherian.*

PROOF. Evidently every left Noetherian ring is left Goldie. By Lemma 4.2 the ring R as a left R^G-module is a submodule of a free module of finite rank. Therefore it is a Noetherian module over R^G and the more so over R. □

It was quite unexpected that the inverse statement is not valid even if the group is finite and the ring has no additive $|G|$-torsion. It needs a more strict restriction.

THEOREM 5.2. *Let G be a finite group of automorphisms of a left Noetherian ring R. If the order of G is invertible in R then the fixed ring R^G is left Noetherian.*

PROOF. Let $I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n \subseteq \cdots$ be an ascending chain of left ideals in the fixed ring. Since R is Noetherian, there is a number k, such that $RI_k = RI_{k+1} = \cdots$. Let us apply the operator $\mathrm{Tr} = |G|^{-1}t$ to both parts of this equality. Since $\mathrm{Tr}(s) = s$ for each fixed element, we get $I_k = I_{k+i} = \cdots$. □

Let us describe here an example by Chuang and Lee [CL77] of a Noetherian ring without additive torsion which has a non Noetherian fixed ring.

EXAMPLE 5.3. Let $A = Z[a_1, b_1, a_2, b_2, \ldots]$ be a ring of polynomials with integer coefficients. Let us denote by K the localization of A with respect to $2A$, i.e.,

$$K = \left\{ \frac{g}{f} \mid g \in A,\ f \in 2A \right\},$$

and let us consider the ring $R = K[[x, y]]$ of power series in the variables x, y. Since K is a principal ideal ring, R is Noetherian. Let the automorphism g on R be given by the formulae:

$$x^g = -x,\ y^g = y,\ a_i^g = -a_i + (a_{i+1}x + b_{i+1}y)y,$$
$$b_i^g = b_i + (a_{i+1}x + b_{i+1}y)x.$$

It is evident that $g^2 = 1$ and R has no additive torsion. One can prove that R^G is not a Noetherian ring. The proof is based on the result by Nagarajan [Na68], who shows that in the ring $(R/2R)^G$ there is a strictly ascending chain of ideals

$$(P_1) \subset (P_1, P_2) \subset \cdots \subset (P_1, \dots, P_n) \subset \cdots,$$

where $P_i = a_{i+1}x + b_{i+1}y + 2R$.

6. Polynomial identities

The relation of fixed ring theory with PI-rings started with an oral question posed independently by V.N. Latishev and J.-E. Björk: if the fixed ring satisfies a polynomial identity (PI) must R also satisfy an identity? This question naturally arose after the S. Amitsur paper [Am68] where there was obtained an affirmative answer for the same question for rings with involution. From Example 1.4, we see that some hypothesis on R or G is necessary. The original Latishev–Bjork question was concerned the case of (not necessary semiprime) rings with no $|G|$-torsion. A positive solution was first found for solvable groups independently by V.E. Barbaumov [Ba73,Ba75] and S. Montgomery [Mo74]. A result for arbitrary finite groups was proved by the author [Kh74].

THEOREM 6.1. *Let G be a finite group of automorphisms of a ring R with no additive $|G|$-torsion. If the fixed ring R^G is PI then so is R.*

This theorem can be proved in three steps. First, if R is a semiprime ring of a finite prime dimension then the fixed ring has finite prime dimension also (see Theorem 1.2 above and the comment to it). The fixed ring is a left Goldie ring, as is any semiprime PI-ring with a finite prime dimension. Thus we can apply Lemma 4.2 and get that the left R^G-module R is a submodule of a free module of a finite rank. Therefore the ring R, considered as a ring of endomorphisms of the left R^G-module R acted on by right multiplications, is a homomorphic image of a subring of the endomorphism ring of a free module. The latter ring is isomorphic to the ring of $n \times n$-matrices over R^G, which satisfies a PI. Thus R also satisfies a polynomial identity. Taking into account the quotient part of Theorem 4.1 one can prove that if the fixed ring satisfies a PI of degree d, then R satisfies a standard identity of degree $|G|d$.

The second step is concerned with an arbitrary semiprime ring R. In this case R is a subdirect product of a family of prime rings. Let $\{P_\alpha\}$ be the set of all prime ideals such that there is no $|G|$-torsion in R/P_α. It is easy to see that $\bigcap P_\alpha = 0$. Let $Q_\alpha = \bigcap_{g \in G} P_\alpha^g$. Then the Q_α are invariant ideals with R/Q_α semiprime rings of finite prime dimension. An action of the group G is induced on R/Q_α, and $t(R/Q_\alpha)$ is a homomorphic image of $t(R)$. This implies that $t(R/Q_\alpha)$, and therefore $(R/Q_\alpha)^G$, satisfy an identity of degree d. Thus we can apply the first step to these quotient rings and find a standard identity of degree $|G|d$ on R.

Finally, if R is not semiprime, one can use a method of Amitsur [Am68]. Let $P(R)$ be the Baer radical of the ring R. This is the smallest ideal such that $R/P(R)$ is a semiprime ring. This implies that there is an induced G-action on the quotient ring, and the quotient ring has no additive $|G|$-torsion. By the above, we have $S_{|G|d}(r_1, \dots, r_{|G|d}) \in P(R)$, for all $r_i \in R$.

Let $A = \{(r_1, \dots, r_{|G|d}) \mid r_i \in R\}$ be the set of all $|G|d$-tuples in R, and let $R_1 = \prod_{\alpha \in A} R_\alpha$, be the complete direct product of copies of R indexed by A. We may extend G to R_1 by letting it act componentwise. For elements $f_1, \dots, f_{|G|d} \in R_1$, we have $u = S_{|G|d}(f_1, \dots, f_{|G|d}) \in P(R_1)$. In particular u it is nil. We choose some particular f_i: let $f_1 = \prod_\alpha (r_1)_\alpha$, where $\alpha = (r_1, \dots, r_{|G|d})$; and $f_2 = \prod_\alpha (r_2)_\alpha; \dots; f_{|G|d} = \prod_\alpha (r_{|G|d})_\alpha$. That is, f_j has the jth component of $\alpha \in A$ as its αth entry.

Now there exists $k > 0$ so that $S_{|G|d}(f_1, \dots, f_{|G|d})^k = 0$. But this says precisely that $S_{|G|d}(r_1, \dots, r_{|G|d})^k = 0$, all $r_i \in R$, which proves the theorem. □

This result is still valid for rings with no non-zero nilpotent elements and arbitrary finite G [Kh75]. For semiprime rings and X-outer groups (see Definition 6.4 set forth below) S. Montgomery has proved that the PI degree of R is equal to that of R^G, provided R^G is left Goldie (which is equivalent for R to be of a finite prime dimension). In general for semiprime rings J.M. Goursaud, J. Osterburg, J.L. Pascaud and J. Valette [GOPV81] have proved the following result, which also can be obtained in the way described above or by using the invariant sheaf.

THEOREM 6.2. *Let G be a finite M-group of automorphisms of a semiprime ring R. If the fixed ring R^G satisfies a polynomial identity of degree d, then R satisfies a standard identity of degree $|G|d$.*

The problems considered in this section (as well as the original Bergman–Isaacs paper) are closely connected with the classical idea of algebraic independence of automorphisms. Indeed, if the fixed ring satisfies an identity, say $f(x_1, \dots, x_d) = 0$, then on the ring R we have an identity

$$f\left(\sum x_1^g, \dots, \sum x_d^g\right) = 0, \tag{4}$$

and it becomes very important to investigate identities with automorphisms. In fact the technique of the proof of the Bergman–Isaacs theorem was actually arrived at by trying to apply to the identity $\sum x^g = 0$ the "standard" argument for proving automorphisms linearly independent [Ar53, p. 35].

We shall start with an example which shows that automorphisms are not necessarily linearly independent in the noncommutative situation any more.

EXAMPLE 6.3. Let D be any division ring which does not satisfy a PI and has a primitive fourth root of unity i in its center. Let R be the ring of 2×2-matrices over D. Let g be the automorphism of R given by a conjugation by $A = \mathrm{diag}(i, -i)$. It is easy to verify that R satisfies the identity $x - x^g + x^{g^2} - x^{g^3} = 0$; that is $1 - g + g^2 - g^3 = 0$. However R does not satisfies a PI.

This example can be easily generalized to an arbitrary ring with an invertible algebraic over the center element A. In this case the powers of the inner automorphism defined by the conjugation by A are linearly dependent. The following result in some extent says that this is the only reason for dependence of automorphisms. (Recall that an automorphism is called *inner* if it is a conjugation by some invertible element.)

DEFINITION 6.4. An automorphism g of a semiprime ring R is said to be X-*outer* if there are no elements $a, b \in R$, such that the identity $axb = b^g x^g a$ is valid on R while $aRb \neq 0$. For a prime ring R this definition is equivalent to g being an outer automorphism when extended to the left Martindale ring of quotients R_F, while for the semiprime case it means that no restriction of g to a subring of the form eR_F with central idempotent $e \in R_F$ is inner. Finally two automorphisms g, h are called *mutually outer* if gh^{-1} is an X-outer automorphism.

THEOREM 6.5. *Let R be a semiprime ring, and let $g_1, \dots, g_m$ be mutually outer automorphisms of R. If the following identity holds on R:*

$$\sum_{i=1}^{m} \sum_{j=1}^{n(i)} a_{ij} x^{g_i} b_{ij} = 0$$

then the following m equalities are valid in the tensor product $R_F \otimes R_F$ over the center of R_F:

$$\sum_{j=1}^{n(i)} a_{ij} \otimes b_{ij} = 0$$

in particular for each i, $1 \leqslant i \leqslant m$, we have an identity (on R_F):

$$\sum_{j=1}^{n(i)} a_{ij} x b_{ij} = 0.$$

This theorem allows one to apply methods of the theory of *generalized polynomial identities* (GPI, i.e. identities with coefficients in the given ring not commuting with the variables) to the investigation of fixed rings. We are going to consider here the Rowen approach [Ro75] connected with the notion of an essential generalized identity. Roughly speaking this is a generalized identity which is not trivial on all the homomorphic images of the given ring R. More precisely a multilinear GPI of the form:

$$\sum_{\pi \in S_d} \sum_{i=(i_1,\dots,i_d)} a_{i,\pi}^{(0)} x_{\pi(1)}^{g_{i_1}} a_{i,\pi}^{(1)} \dots a_{j,\pi}^{(n-1)} x_{\pi(n)}^{g_{i_d}} a_{i,\pi}^{(d)} = 0$$

is called *essential* if the two sided ideal I_f generated by all values of its generalized monomials f_π:

$$f_\pi = \sum_{i=(i_1,\dots,i_d)} a_{i,\pi}^{(0)} x_{\pi(1)}^{g_{i_1}} a_{i,\pi}^{(1)} \dots a_{j,\pi}^{(n-1)} x_{\pi(n)}^{g_{i_d}} a_{i,\pi}^{(d)}$$

coincides with R.

For instance, the generalized monomials of the multilinear polynomial identity $\sum_{\pi \in S_d} \alpha_\pi x_{\pi(1)} x_{\pi(2)} \ldots x_{\pi(d)} = 0$ are the ordinary monomials $\alpha_\pi x_{\pi(1)} x_{\pi(2)} \ldots x_{\pi(d)}$. Therefore the polynomial identity will be essential iff $\sum \alpha_\pi R^d = R$. The generalized monomials of the multilinear polynomial identity with automorphisms (4) are $\alpha_\pi t(x_{\pi(1)}) \ldots t(x_{\pi(d)})$. Thus the identity is essential iff $Rt(R)^d R = R$.

THEOREM 6.6. *Let R be a ring with* 1. *If it satisfies a multilinear essential GPI with automorphisms than it is a PI-ring.*

This theorem was proved by L. Rowen [Ro75] for generalized identities without automorphisms and by the author for GPI with automorphisms [Kh75b]. For semiprime rings the general form can be easily obtained from the Rowen version by using Theorem 6.5, that allows one to get an ordinary generalized identities from ones with automorphisms. Then it is possible to apply the Amitsur method.

This theorem implies a generalization of Theorem 6.1.

THEOREM 6.7. *Let R be a ring with* 1 *and G be a finite group of automorphisms. If the fixed ring has a polynomial identity of degree d and $Rt(R)^d R = R$ then R is a PI-ring.*

Theorem 6.1 can be obtained from this one even if R has no unit element. In this case it is enough to adjoin an element $1/n$ to R and note that the fixed ring will be still a PI-ring. In particular case when the ring R has an element γ with $t(\gamma) = 1$ (in this case evidently $Rt(R)^d R = R$) Theorem 6.7 is also formulated by K.I. Beidar in a short publication [Bd77].

7. Radicals

Let us consider several well-known examples of radicals and radical properties (for a definition of radicals see [Kh95, 1.6, p. 769]).

(a) The Baer radical.

(b) The Levitzki locally-nilpotent radical, that is determined by the property of local nilpotency.

(c) The locally-finite radical, that is determined by the property of local finiteness of an algebra.

(d) The Koethe or upper nil-radical, that is determined by the property of nilpotency of all elements of an algebra.

(e) The algebraic kernel is determined by the property of algebraicity of all elements of an algebra.

(f) The Jacobson radical, the most well-known, can be determined by the property of quasi-regularity of elements: $\forall x \exists y \; xy + x + y = 0$.

All the radicals listed above (the Koethe radical and the algebraic kernel under the condition that **k** is an uncountable field) obey additional conditions – they are overnilpotent and strict. Let us recall the definitions.

DEFINITION 7.1. A radical is called *overnilpotent* if every algebra with the zero multiplication is radical. A radical is called *strict* if every one-sided radical ideal of an algebra is contained in the radical of this algebra.

All the radicals listed, with the exception of the Jacobson radical, are *absolutely hereditary*, that is each subalgebra of a radical algebra is radical.

THEOREM 7.2. *Let ρ be an overnilpotent strict absolutely hereditary radical. If an algebra R is acted on by a finite group G whose order is invertible in* $\mathbf{k}$ *then*

$$\rho(R^G) = \rho(R) \cap R^G = \rho(R), \tag{5}$$

and, besides, if R^G is ρ-radical, then R is also ρ-radical.

The idea of the proof is based on the Shirshov finiteness theorem. First, if R^G is ρ-radical, but $\rho(R) \neq R$, then there is an action of G induced on the quotient algebra $R/\rho(R)$. This algebra is semiprime and $(R/\rho(R))^G = R^G/\rho(R) \cap R^G$. Thus one can suppose $\rho(R) = 0$. If the element a in R is left finite over R^G then it can be proved that there exists an epimorphism of algebras $\pi : S \twoheadrightarrow aR$, where S is a subalgebra of a matrix algebra M_n over R^G. Since ρ is an overnilpotent and absolutely hereditary radical both M_n and S are ρ-radical. Thus $0 \neq aR \subseteq \rho(R)$, since the radical is strict; a contradiction, which proves the second part of the theorem.

It is clear that $\rho(R^G) \supseteq \rho(R) \cap R^G = \rho(R)^G$. If we consider the right invariant ideal $L = \rho(R)R + \rho(R)$ as an algebra acted on by G, then $L^G = t(\rho(R)R + \rho(R)) = \rho(R)$. Therefore L is ρ-radical. This implies $L \subseteq \rho(R)$, hence $\rho(R^G) \subseteq \rho(R)$, which is required proof.

The same statement is valid for the Jacobson radical as well.

THEOREM 7.3. *If a finite group whose order is invertible in* $\mathbf{k}$ *acts on the algebra R, then*

$$J(R^G) = J(R) \cap R^G = J(R)^G,$$

where J is the Jacobson radical. If R^G is quasi-regular, then the algebra R is also quasi-regular.

This theorem was first proved by S. Montgomery [Mo76]. There is also a proof based on Shirshov finiteness and Pascaud's ideas [Pc81], see [Kh91, p. 297]. In the short Beidar publication [Bd77] it is claimed that if G is an arbitrary finite group (possibly, $|G|\mathbf{k} = 0$) and R has a central element γ with $t(\gamma) = 1$ then the same statement is valid for the radicals (a), (b), (d), (f) listed above. In the original paper [Kh76] Theorem 7.2 is proved for an arbitrary finite group with $Rt(R) = R$.

It is well known that the Koethe radical and the algebraic kernel are strict radicals for algebras over an uncountable field. If the base field is countable than it is open problem whether these radicals are strict. For PI-rings the Koethe radical and the algebraic kernel in relation with fixed ring theory have been considered by S. Montgomery [Mo80] and

E.P. Armendariz [An77], respectively. In the class of PI-algebras the Koethe radical coincides with the Baer radical, while the algebraic kernel coincides with the locally-finite radical. Thus one can apply Theorem 7.2.

Evidently every nil ring of bounded index is PI. A variation of an argument of S.A. Amitsur allows one to prove the following result of M. Lorenz and D.S. Passman [LP80].

THEOREM 7.4. *Let R have no $|G|$-torsion and assume that R^G is nil of bounded index. Then R in nil of bounded index.*

The Jacobson and Baer radicals for (Jordan) fixed rings of Jordan isomorphisms were considered by W.S. Martindale and S. Montgomery [MM77] where equality (3) was proved for the Baer radical. The Baer radical has also been considered by J. Fisher and J. Osterburg in [FO78].

By a careful refinement of the proof of Theorem 7.3, W.S. Martindale [Ma78] has shown that $|G|J(R^G) \subseteq J(R)$ and $J(R)^G \subseteq J(R^G)$. Later E. Puczylowski [Pu84] found a very short proof of this fact. Moreover the Puczylowski method allows him to find short proof of the Bergman–Isaacs theorem (without any concrete bound of nilpotency) with the help of known structure results.

The Martindale result implies that if R has no additive $|G|$-torsion, then R^G is semisimple if and only if this is the case for R. This fact for X-outer groups was obtained by J. Fisher and S. Montgomery [FM78] and by M. Cohen [Co82]. The general result is due to J.M. Goursaud, J.L. Pascaud and J. Valette [GPV83].

THEOREM 7.5. *Let G be an M-group of automorphisms of a semiprime ring R. The fixed ring is Jacobson semisimple if and only if R is Jacobson semisimple.*

However, even if we assume that R has no $|G|$-torsion, it is false that $J(R^G) = J(R) \cap R^G$. The first example of this behavior is due to Martindale; in [Ma78] he constructs a domain R of characteristic 0, with a cyclic group G of automorphisms of order n, such that $J(R^G)$ is not contained in $J(R)$. The following example has been constructed by A. Page [Pa79].

EXAMPLE 7.6. Let $R = A_{2\times 2}$, where $A = \mathbf{Z}_{(2)}$, the integers localized at 2. Since A is a local ring, $J(A) = 2A$, and thus $J(R) = (2A)_{2\times 2}$.

Now let g denote a conjugation by the matrix $e_{12} + e_{21} \in R$. The group G generated by g has order 2, and $R^G = \{a(e_{11} + e_{22}) + b(e_{21} + e_{12}) \mid a, b \in A\}$. It is straightforward to check that $I = \{a(e_{11} + e_{22} + e_{21} + e_{12}) \mid a \in A\}$ is an ideal of R^G which is quasi-regular, and so $I \subseteq J(R^G)$. However, by the above, I is not contained in $J(R)$; thus $J(R)$ does not contain $J(R)^G$.

THEOREM 7.7. *Let G be a finite group of automorphisms of a semiprime ring R with no additive $|G|$-torsion. The fixed ring R^G is semisimple Artinian if and only if the ring R is semisimple Artinian.*

The "if" part of this theorem is an old theorem of Levitzki [Le35]. An elementary proof of the "only if" part was given by M. Cohen and S. Montgomery [CM75]. This theorem is still valid for M-groups [Kh91, p. 298].

The following results concerning primitive rings have been obtained by the efforts of a number of authors.

THEOREM 7.8. *Let G be a finite group of automorphisms of a semiprime ring R with no G-torsion (or, more generally, an M-group). Then the fixed ring R^G is a finite subdirect product of left primitive rings if and only if this is the case for R.*

This theorem was first proved by S. Montgomery [Mo76] for the case $|G|^{-1} \in R$ and prime R by using the skew group ring method (see also the short Beidar publication [Bd77]). Later J.-L. Pascaud [Pc81] found a short proof of the result for prime R and group with G-simple algebra $\mathbf{B}(G)$ based on Shirshov finiteness. Under these conditions the fixed ring is prime and therefore they get a criterium for R^G to be primitive. The proof of the general result presented here based on ideas of Pascaud can be found in [Kh91, p. 300].

For prime rings the "only if" part of this theorem is true in more general form.

THEOREM 7.9. *Let G be a finite group of automorphisms of a prime ring R with no G-torsion (or, more generally, an M-group). If R^G has a non-zero ideal which is primitive as a ring then R is primitive.*

The same results are valid for primitive rings with non-zero one-sided minimal ideal. The following theorem in one direction was proved by M. Hacque [Ha87] and in this general form it can be found in [Kh91, p. 306].

THEOREM 7.10. *A fixed ring of an M-group of automorphisms of a semiprime ring R is a finite subdirect product of primitive rings with one sided minimal ideals if and only if this is the case for R.*

8. Maximal ring of quotients

Let us recall some definitions [La66]. A left ideal D of a ring R is called *dense* if for every $r \in R$ the right annihilator in the ring R of the left ideal $Dr^{-1} \overset{\mathrm{df}}{=} \{x \in R \mid xr \in D\}$ equals zero. The set $\mathbf{S}$ of all dense left ideals forms an *idempotent filter* or, equivalently, defines a *Gabriel topology* (see, for instance, [St71]). That is the following statements are valid for $\mathbf{S}$:

(1) if $D \in \mathbf{S}$ and $r \in R$, then $Dr^{-1} \in \mathbf{S}$;

(2) if L is a left ideal of the ring R and there exists a left ideal $D \in \mathbf{S}$, such that for all $d \in D$ we have $Ld^{-1} \in \mathbf{S}$, then $L \in \mathbf{S}$.

From these statements one can easily deduce that an intersection of a finite set of dense left ideals is a dense left ideal. A two-sided ideal I will be dense iff its right annihilator is zero, since $Ir^{-1} \supseteq I$.

The maximal (left) ring of quotients $Q_{\max}(R)$ is defined as the direct limit:

$$Q_{\max}(R) = \varinjlim_{D \in \mathbf{S}} \operatorname{Hom}(D, R).$$

LEMMA 8.1. *Any automorphism of a ring R has a unique extension to $Q_{\max}(R)$.*

PROOF. If $D \in \mathbf{S}$ and $f \in \operatorname{Hom}(D, R)$, then $D^g \in \mathbf{S}$ for an arbitrary automorphism g of the ring R. The element $f^g \in \operatorname{Hom}(D^g, R)$ is defined by the formula $f^g(x) = (f(x^{g^{-1}}))^g$. It is easy to see that $f \to f^g$ is the sought extension of g. $\square$

Thus we can suppose that the group G acts on the maximal quotient ring $Q_{\max}(R)$.

THEOREM 8.2. *Let G be an M-group of automorphisms of a semiprime ring R. Then $Q_{\max}(R^G) = (Q_{\max}(R))^G$.*

This theorem was obtained by J.M. Goursaud, J.L. Pascaud, J. Valette in [GPV83]. The proof is based on the following two lemmas.

LEMMA 8.3. *Let G be an M-group of automorphisms of a semiprime ring R. If D is a dense left ideal of R, then $D \cap R^G$ is a dense left ideal of R^G.*

LEMMA 8.4. *Let G be an M-group of automorphisms of a semiprime ring R. If A is a dense left ideal of R^G, then RA is a dense left ideal of R.*

We are going to present here a proof of Theorem 8.2 with the help of these lemmas for the case of a finite group G and a ring R with no $|G|$-torsion, which is based on the authors ideas from [Kh74].

PROOF. Let us first show that $Q_{\max}(R^G)$ is naturally embedded into $Q_{\max}(R)$. Let $f \in \operatorname{Hom}(A, R^G)$, where A is a dense left ideal of the fixed ring. Let us define a correspondence $f^h : RA \to R$ by the formula

$$\Big(\sum r_i a_i\Big) f^h = \sum r_i (a_i f).$$

This correspondence is a mapping. Indeed, the set $V = \{\sum r_i (a_i f) \mid \sum r_i a_i = 0\}$ is a left ideal and its trace is zero:

$$t\Big(\sum r_i (a_i f)\Big) = \sum t(r_i)(a_i f) = \Big[\sum t(r_i) a_i\Big] f = t\Big(\sum r_i a_i\Big) f = 0.$$

By the Bergman–Isaacs theorem one has $V = 0$. Therefore f^h is a homomorphism of left R-modules. This homomorphism is defined on a dense left ideal RA. Hence, f^h determines an element in $Q_{\max}(R)$. It is now obvious that the mapping $h : f \to f^h$ is an embedding of $Q_{\max}(R^G)$ into $Q_{\max}(R)$.

Now we are to show that the image of h coincides with $(Q_{\max}(R))^G$. If $f \in Q_{\max}(R^G)$, $x = \sum r_i a_i \in RA$, then

$$x(f^h)g = \left[\left(\left(\sum r_i a_i\right)^{g^{-1}}\right)f\right]^g = \left[\sum r_i^{g^{-1}}(a_i f)\right]^g = \left(\sum r_i a_i\right)f^h = xf^h.$$

Therefore, $RA[f^h - (f^h)^g] = 0$ and, hence, $f^h = (f^h)^g$ since RA is dense.

Conversely, let $\xi \in (Q_{\max}(R))^G$, $\xi : D \to R$. Since $(D \cap R^G)\xi \subseteq Q_{\max}(R)^G \cap R$, the restriction f of ξ on $(D \cap R^G)$ belongs to $Q_{\max}(R^G)$. Furthermore $f^h = \xi$, since $R(D \cap R^G)$ is a dense left ideal, while the difference $f^h - \xi$ maps it to zero. □

The theorem just proved can be applied to other rings of quotients. For example the left Martindale ring of quotients R_F of a semiprime ring R is a subring of the maximal ring of quotients, $Q_{\max}(R)$, in which case $R_F^G = Q_{\max}(R)^G \cap R_F$ and we have as a result:

THEOREM 8.5. *Let R be a semiprime ring, and let G be an M-group of its automorphisms. In this case the following equalities are valid*:

$$(R_F)^G = \left(R^G\right)_F, \qquad Q_s(R)^G = Q_s\left(R^G\right),$$

where $Q_s(R)$ is the symmetric Martindale ring of quotients.

The ring R_F can be characterized as the subring of $Q_{\max}(R)$ consisting of those elements for which the set of left denominators contains a two-sided essential ideal of the ring R, while $Q_s(R)$ is the subring of elements of R_F which have an essential ideal of right denominators. Therefore, for the proof it is enough to show that for each two-sided essential ideal I of R the intersection $I \cap R^G$ is a two-sided essential ideal and vice versa: for each two-sided essential ideal A of R^G the left ideal RA (as well as the right one AR) contain a two-sided essential ideal of the ring R.

Let us show the idea of the proofs of these two facts by consideration of the case of finite G and an R with no $|G|$-torsion. If I is an essential two-sided ideal then $J = \bigcap_{g\in G} I^g$ is an invariant essential two-sided ideal with $t(J) \subseteq I \cap R^G$. Now if $0 \neq a \in \operatorname{ann}_{R^G}(I \cap R^G)$ then $t(aJ) = at(J) = 0$, which is impossible by the Bergman–Isaacs theorem.

Conversely, if A is an essential two-sided ideal of R^G, then AR is an invariant left ideal, therefore it can be considered like a ring acted on by the group G. This is a semiprime ring: if $uARu = 0$, then $uARuA = 0$ and $uA = 0$. But the left annihilator of A in R is a left invariant ideal of R, which has zero intersection with R^G, so by the Bergman–Isaacs theorem $u = 0$.

The same is valid for the right ideal RA. If we apply equality (1) to the ring RA, we will find a non-zero element $a \in RA$, such that $RA \cdot a \subseteq \sum_k \tau_k(RARA)RA$. Since the τ_k are bimodule homomorphisms it follows that $\tau_k(RARA) \subseteq \tau_k(RAR)A \subseteq A$. Now we have $RA \cdot a \cdot R \subseteq ARA \cdot R \subseteq AR$ and AR contains the nonzero two-sided ideal $RA \cdot a \cdot R$ of R. As has been noted, the set of all a's in formulae of type (1) forms an essential two-sided ideal I of R. We have that the essential ideal $RA \cdot I \cdot R$ is contained in AR which is what is required.

Lemma 8.4 is a deep and difficult result. Surely it is very important achievement of [GPV83]. I. Kitamura [Ki76] has proved this lemma for G-Galois extensions. This implies the following result.

THEOREM 8.6. *If R is a G-Galois extension of the fixed ring then $Q_{\max}(R^G) = (Q_{\max}(R))^G$ and moreover $Q_{\max}(R) = R$ iff $Q_{\max}(R^G) = R^G$.*

The original proof of Lemma 8.4 is based on the very important Morita context technique related to an action of (reduced-)finite automorphism group. Another proof using the invariant sheaf construction can be found in [Kh91, 6J].

Lemma 8.4 becomes much simpler if R is nonsingular. Recall that the ring R is called *(left) nonsingular* if it has no non-zero elements with essential left annihilators.

In this case every essential left ideal is dense (and, certainly, vise versa, see [La66, 4.3–4.5]). If A is an essential left ideal of the fixed ring and R is a semiprime ring without additive $|G|$-torsion then by the Bergman–Isaacs theorem RA has non-zero intersection with any invariant ideal: $RA \cap L \supseteq R(A \cap t(L)) \neq 0$. That is RA is an essential $R * G$-submodule in R. One of the variants of the Maschke's theorem (the so-called *essential version*, see [Pm89, Prop. 4.3] or the original papers [LP80a,Pm83]) states that RA is an essential left R-module as well.

For nonsingular rings Theorem 8.2 has been proved by J.M. Goursaud, J. Osterburg, J.L. Pascaud, J. Valette in [GOPV81] and announced by K.I. Beidar in [Bd77] (provided that there exists a central element γ with $t(\gamma) = 1$; this guarantees the validity of the essential version of the Maschke theorem).

The nonsingularity of R is equivalent to $Q_{\max}(R)$ being a regular ring (i.e. $\forall a \exists x\ axa = a$). The nonsingularity of R implies also $Q_{\max}(R)$ is left self-injective. This leads us to another corollary of Theorem 8.2 concerned with regular self-injective rings. This class of rings can be characterized as nonsingular semisimple rings which coincide with their left maximal quotient rings. Thus we have

THEOREM 8.7. *Let R be a regular left self-injective ring. If G is an M-group of automorphisms of it, then R^G is also regular and left self-injective.*

This theorem (under the restriction $|G|^{-1} \in R$) has been proved by A. Page in [Pa79]. If R has no non-zero nilpotent elements or R^G is central it was obtained by R. Diop [Di78]. In this general form the theorem appeared in [GOPV81]. Self-injective fixed rings were considered also by D. Handelman and G. Renault in [HR80]. They proved the following result (note that if a regular unitary ring has no additive n-torsion, then n is invertible).

THEOREM 8.8. *If R is a regular self-injective ring and G is a finite group of invertible order than*

(1) *R is injective as an R^G-module;*
(2) *R is projective if and only if it is finitely generated;*
(3) *R^G is biregular if and only if this is the case for R.*

Finally, note that the equality $Q_{\max}(R^G) = (Q_{\max}(R))^G$ is not always valid even if the ring R is commutative (see C. Faith [Fa76]).

9. Prime spectrum

Let $S \subseteq R$ be a finite integral extension of commutative domains. In algebraic number theory one studies, among other things, the relationship between the prime ideals of these two rings, obtaining in particular the classical properties known as *lying over, going up, going down* and *incomparability*.

In the theory of fixed rings, we are also interested in the relationships between the prime ideals for Galois extensions $R^G \subseteq R$. Here we shall consider the case when G is a finite group such that its order is invertible in R. In this case the fixed elements are determined by the equality $x = T(x)$, where T is the normalized trace function $T = \frac{1}{|G|}\sum_G g$. Therefore, if I is an invariant ideal then we have $(R/I)^G = R^G/I \cap R^G$.

In particular, if P is a prime ideal, then $\bigcap_{g\in G} P^G$ is an invariant ideal and we can go over to considering the quotient ring $R/\bigcap P^g$, which is a subdirect product of a finite number of prime rings, that is it has finite prime dimension. Let us start with well-known properties of such rings and general properties of prime ideals.

LEMMA 9.1. *Each prime ideal P contains a minimal prime ideal $P' \subseteq P$.*

THEOREM 9.2. *Let a ring R have an irreducible presentation as a subdirect product of n prime rings (that is its prime dimension is n). Then the following statements are valid.*

(a) *Every irreducible presentation of R as a subdirect product of prime rings contains exactly n cofactors.*

(b) *A presentation of R as a subdirect product of n prime rings is unique; that is if $R = S_{i=1}^n R_i = S_{i=1}^n R_i'$, then there exists a permutation π, such that $\ker \xi_i' = \ker \xi_{\pi(i)}$, where ξ_i, ξ_i' are the projections onto R_i, R_i', respectively.*

(c) *The ring R has exactly n minimal prime ideals and their intersection equals zero.*

(d) *Every not minimal prime ideal is essential.*

DEFINITION 9.3. The set Spec R of all prime ideals of the ring R considered as a topological space with the following closure operation

$$\overline{\{P_\alpha\}} = \left\{P \in \operatorname{Spec} R \mid P \supseteq \bigcap P_\alpha\right\}$$

is called *spectrum* (or *prime spectrum*) of the ring R.

The lemmas formulated above are quite obvious from the point of view of this topological space. For instance, a decomposition of a semiprime ring into a subdirect product of prime rings means that the set of kernels of the corresponding projections is dense in Spec R. The prime dimension is the least possible number of elements of a dense subset. Theorem 9.2 says if the prime dimension is finite then there exists a smallest finite dense subset in Spec R.

If a group G acts on the ring R, there is an induced action of this group on the space Spec R. In this case there arises the *space of orbits* Spec R/G which is defined as the quotient space of Spec R by the equivalence relation $P \approx Q \leftrightarrow \exists g \in G\colon\ P^g = Q$.

If P is a prime ideal of R and q is a prime of R^G, we say that P *lies over* q or equivalently q *lies under* P if q is a minimal covering prime of $P \cap R^G$, that is $q/P \cap R^G$ is a minimal prime ideal of the quotient ring $R^G/P \cap R^G$.

We can now list the *Krull relations* in the extension $R^G \subseteq R$. It is convenient to do this diagrammatically. For example the middle diagram in (iv) below is read as follow. Suppose $q_1 \supseteq q_2$ are primes of R^G and P_2 is a prime of R lying over q_2. Then there exists a prime P_1 of R such that P_1 lies over q_1 and $P_1 \supseteq P_2$.

THEOREM 9.4. *Let G act on R, and suppose that $|G|^{-1} \in R$. The following basic relations hold between the prime ideals of R and of R^G.*

(i) Cutting down. *If P is a prime ideal of R, then there are $k \leqslant |G|$ primes $q_1, q_2, \ldots, q_k$ of R^G minimal over $P \cap R^G$ and we have $P \cap R^G = q_1 \cap q_2 \cap \cdots \cap q_k$.*

(ii) Lying over. *If q is a prime ideal of R^G, then there exists a prime P of R, unique up to G-conjugation, such that P lies over q. Furthermore, the distinct P^g with $g \in G$ are incomparable.*

(iii) Incomparability. *Given the lying over diagram*

$$\begin{tikzcd}
& P_1 \arrow[dl, dash] \arrow[d, dash] \\
q_1 \arrow[d, dash] & P_2 \arrow[dl, dash] \\
q_2 &
\end{tikzcd}$$

Then $P_1 = P_2$ if and only if $q_1 = q_2$.

(iv) Going up and going down. *We have finally*

$$\begin{tikzcd}
& P_1 \arrow[dl, dash] \arrow[d, dash] \\
q_1 \arrow[d, dash, dotted] & P_2 \arrow[dl, dash, dotted] \\
q_2 &
\end{tikzcd}
\qquad
\begin{tikzcd}
& P_1 \arrow[dl, dash, dotted] \arrow[d, dash, dotted] \\
q_1 \arrow[d, dash] & P_2 \arrow[dl, dash] \\
q_2 &
\end{tikzcd}
\qquad
\begin{tikzcd}
& P_1 \arrow[dl, dash] \arrow[d, dash, dotted] \\
q_1 \arrow[d, dash] & P_2 \arrow[dl, dash, dotted] \\
q_2 &
\end{tikzcd}$$

This theorem is due to S. Montgomery [Mo81] (it was also announced by K.I. Beidar [Bd77] in a different form). The original proof [Mo81] is a most successful application of the third main method in the fixed ring theory, the skew group rings method (see the key papers by M. Lorenz and D. Passman [LP80], D. Passman [Pm83], and the book [Pm89]). Another proof based on Shirshov finiteness can be found in [Kh91, p. 321].

The following example by S. Montgomery and L. Small [MS84] shows that the missing going up result in Theorem 9.4 does indeed fail.

EXAMPLE 9.5. Let A be a simple domain over a field $\mathbf{k}$ with char $\mathbf{k} \neq 2$ and assume that A is not a division ring. Then we can choose $I \neq A$ to be a non-zero left ideal. For example we could take A to be the *Weyl algebra* $A_1(\mathbf{k}) = \mathbf{k}\langle x, y \mid xy - yx = 1\rangle$ with char $\mathbf{k} = 0$ and $I = Ax$. Now define

$$R = \begin{pmatrix} \mathbf{k}+I & A \\ I & A \end{pmatrix} \subseteq A_{2\times 2}.$$

It follows easily that R is a prime ring. Thus $P_2 = 0$ is a prime ideal of R. Also observe that $P_1 = \left(\begin{smallmatrix} I & A \\ I & A \end{smallmatrix}\right)$ is a maximal two-sided ideal of R with $R/P_1 = \mathbf{k}$. Let G be the group of automorphisms of R generated by a conjugation by $\mathrm{diag}(1, -1)$. Then $|G| = 2$ and $R^G = \mathrm{diag}(\mathbf{k} + I, A)$. Thus R^G has two minimal primes, one of which is $q_2 = \mathrm{diag}(\mathbf{k} + I, 0)$. We have $R^G/q_2 \cong A$, so q_2 is also maximal. Thus we see that there exists no prime q_1 of R^G which completes the diagram

$$\begin{tikzcd} & P_1 \arrow[dl, dotted, no head] \arrow[d, no head] \\ q_1 \arrow[d, dotted, no head] & P_2 . \arrow[dl, no head] \\ q_2 & \end{tikzcd}$$

Now it is clear that Theorem 9.4 yields a one-to-one correspondence between the G-conjugacy classes of prime ideals of R and certain finite subsets of primes of R^G. To be precise, if q_1 and q_2 are prime ideals of R^G, write $q_1 \overset{M}{\approx} q_2$ if and only if q_1 and q_2 lie under the same prime P of R. In this case the primes q_1, q_2 are said to be *Montgomery equivalent*. This equivalence defines an equivalence relation whose classes are finite of size $\leqslant |G|$ and there is a one-to-one correspondence between these classes and the G-conjugacy classes of primes of R.

THEOREM 9.6. *The described correspondence is a monotonic homeomorphism of ordered topological spaces.*

$$\mathrm{Spec}\, R/G \cong \mathrm{Spec}\, R^G / \overset{M}{\approx}.$$

It is natural to consider which ring theoretic properties are shared by Montgomery equivalent primes. Obviously these include those properties which are inherited by lying over and lying under primes. Some examples due to M. Lorenz and D.S. Passman [LP80] and S. Montgomery [Mo81] are as follows. Recall that the *height* of a prime P of R is the largest n such that $P_0 \subset P_1 \subset \cdots \subset P_n = P$ is a chain of primes in R. If no such maximum exists, then the height is infinite. The *depth* of P can be defined similarly by looking at primes containing P. Equivalently, this is the prime length of R/P.

THEOREM 9.7. *Let G act on R with $|G|^{-1} \in R$ and let P be a prime ideal of R, which lies over q.*

(i) *P is primitive if and only if q is primitive.*

(ii) *P and q have the same height.*

(iii) *R and R^G have the same prime length and the same primitive length.*

EXAMPLE 9.8. Let R be the ring of all linear transformations of a countable-dimension space V over a field $\mathbf{k}$ with $\operatorname{char} \mathbf{k} = 0$. This ring is naturally identified with a ring of infinite finite-row matrices (see [Kh95, p. 763]). Let us consider the conjugation g by the diagonal matrix $\operatorname{diag}(-1, 1, 1, \ldots)$. The fixed ring for $G = \{g, 1\}$ is isomorphic to $\mathbf{k} \oplus R$. Therefore we have two Montgomery equivalent ideals, $q_1 = \mathbf{k} \oplus 0$, $q_2 = 0 \oplus R$. One of them is maximal, the other is not. The same example shows the condition "R/P is a Goldie ring" to be Montgomery unstable, in spite of the fact that it is transferred onto rings of invariants and vice versa.

Montgomery equivalence turns out to be closely related with Morita contexts ([Kh91]).

Let us consider the properties of prime ideals, which are determined by quotient rings. Setting such a property is equivalent to singling out a class R of prime rings. In this case Montgomery equivalence defines a relation between prime rings (which possibly is not an equivalence): $R \overset{M}{\approx} S$ if and only if there exists a ring A and a group G of invertible order, such that $R \cong A^G/q_1$, $S \cong A^G/q_2$ for Montgomery equivalent primes q_1, q_2 of the fixed ring A^G.

It should be recalled that a *Morita context* is a tuple (R, V, W, S), where R, S are rings; V is an (R, S)-bimodule; W is an (S, R)-bimodule, and there are defined multiplications $V \otimes_S W \to R$, $W \otimes_R V \to S$, so that the set of all matrices of the type $\begin{pmatrix} R & V \\ W & S \end{pmatrix}$ forms an associative ring relative the matrix operations of multiplication and addition. Two unitary rings R and S are called *Morita equivalent* if there is a Morita context (R, V, W, S), such that $VW = R$, $WV = S$. This notion has been studied in detail and its essence is that the categories of left modules over R and S are equivalent if and only if R and S are Morita equivalent.

The Morita context (R, V, W, S) is called *prime* if the corresponding ring $M = \begin{pmatrix} R & V \\ W & S \end{pmatrix}$ is prime. If the rings R and S contain $1/2$, then, as in Example 9.8, we can consider the conjugation g by the element $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. For the group $G = \{1, g\}$ we have $M^G = R \oplus S$, i.e. $R \overset{M}{\approx} S$. The following theorem shows that the inverse statement is also, to a certain extent, valid.

Let us consider two conditions on a class of prime rings $\mathbf{R}$.

Inv. If R is a prime ring and $G \subseteq \operatorname{Aut} R$ is a group of automorphisms of invertible order, such that R^G is prime, then $R^G \in \mathbf{R}$ if and only if $R \in \mathbf{R}$.

Mor. If two nonzero prime rings R and S are related by a prime Morita context and $R \in \mathbf{R}$, then $S \in \mathbf{R}$.

THEOREM 9.9. *If a class of prime rings $\mathbf{R}$ satisfies conditions **Inv** and **Mor**, then it is stable under Montgomery equivalence.*

It can be proved that the classes of prime rings that are semisimple for the radicals considered in Section 7 are Montgomery stable. And so are the class of prime subdirectly indecomposible rings, the class of primitive rings and that of primitive rings with one sided minimal ideal. The classes of simple Artinian rings, Goldie rings, Noetherian and simple rings are Montgomery unstable.

Here one essential remark should be made. Montgomery equivalence can be considered in some classes of prime rings not equal to the class of all prime rings. In this case unstable properties can become stable. For instance, the Gel'fand–Kirillov dimension becomes a stable characteristic in the class of PI-rings (S. Montgomery and L. Small [MS86]). The classical Krull dimension is stable in the class of affine PI-rings (J. Alev [Al83]), as well as in the class of Noetherian PI-rings (S. Montgomery and L. Small [MS86]).

In the case of Goldie rings, we are also concerned with the relationship between the Goldie ranks of R/P and of the corresponding factor rings R^G/q_i. The following *additivity principle* applied to fixed rings has been proved by M. Lorenz, S. Montgomery and L.W. Small [LMS82].

THEOREM 9.10. *Let G act on R, a ring with no $|G|$-torsion, and suppose that R is G-prime and Goldie. If P is a minimal prime of R with $H = G_P$, the stabilizer of P in G, and if $q_1, q_2, \dots, q_k$ are minimal primes of R^G, then*

$$\operatorname{rank}(R/P) = \sum_{i=1}^{k} z_i \cdot \operatorname{rank}\bigl(R^G/q_i\bigr)$$

for suitable integers z_i satisfying $1 \leqslant z_i \leqslant |H|$.

10. Simple and subdirectly indecomposable rings

THEOREM 10.1. *Let G be a finite group of automorphisms of a direct sum of n simple rings having no additive $|G|$-torsion. Then the fixed ring is a direct sum of not more than $n|G|$ simple rings.*

This theorem was proved firstly by T. Sundström [Su74] for solvable groups. Then by the author [Kh75] for arbitrary groups with a bound of the number of summands given by the Bergman–Isaacs theorem. The exact number of summands was found by J. Osterburg [Os78,Os79]. This is equal to the number s of G-simple components of the algebra $\mathbf{B}(G)$ of the group G (see the comments to Theorem 1.2 above). A proof of this theorem based on the skew group rings method was found by M. Lorenz and D. Passman [LP80].

Simple rings have an interesting characterization as hearts of subdirectly indecomposable prime rings. Recall that a ring R is subdirectly indecomposable if and only if it has a smallest non-zero ideal. This ideal is called a *heart* of R and it either has zero multiplication or is a simple ring. The latter case is equivalent to the ring R being prime.

THEOREM 10.2. *Let G be a finite group of automorphisms of a semiprime ring R with no additive $|G|$-torsion. The fixed ring is a subdirect product of finite number of prime subdirectly indecomposable rings if and only if this is the case for R.*

This result is also valid for M-groups [Kh91] unlike Theorem 10.1. Let us describe here a counter example by James Osterburg [Os78], which is based on an example by A.I. Zalesskii and O.M. Neroslavskii [ZN77].

EXAMPLE 10.3. Let $\mathbf{k}$ be a field of characteristic 2. Let us consider the algebra $R_1 = \mathbf{k}(y)[x, x^{-1}]$ over a field of rational functions in a variable y. Let g be the automorphism of this algebra which transfers x into xy, and let G be the infinite cyclic group generated by the automorphism g. Let $R_2 = R_1 * G$ be the skew group ring, and let h be the $\mathbf{k}(y)$-automorphism of R_2, defined by $h(x) = x^{-1}$ and $h(g) = g^{-1}$. One can show that R_2 is a simple ring, h is an outer automorphism, and that the ideal $t(R_2) = \{x + x^h \mid x \in R_2\}$ of the fixed ring contains no unit, that is $R_2^{\langle h \rangle}$ is not simple.

11. Modular lattices

Many properties studied in the ring theory can be formulated in terms of the lattice of left (right) or two-sided ideals (for instance, ascending chain condition, descending chain condition, Krull dimension, Goldie dimension, etc.). If a fixed ring R^G satisfies such a property, then this fact gives information only on certain invariant (left or two-sided) ideals and there naturally arises the problem of relations between the lattice of all ideals with that of invariant ideals. Since the lattice of ideals is modular (that is $I + (J \cap K) = (I + J) \cap K$ provided that $I \subseteq K$), we get the problem of studying fixed points of finite groups acting on modular lattices. The following two theorems proved by J. Fisher [Fi79], and P. Grzeszczuk and E. Puczylowski [GP86] give us a lot of information on relations between L and L^G.

Recalled that a lattice is an algebraic system L with two binary commutative associative operations $\wedge$ and $\vee$, satisfying the identities:

P1. $(x \vee y) \wedge x = x$,

P2. $(x \wedge y) \vee x = x$,

P3. $x \wedge x = x, \ x \vee x = x$.

On a lattice one can introduce a relation of a partial order by $a \leqslant b \Leftrightarrow a \wedge b = a$. A mapping of lattices $f : L \to L_1$ is called *strictly monotone* if the inequalities $a < b, \ a \neq b$ imply $f(a) < f(b), \ f(a) \neq f(b)$. The lattice L called *modular* provided that for all $a \leqslant c$ the following equality holds:

M1. $a \vee (b \wedge c) = (a \vee b) \wedge c$.

THEOREM 11.1. *Let G be a finite group of automorphisms of a modular lattice L. Then there exists a strictly monotone mapping $f : L \to L^G \times \cdots \times L^G$ of the lattice L into a direct product of a finite number of copies of the sublattice L^G of fixed elements.*

Since the lattice of all left (two-sided) ideals is modular, we have the result.

COROLLARY 11.2. *Let G be a finite group of automorphisms of a ring R. If R satisfies the* ACC (DCC) *for invariant left (two-sided) ideals, then R satisfies the same condition for all left (two-sided) ideals.*

It can be proved that the set of all semiprime ideals of a ring is a distributive (thus, modular) lattice with respect to the operations $I \wedge J = I \cap J$ and $I \vee J = \mathbf{B}(I + J) =$ the intersection of all prime ideals containing $I + J$. This implies the corollary.

COROLLARY 11.3. *Let G be a finite group of automorphisms of a ring R. If R satisfies the* ACC (DCC) *for invariant semiprime ideals, then R satisfies the same condition for all semiprime ideals.*

COROLLARY 11.4. *Let the order of the group G be invertible in R. If the fixed ring R^G satisfies the* ACC (DCC) *for semiprime ideals, then so does R.*

Indeed, if $I \subseteq J$ are semiprime invariant ideals and $I^G = J^G$, then in the quotient ring $\overline{R} = R/I$ we have $\overline{J}^G = \overline{0}$. By the Bergman–Isaacs theorem $J^m \subseteq I$. Hence, $J = I$. Therefore, $I \to I^G$ is a strictly monotone mapping of the lattice of invariant semiprime ideals of R into the lattice of semiprime ideals of the fixed ring. Finally one has to use Corollary 11.3.

Now we turn to the Grzeszczuk–Puczyłowski theorem.

DEFINITION 11.5. A set of nonzero elements $\{x_1, \ldots, x_n\}$ is called *independent* if $x_i \wedge (x_1 \vee \cdots \vee x_{i-1} \vee x_{i+1} \vee \cdots \vee x_n) = 0$ for all i, $1 \leqslant i \leqslant n$. Here 0 is the smallest element of the lattice: that is $0 \wedge x = 0$, $0 \vee x = x$, all x. The *Goldie dimension* $\mathrm{Gd}(L)$, of a lattice L is the largest number of elements in independent sets. It can be proved that if in a modular lattice with 0 numbers of elements of independent sets are not bounded then there exists an infinite independent set (i.e. a set all of whose finite subsets are independent). In this case the Goldie dimension is considered to be infinite.

THEOREM 11.6. *If G is a finite group of automorphisms of a modular lattice with* 0 *then*

$$\mathrm{Gd}\big(L^G\big) \leqslant \mathrm{Gd}(L) \leqslant |G| \cdot \mathrm{Gd}\big(L^G\big).$$

From this theorem there results a statement that first had a complex proof [Kh74] and then was reproved by C. Lanski [Mo80, Ch. 5].

COROLLARY 11.7. *Let G be a finite group of automorphisms of a ring R. If R has an infinite direct sum of non-zero left ideals than it has an infinite direct sum of non-zero invariant left ideals.*

12. Noncommutative invariant theory

By the term "invariant theory", most mathematicians mean the following set-up. One is given a finite dimensional vector space V over some field $\mathbf{k}$ and a subgroup $G \subseteq \mathrm{GL}(V)$. One analyzes the induced action of G on the polynomial algebra $\mathbf{k}[V]$ and its homomorphic images. "Noncommutative invariant theory" deals with the noncommutative polynomial algebra $\mathbf{k}\langle V\rangle$ that is the *tensor algebra* of the space V or, equivalently, the *free algebra* of rank n equal to the dimension of V. Probably the first paper in noncommutative invariant theory was by Margarete C. Wolf, who considered invariants for the symmetric group acting by permutation of variables [Wo36].

Of course every (finitely-generated) algebra is a homomorphic image of an algebra of the type $\mathbf{k}\langle V\rangle$. Therefore, all fixed ring theory could be logically considered as noncommutative invariant theory. The tradition is, however, that only results on *relatively free* algebras are included in noncommutative invariant theory. This means that only quotients (homomorphic images) of the free algebra by T-ideals are considered.

Recall that an ideal I of $\mathbf{k}\langle V\rangle$ is called T*-ideal* if it is invariant under all endomorphisms of $\mathbf{k}\langle V\rangle$. This is equivalent for $\mathbf{k}\langle V\rangle/I$ to be a free algebra of a certain variety of $\mathbf{k}$-algebras.

Up to now there are three main divisions in noncommutative invariant theory inspired by classical results. That are finite generation, relative freeness and Hilbert series of rings of invariants.

12.1. *Finite generation*

The main results in commutative invariant theory of this division are the Noether theorem [No16] and the Nagata–Hilbert theorem (see, for instance, [DC71]). The first one says that $\mathbf{k}[V]^G$ is finitely generated provided G is finite, while the second one gives sufficient conditions for $\mathbf{k}[V]^G$ to be finitely generated (even if G is not finite). This condition is the reductivity of the group G. Recall that a rational group $G \subseteq \mathrm{GL}(V)$ is said *reductive* if each invariant submodule U of codimension one of every finitely dimensional rational G-module W, such that $w - w^g \in U,\ w \in W,\ g \in G$, has an invariant complement.

The complete solution of the finite generation problem for the free algebra $\mathbf{k}\langle V\rangle$ was given by A.N. Korjukin [Ko84] in terms of a support space.

DEFINITION 12.1.1. The *support space* of a subset $A \subseteq \mathbf{k}\langle V\rangle$ is the smallest subspace $W \subseteq V$ with $A \subseteq \mathbf{k}\langle W\rangle$.

It is obvious that for subspaces U and W the equality $\mathbf{k}\langle U\rangle \cap \mathbf{k}\langle W\rangle = \mathbf{k}\langle U \cap W\rangle$ is valid. Therefore, if $A \subseteq \mathbf{k}\langle U\rangle$ and $A \subseteq \mathbf{k}\langle W\rangle$ then $A \subseteq \mathbf{k}\langle U \cap W\rangle$. This implies the existence of support spaces. Moreover if A is a G-invariant subspace and $A \subseteq \mathbf{k}\langle W\rangle$ then

$$A^g = A \subseteq \mathbf{k}\langle W\rangle \cap \mathbf{k}\langle W^g\rangle$$

and, hence, the support space is invariant.

In particular the support space of $\mathbf{k}\langle V\rangle^G$ for every irreducible group $G \subseteq \mathbf{k}\langle V\rangle$ is either 0 (provided $\mathbf{k}\langle V\rangle^G = 0$) or V. Another example of interest is given by the following lemma.

LEMMA 12.1.2. *If G is an almost special group, that is it has a subgroup H of finite index with* $\det h = 1,\ h \in H$, *then the support space of* $\mathbf{k}\langle V\rangle^G$ *is equal to V.*

PROOF. Let f be the standard polynomial, $f = \sum_{\pi \in S_n} (-1)^{\pi} x_{\pi(1)} \dots x_{\pi(n)}$. One can easily see that $f^g = \det g \cdot f$. Since $\det G$ is a finite group we can find a number m such that $(\det g)^m = 1$. Therefore $f^m \in \mathbf{k}\langle V\rangle^G$. Let W be the support space of f^m. Let us choose a basis $v_1, \ldots, v_s$ and extend it to a basis $v_1, \ldots, v_n$ of V. For the linear transformation $h : v_i \to x_i$ we have

$$f(x_1, \ldots, x_n) = f(v_1, \ldots, v_n)^h = \det h \cdot f(v_1, \ldots, v_n).$$

Hence, $f^m(v_1, \ldots, v_n,) \in \mathbf{k}\langle W\rangle = \mathbf{k}\langle v_1, \ldots, v_s\rangle$ which is impossible for $s < n$. □

THEOREM 12.1.3. *The algebra $\mathbf{k}\langle V\rangle^G$ is finitely generated if and only if G acts on the support space of $\mathbf{k}\langle V\rangle^G$ as a finite cyclic group of scalar transformations.*

As a corollary we immediately obtain that no irreducible group has a *non-trivial* finitely generated algebra of invariants (except in the case $\dim V = 1$). Lemma 12.1.2 implies that no almost special group has a finitely generated algebra of invariants, but the scalar one: $G = \{g, g^2, \ldots, g^m\}$, $x_i^g = wx_i$, $w^m = 1$. In particular for a finite group G the algebra $\mathbf{k}\langle V\rangle^G$ is finitely generated only if G is scalar. This latter fact was first proved by W. Dicks and E. Formanek [DF82] with the help of Hilbert series (see below in 12.3) and independently by the author [Kh84]. Both proofs were based on the Galois correspondence theorem for free algebras (see, Theorem 12.2.1 below) while the original Korjukin proof is purely combinatorial. His proof shows that the main obstacle for finite generation are the actions of symmetric groups on homogenous components:

$$\pi(v_1 \otimes \cdots \otimes v_m) = v_{\pi(1)} \otimes \cdots \otimes v_{\pi(m)}.$$

This action commutes with the action of $\mathrm{GL}(V)$, and hence carries $\mathbf{k}\langle V\rangle^G$ into itself for all $G \subseteq \mathrm{GL}(V)$. Since the construction of the invariant f^π by an invariant f and a permutation π is of no calculation difficulty, it is natural to study $\mathbf{k}\langle V\rangle^G$ as a graded algebra with actions of symmetric groups on the homogenous components. Such an algebra is called an *S-algebra*. Now the noncommutative variant of the Nagata–Hilbert theorem has the form.

THEOREM 12.1.4. *The algebra $\mathbf{k}\langle V\rangle^G$ is finitely generated as an S-algebra provided that the group G is reductive.*

The original proof by Korjukin is based on a combinatorial lemma of Higman [Hi52]. This lemma allows one to prove the ACC for S-ideals of $\mathbf{k}\langle V\rangle$. Another proof based on the commutative invariant theory is found by T. Tambour [Ta90].

Invariants of reductive groups rationally acting on finitely generated left Noetherian PI-algebras have been considered by N. Vonessen [Vo89].

THEOREM 12.1.5. *Suppose that R is finitely generated left Noetherian PI-algebra and G linearly reductive. Then the following properties hold.*

(1) *R^G is a finitely generated left Noetherian algebra.*
(2) (Lying over) *For every prime ideal p of R^G, there is a prime P of R such that p is a minimal prime over $P \cap R^G$.*
(3) (Separation by invariants) *If I_1 and I_2 are G-stable ideals of R such that $R = I_1 + I_2$. Then $R^G = I_1^G + I_2^G$.*

Conversely, suppose that G is a linear algebraic group which is not linearly reductive. Then there is a finitely generated prime Noetherian PI-algebra with a rational action of G such that none of the above properties hold.

Note that the first statement for finite G and R not necessary PI was obtained by S. Montgomery and L. Small [MS81] (for details see the survey [Mo82]).

If $\mathbf{k}\langle V\rangle/T$ is a relatively free algebra and G is a subgroup of $\mathrm{GL}(V)$, then the finite or infinite generation of $(\mathbf{k}\langle V\rangle/T)^G$ is a function of two variables T and G. The following results [Kh84,DD98] determine those T-ideals for which $(\mathbf{k}\langle V\rangle/T)^G$ is finitely generated, respectively for all finite G and for all reductive $G \subseteq \mathrm{GL}(V)$.

THEOREM 12.1.6. *Let T be a T-ideal. Then $(\mathbf{k}\langle V\rangle/T)^G$ is finitely generated for every finite subgroup G of* $\mathrm{GL}(V)$ *if and only if* $\mathbf{k}\langle V\rangle/T$ *satisfies* ACC *on two-sided ideals.*

THEOREM 12.1.7. *Let T be a T-ideal. Then $(\mathbf{K}\langle V\rangle/T)^G$ is finitely generated for every reductive subgroup G of* $\mathrm{GL}(V)$ *if and only if* $\mathbf{K}\langle V\rangle/T$ *is a Lie nilpotent algebra.*

It is quite rare for $\mathbf{k}\langle V\rangle/T$ to satisfy ACC on two-sided ideals. For example if $T = C^3$, where C is the commutator ideal (generated by $xy - yx$) or if T is the ideal $M(2)$ of identities of the algebra of 2×2 matrices over $\mathbf{k}$ then $\mathbf{k}\langle V\rangle/T$ does not satisfy ACC on two-sided ideals, although $\mathbf{k}\langle V\rangle/C^2$ does.

One of the results on $\mathbf{k}\langle V\rangle/M(2)$ is an accidental looking positive result proved by E. Formanek and A.H. Schofield [FS85].

THEOREM 12.1.8. *Let G be a finite subgroup of determinant one matrices in* $\mathrm{GL}(2,k)$. *If* $\mathrm{char}(\mathbf{k})$ *does not divide $|G|$ then $(\mathbf{k}\langle x_1, x_2\rangle/M(2))^G$ is finitely generated.*

The proof is based on very special properties of the ring of two 2×2 generic matrices. A key ingredient of the proof is the fact that in $\mathbf{k}\langle x_1, x_2\rangle$ the commutator $x_1x_2 - x_2x_1$ is an $\mathrm{SL}(2,k)$-invariant.

The next result by J. Fisher and S. Montgomery [FM86] suggests that further examples with $(\mathbf{k}\langle V\rangle/T)^G$ finitely generated will be hard to find.

THEOREM 12.1.9. *Suppose that $G = \langle g\rangle$ is a finite cyclic subgroup of* $\mathrm{GL}(V)$ *and that G does not act by scalar multiplication. If $|G|$ is invertible in $\mathbf{k}$ and $r \geqslant |G| - \sqrt{|G|} + 1$ then $(\mathbf{k}\langle V\rangle/M(r))^G$ is not finitely generated. Moreover, if g has a characteristic root α such that $\alpha^q = 1$, some q with $0 < q < |G|$, then $(\mathbf{k}\langle V\rangle/M(r))^G$ is not finitely generated whenever $r \geqslant 2$.*

12.2. *Relative freeness*

If G is a subgroup of $\mathrm{GL}(V)$, it is natural to ask when $\mathbf{k}[V]^G$ is itself a polynomial ring, or when $(\mathbf{k}\langle V\rangle)^G$ is a free algebra. The first question is answered by the classic Shephard–Todd [ST54] and Chevalley [Ch55] theorems: if G is a finite group then $\mathbf{k}[V]^G$ is a polynomial ring if and only if G is generated by pseudoreflections. Recall that a linear transformation $g \in \mathrm{GL}(V)$ is a *pseudoreflection* if $\operatorname{rank}(g-1)=1$. The second question has an unexpected solution:

THEOREM 12.2.1. *The algebra* $\mathbf{k}\langle V\rangle^G$ *is always free.*

This fact was obtained independently by D.L. Lane [La76] and the author [Kh78]. For finite G there exists a Galois correspondence [Kh78]. The first results in the Galois correspondence direction were obtained by A.T. Kolotov [Ko78], who proved that in the case of two variables the invariants of the symmetric group form a maximal free subalgebra of $\mathbf{k}\langle x, y\rangle$.

THEOREM 12.2.2. *Let G be a finite subgroup of* $\mathrm{GL}(V)$. *The map* $H \to \mathbf{k}\langle V\rangle^H$ *gives a one-to-one correspondence between all subgroups of G and all free subalgebras of* $\mathbf{k}\langle V\rangle$, *that contain* $\mathbf{k}\langle V\rangle^G$.

Let now T be a T-ideal of $\mathbf{k}\langle V\rangle$. One can ask when $(\mathbf{k}\langle V\rangle/T)^G$ is a relatively free algebra. It turns out that there are beautiful answers for $\mathbf{k}[V]$ and $\mathbf{k}\langle V\rangle$, but only negative results for other relatively free algebras. If, for instance, $T=M(r)$, is the T-ideal of identities satisfied by $r\times r$ matrices over infinite $\mathbf{k}$ then it can be shown that the $(\mathbf{k}\langle V\rangle/M(r))^G$ generate the same variety as do the $\mathbf{k}\langle V\rangle/M(r)$ (see [LK82, Corollary 5]). This fact and the fact that the center of $\mathbf{k}\langle V\rangle/M(r)$ is a commutative domain of rank $r^2(\dim V-1)+1$ (see, for instance [Ro83, p. 197]) show that if $(\mathbf{k}\langle V\rangle/M(r))^G$ is relatively free algebra then it is isomorphic to $\mathbf{k}\langle V\rangle/M(r)$. Using this as a starting point, R.M. Guralnick [Gu85] has shown:

THEOREM 12.2.3. *Let G be a nontrivial finite subgroup of* $\mathrm{GL}(V)$, *and let $M(r)$ be the T-ideal of identities satisfied by the $r\times r$ matrices over an infinite field* $\mathbf{k}$, *where $r\geqslant 2$ and* $\dim V\geqslant 2$, *then* $(\mathbf{k}\langle V\rangle/M(r))^G$ *is not a relatively free algebra.*

From this result it is easy to see that $(\mathbf{k}\langle V\rangle/T)^G$ cannot be relatively free except possibly when the radical of T is the commutator ideal.

12.3. *Hilbert series*

Suppose that

$$R=\mathbf{k}\oplus R_1\oplus R_2\oplus\cdots$$

is a graded $\mathbf{k}$-algebra in which each homogenous component is finite-dimensional over $\mathbf{k}$. The *Hilbert (or Poincaré) series* of R is the formal power series in t defined by

$$H(R)=1+\sum_{i\geqslant 1}(\dim R_i)t^i.$$

The coefficient at t^i therefore gives quantitative information about R_i, namely its dimension. If, for example, $R = \mathbf{k}\langle V\rangle$ is the tensor algebra then $\dim R_i = n^i$, where $n = n_0 = \dim V$. Therefore $H(\mathbf{k}\langle V\rangle) = 1 + nt + n^2t^2 + n^3t^3 + \cdots = \frac{1}{1-nt}$.

For $R = \mathbf{k}[V]$ the dimension of R_i is equal to the number of commutative monomials $x_1^{d_i} \dots x_n^{d_n}$ of degree i in a basis of V. Every monomial can be put into correspondence with one power t^i in the development of the product of power series

$$\left(1+t+t^2+\cdots+t^{d_1}+\cdots\right)\left(1+t+t^2+\cdots+t^{d_2}+\cdots\right)\cdots$$
$$\left(1+t+t^2+\cdots+t^{d_n}+\cdots\right).$$

Thus $H(\mathbf{k}[V]) = 1/(1-t)^n$.

Let now $G \subseteq \mathrm{CL}(V)$. By Theorem 12.2.1 $R = \mathbf{k}\langle V\rangle^G$ is a homogenous free subalgebra of $\mathbf{k}\langle V\rangle$. In particular, one can find a homogenous subspace $U \subseteq \mathbf{k}\langle V\rangle^G$, such that $\mathbf{k}\langle V\rangle^G \cong \mathbf{k}\langle U\rangle$. Let us calculate the auxiliary series

$$P(U) = \sum_{n\geqslant 1} \dim(U \cap R_n)t^n.$$

From the obvious decomposition of graded spaces

$$\mathbf{k}\langle U\rangle \cong \mathbf{k} \oplus \left(U \otimes \mathbf{k}\langle U\rangle\right)$$

we get $H(\mathbf{k}(U)) = 1 + P(U)H(\mathbf{k}\langle U\rangle)$. Hence, in the ring of series we have the relation

$$H\left(\mathbf{k}\langle V\rangle^G\right)^{-1} = 1 - P(U).$$

In particular, this yields that $\mathbf{k}\langle V\rangle^G$ is finitely generated if and only if its Hilbert series has the form $1/(1-f(t))$, where f is a polynomial. We know (Theorem 12.1.3), that as a rule, a Hilbert series has no form of this kind. However, calculations of Hilbert series in explicit form give essential positive information on the algebra of invariants (unlike the mere fact of being not finitely generated).

Hilbert series for algebras of commutative invariants were calculated as far back as the 18th century. For instance, there is the well known Molien theorem [Mo79]: if the field $\mathbf{k}$ has zero characteristic and the group G is finite, then the Hilbert series of the algebra $\mathbf{k}[V]^G$ has the form:

$$H(t) = \frac{1}{|G|}\sum_{g\in G}\frac{1}{\det(1-gt)}.$$

An analogous result is also valid for the algebras of non-commutative invariants (see [ADF85]).

THEOREM 12.3.1. *Let* $\mathbf{k}$ *be a field of zero characteristic, and let* G *be a finite subgroup of* $\mathrm{GL}(V)$. *Then we have*

$$H(t) = \frac{1}{|G|} \sum \frac{1}{1 - \operatorname{tr}(g)t},$$

where tr *is the trace of a linear transformation.*

In the cited work by G. Almkvist, W. Dicks and E. Formanek, this formula was transferred to the case of compact groups in $\mathrm{GL}(V)$ for the field of complex numbers. The authors called their result:

NON-COMMUTATIVE MOLIEN–WEYL THEOREM 12.3.2. *Let* G *be a compact subgroup of* $\mathrm{GL}(n, C)$. *In this case*

$$H(\mathbf{k}\langle V\rangle^G) = \int_G \frac{\mathrm{d}\mu}{1 - t\operatorname{tr}(g)},$$

where $\mathrm{d}\mu$ *is the normalized Haar measure on* G *and* $|t| < 1/n$.

In case $R = \mathbf{k}\langle V\rangle/T$ is a relatively free algebra of rank n, each R_i is a $\mathrm{GL}(n, \mathbf{k})$-module and there is the more refined *character series*. For this definition we need the classical Schur theorem on representations of $\mathrm{GL}(n, \mathbf{k})$ (see [Gr80, Ch. 3]). First of all the *Grothendieck ring of* $\mathrm{GL}(n, \mathbf{k})$*-modules* is defined in the usual way, with $[M]$ denoting the element of the Grothendieck ring represented by a module M. Addition and multiplication are given by

$$[M] + [N] = [M \oplus N], \quad [M][N] = [M \otimes N],$$

where the linear group acts diagonally on $M \otimes N$.

THEOREM 12.3.3. *If* $\mathbf{k}$ *is a field of a zero characteristic then there is an isomorphism (called the character map)*

$$\chi : \mathrm{Gr}\big(\mathrm{GL}(n, k)\big) \to \mathbf{Z}[t_1, \ldots, t_n]^{S_n}$$

between the Grothendieck ring of finite-dimensional polynomial $\mathrm{GL}(n, \mathbf{k})$*-modules and symmetric functions in* n *commuting variables* $t_1, \ldots, t_n$. *In this case*

$$\dim M = \chi[M](1, 1, \ldots, 1).$$

Now the *character series* of R can be defined by the formula

$$\chi(R) = 1 + \sum_{i \geqslant 1} \chi[R_i]t^i,$$

where χ is the character map defined by the Schur theorem. Therefore $\chi(R)$ is a formal power series in t with coefficients in $\mathbf{Z}[t_1, \ldots, t_n]^{S_n}$. By Theorem 12.3.3 it is easy to see that the Hilbert series of R is $\chi(R)(1, 1, \ldots 1; t)$.

If g is a linear transformation of V with eigenvalues $\alpha_1, \ldots, \alpha_n$, then one can define a series

$$\chi(R)(g) = \chi(R)(\alpha_1, \ldots, \alpha_n; t).$$

This allows one to formulate beautiful theorem due to E. Formanek [Fo85] called by him the Molien theorem for relatively free algebras.

THEOREM 12.3.4. *Let T be a T-ideal of $\mathbf{k}\langle V\rangle$ and let G be a finite subgroup of* $\mathrm{GL}(V)$. *If* $\mathrm{char}(\mathbf{k}) = 0$ *then the Hilbert series of* $(\mathbf{k}\langle V\rangle)^G$ *is*

$$H\left(\left(\mathbf{k}\langle V\rangle/T\right)^G\right) = \frac{1}{|G|}\sum_{g\in G} \chi\left(\mathbf{k}\langle V\rangle/T\right)(g),$$

where $\chi(\mathbf{k}\langle V\rangle/T)$ is the character series of $\mathbf{k}\langle V\rangle/T$.

The detailed investigation of noncommutative invariants of the groups $\mathrm{SL}(n, \mathbf{k})$ and $\mathrm{GL}(n, \mathbf{k})$ is called *classical noncommutative invariant theory*. It was developed in the paper [ADF85], already quoted above, by G. Almkvist, W. Dicks and E. Formanek and also in [Te88] by Y. Teranishi and in [Ta91] by T. Tambour.

In his paper Y. Teranishi investigated invariants of $H = \mathrm{SL}(V)$ acting on the tensor algebra of the space $S^r(V)$ of commutative polynomials of degree r. He has shown that free generators of the algebra of invariants correspond to certain Young tableaux which he calls *indecomposable*. He proved also that a certain Hilbert series associated with $\mathbf{k}\langle S^r(V)\rangle^H$ is rational and satisfies a functional equation.

In [Ta91] T. Tambour has considered invariants and covariants of $\mathrm{GL}(2, \mathbf{k})$ using the S-algebra notion. He developed a noncommutative variant of the symbolic method, found explicit generators for the S-algebra of invariants for $G = \mathrm{GL}(2, \mathbf{k})$ acting on the tensor algebra of a finite-dimensional simple G-module, and proved that the S-algebra of covariants is generated by a fundamental covariant and transvectants.

References

[Al83] J. Alev, *Sur l'extension $R^G \subseteq R$*, Seminar a'Algebre Dubreil et Malliavin, Lecture Notes in Math., Vol. 1029, Springer, Berlin (1983), 267–281.

[ADF85] G. Almkvist, W. Dicks and E. Formanek, *Hilbert series of fixed free algebras and noncommutative classical invariant theory*, J. Algebra **93** (1) (1985), 189–214.

[Am68] S.A. Amitsur, *Rings with involution*, Israel J. Math. **6** (2) (1968), 99–106.

[An77] E.P. Armendariz, *Generalized regular fixed rings of automorphisms*, Notices Amer. Math. Soc., no. 742-16-5, January 1977.

[Ar53] E. Artin, *Galois Theory*, Notre Dame Mathematics Lectures No. 2, 2nd edn, A.N. Milgram, ed., Notre Dame (1953).

[Az46] G. Azumaya, *New foundations of the theory of simple rings*, Proc. Japan Acad. **22** (23) (1946), 325–332.

[Ba73] V.E. Barbaumov, *Automorphisms, and* PI-*algebras*, Uspekhi Mat. Nauk. **28** (1) (169) (1973), 231–232.

[Ba75] V.E. Barbaumov, *Algebraic automorphisms, and* PI-*algebras*, Mat. Sb. **97** (1) (1975), 59–76.

[Bd77] K.I. Beidar, *The fixed ring under the action of a finite group of automorphisms of a ring*, Uspekhi Mat. Nauk. **32** (1) (193) (1977), 159–160.

[Be71] G.M. Bergman, *Groups acting on hereditary rings*, Proc. London Math. Soc. **3** (2) (1971), 70–82.

[BI73] G.M. Bergman and I. Isaacs, *Rings with fixed-point-free group actions*, Proc. London Math. Soc. **27** (1) (1973), 69–87.

[Ch55] C. Chevalley, *Invariants of finite groups generated by reflections*, Amer. J. Math. **77** (1955), 778–782.

[Co75] M. Cohen, *Semiprime Goldie centralizers*, Israel J. Math. **20** (1) (1975), 37–45; addendum **24** (2) (1976), 89–93.

[Co82] M. Cohen, *A Morita context related to finite automorphism groups of rings*, Pacific J. Math. **98** (1) (1982), 37–54.

[CL77] C.L. Chuang and P.H. Lee, *Noetherian rings with involutions*, Chinese J. Math. **5** (1977), 15–19.

[CM75] M. Cohen and S. Montgomery, *Semisimple Artinian rings of fixed points*, Canadian Math. Bull. **18** (2) (1975), 189–190.

[CM82] M. Cohen and S. Montgomery, *Trace like functions on rings with no nilpotent elements*, Trans. Amer. Math. Soc. **273** (1) (1982), 131–145.

[Di78] R. Diop, *Actions de groupes*, These de 3e cycle, Poitiers (1978).

[DD98] M. Domokos and V. Drensky, *A Hilbert–Nagata theorem in noncommutative invariant theory*, Trans. Amer. Math. Soc. **350** (7) (1998), 2797–2811.

[DC71] J. Dieudonne and J. Carrell, *Invariant Theory, Old and New*, Academic Press, New York (1971).

[DF82] W. Dicks and E. Formanek, *Poincare series and a problem of S. Montgomery*, Linear and Multilinear Algebra **12** (1) (1982), 21–30.

[Fa72] C. Faith, *Galois subrings of Ore domains are Ore domains*, Bull. Amer. Math. Soc. **78** (6) (1972), 1077–1080.

[Fa76] C. Faith, *Galois extensions of commutative rings*, Math. J. Okayama Univ. **18** (2) (1976), 113–116.

[Fi79] J. Fisher, *Chain condition for modular lattices with finite group actions*, Canad. J. Math. **31** (3) (1979), 558–564.

[Fo85] E. Formanek, *Noncommutative invariant theory*, Contemp. Math. **43** (1985), 87–121.

[FM78] J. Fisher and S. Montgomery, *Semiprime ideals in rings with finite group action*, J. Algebra **50** (2) (1978), 488–502.

[FM86] J. Fisher and S. Montgomery, *Invariants of finite cyclic groups acting on generic matrices*, J. Algebra **99** (2) (1986), 430–437.

[FO78] J. Fisher and J. Osterburg, *Semiprime ideals in rings with finite group action*, J. Algebra **52** (1) (1978), 241–247.

[FS77] D. Farcas and R. Snider, *Noetherian fixed rings*, Pacific J. Math. **69** (2) (1977), 347–353.

[FS85] E. Formanek and A.H. Schofield, *Groups acting on the ring of two* 2×2 *generic matrices and a coproduct decomposition of its trace ring*, Proc. Amer. Math. Soc. **95** (2) (1985), 179–183.

[Gr80] J.A. Green, *Polynomial Representations of* GL_n, Lecture Notes in Math., Vol. 830, Springer, New York (1980).

[Gu85] R.M. Guralnick, *Invariants of finite linear groups on relatively free algebras*, Linear Algebra Appl. **72** (1985), 85–92.

[GPV82] J.M. Goursaud, J.L. Pascaud and J. Valette, *Sur la travaus de V.K. Kharchenko*, Lecture Notes in Math., Vol. 924, Springer, Berlin (1982), 322–355.

[GOPV81] J.M. Goursaud, J. Osterburg, J.L. Pascaud and J. Valette, *Points fixes des anneaux reguliers autoinjectifs a gauche*, Comm. Algebra **9** (13) (1981), 1343–1396.

[GPV83] J.M. Goursaud, J.L. Pascaud and J. Valette, *Actions de groups et contextes de Morita*, Comm. Algebra **11** (18) (1983), 2069–2105.

[GP86] P. Grzeszczuk and E. Puczyłowski, *Goldie dimension and chain conditions for modular lattices with finite group actions*, Canad. Math. Bull. **29** (3) (1986), 274–280.

[Ha87] M. Hacque, *Thorie de Galois des anneaux presque-simples*, J. Algebra **108** (2) (1987), 534–577.

[HR80] D. Handelman and G. Renault, *Actions of finite groups in self-injective rings*, Pacific J. Math. **89** (1) (1980), 69–80.

[Hi52] G. Higman, *Ordering by divisibility in abstract algebras*, Proc. London Math. Soc. **2** (7) (1952), 326–336.

[Kh74] V.K. Kharchenko, *Galois extensions and quotient rings*, Algebra i Logika **13** (4) (1974), 460–484.

[Kh75] V.K. Kharchenko, *Generalized identities with automorphisms*, Algebra i Logika **14** (2) (1975), 215–237.

[Kh75a] V.K. Kharchenko, *Fixed elements under a finite group acting on a semiprime ring*, Algebra i Logika **14** (3) (1975), 328–344.

[Kh75b] V.K. Kharchenko, *Generalized identities with automorphisms of associative rings with unit*, Algebra i Logika **14** (6) (1975), 681–696.

[Kh76] V.K. Kharchenko, *Galois extensions of radical algebras*, Mat. Sb. **101** (143) **4** (12) (1976), 500–507.

[Kh77] V.K. Kharchenko, *Galois theory of semiprime rings*, Algebra i Logika **16** (3) (1977), 313–363.

[Kh78] V.K. Kharchenko, *Algebras of invariants of free algebras*, Algebra i Logika **17** (4) (1978), 478–487.

[Kh81] V.K. Kharchenko, *Centralizers of finite-dimensional algebras*, Algebra i Logika **20** (2) (1981), 231–247.

[Kh84] V.K. Kharchenko, *Noncommutative invariants of finite groups and Noetherian varieties*, J. Pure Appl. Algebra **31** (1–3) (1984), 83–90.

[Kh91] V.K. Kharchenko, *Automorphisms and Derivations of Associative Rings*, Kluwer Academic Publishers, Dordrecht (1991).

[Kh95] V.K. Kharchenko, *Simple, prime and semiprime rings*, Handbook of Algebra, Vol. 1, Elsevier, Amsterdam (1995).

[Ki76] Y. Kitamura, *Note on the maximal quotient ring of a Galois subring*, Math. J. Okayama Univ. **19** (1) (1976), 55–60.

[Ko78] A.T. Kolotov, *On free subalgebras of free associative algebras*, Sibirsk. Math. Zh. **19** (2) (1978), 328–335.

[Ko84] A.N. Korjukin, *On noncommutative invariants of reductive groups*, Algebra i Logika **23** (4) (1984), 419–429.

[La66] J. Lambek, *Lectures on Rings and Modules*, Waltham, Massachusetts (1966).

[La76] D.R. Lane, *Free algebras of rank two and their automorphisms*, PhD thesis, Bedford College, London (1976).

[Le35] J. Levitzki, *On automorphisms of certain rings*, Ann. of Math. **36** (4) (1935), 984–992.

[LK82] I.V. L'vov and V.K. Kharchenko, *Normal elements of the algebra of generic matrices are central*, Sibirsk. Math. Zh. **23** (1) (1982), 193–195.

[LP80] M. Lorenz and D.S. Passman, *Observations on crossed products and fixed rings*, Comm. Algebra **8** (8) (1980), 743–779.

[LP80a] M. Lorenz and D.S. Passman, *Two applications of Maschke's theorem*, Comm. Algebra **8** (19) (1980) 1853–1866.

[LMS82] M. Lorenz, S. Montgomery and L. Small, *Prime ideals in fixed rings. II*, Comm. Algebra **10** (5) (1982), 449–455.

[Ma78] W.S. Martindale and S. Montgomery, *Fixed rings of automorphisms and the Jacobson radical*, J. London Math. Soc. **17** (1) (1978), 42–46.

[Mi67] Y. Miyashita, *Local finite outer Galois theory*, J. Fac. Sci. Hokkaido Univ. **10** (1) (1967), 1–26.

[MM77] W.S. Martindale and S. Montgomery, *Fixed elements of Jordan automorphisms*, Pacific J. Math. **72** (1) (1977), 181–196.

[Mo74] S. Montgomery, *Centralizers satisfying a polynomial identity*, Israel J. Math. **18** (2) (1978), 207–209.

[Mo78] S. Montgomery, *Outer automorphisms of semiprime rings*, J. London Math. Soc. **18** (2) (1978), 209–221.

[Mo76] S. Montgomery, *The Jacobson radial and fixed rings of automorphisms*, Comm. Algebra **4** (5) (1976), 459–466.

[Mo79] S. Montgomery, *Automorphism groups of rings with no nilpotent elements*, J. Algebra **60** (1) (1979), 238–248.

[Mo80] S. Montgomery, *Fixed Rings of Finite Automorphism Groups of Associative Rings*, Lecture Notes in Math., Vol. 818, Springer, Berlin (1980).

[Mo81] S. Montgomery, *Prime ideals in fixed rings*, Comm. Algebra **9** (4) (1981), 423–449.

[Mo82] S. Montgomery, *Trace functions and affine fixed rings in Noncommutative rings*, Seminaire d'Algebre Paul Dubreil et Marie-Paule Malliavin, Lecture Notes in Math., Vol. 924, Springer, Berlin (1982), 356–374.

[MP84] S. Montgomery and D. Passman, *Galois theory of prime rings*, J. Pure Appl. Algebra **31** (1–3) (1984), 139–184.

[MS81] S. Montgomery and L. Small, *Fixed rings of Noetherian rings*, Bull. London Math. Soc. **13** (1) (40) (1981), 33–38.

[MS84] S. Montgomery and L. Small, *Integrality and prime ideals in fixed rings of PI-rings*, J. Pure Appl. Algebra **31** (1–3) (1984), 185–190.

[MS86] S. Montgomery and L. Small, *Some remarks on affine rings*, Proc. Amer. Math. Soc. **98** (4) (1986), 537–544.

[Na68] K. Nagarajan, *Groups acting on Noetherian rings*, Neuw. Archief. voor Wiskunde. **16** (1) (1968), 25–29.

[No16] E. Noether, *Der Endlichkeitssatz der invarianten endlichen Gruppen*, Math. Ann. **77** (1916), 89–92; reprinted in Collected Papers, Springer, New York (1983), 181–184.

[Os78] J. Osterburg, *Fixed rings of simple rings*, Comm. Algebra **6** (17) (1978), 1741–1750.

[Os79] J. Osterburg, *The influence of the algebra of the group*, Comm. Algebra **7** (13) (1979), 1377–1396.

[Pa79] A. Page, *Actions de groupe*, Seminaire d'Algebre P. Dubreil, Lecture Notes in Math., Vol. 740, Springer, Berlin (1979).

[Pc81] J.-L. Pascaud, *Two results on fixed rings*, Proc. Amer. Math. Soc. **82** (4) (1981), 517–520.

[Pc82] J.-L. Pascaud, *Actions de gropes et traces*, Comm. Algebra **10** (10) (1982), 1101–1117.

[Pm81] D. Passman, *Fixed rings and integrality*, J. Algebra **68** (2) (1981), 510–519.

[Pm83] D. Passman, *It is essentially Maschke's theorem*, Rocky Mountain J. Math. **13** (1) (1983), 37–54.

[Pm89] D. Passman, *Infinite Crossed Products*, Academic Press, Boston (1989).

[PS78] R. Paré and W. Schelter, *Finite extensions are integral*, J. Algebra **53** (2) (1978), 477–479.

[Pu84] E. Puczylowski, *On fixed rings of automorphism*, Proc. Amer. Math. Soc. **90** (4) (1984), 517–518.

[Qu89] D. Quinn, *Integrality over fixed rings*, J. London Math. Soc. **40** (2) (1989), 206–214.

[Re79] G. Renault, *Action de gropes et anneaux regulars injectifs*, Lecture Notes in Math., Vol. 734 (1979), 236–248.

[Ro75] L. Rowen, *Generalized polynomial identities*, J. Algebra **37** (3) (1975), 441–153.

[Ro83] L. Rowen, *Polynomial Identities in Ring Theory*, Academic Press, New York (1983).

[Sh57] A.I. Shirshov, *On rings with identities*, Mat. Sb. **43** (85) (1957), 277–283.

[ST54] G.C. Shephard and J.A. Todd, *Finite unitary reflection groups*, Canad. J. Math. **6** (2) (1954), 274–304.

[St71] B. Stenström, *Rings and Modules of Quotients*, Lecture Notes in Math., Vol. 237, Springer (1971).

[Su74] T. Sundström, *Groups of automorphisms of simple rings*, J. Algebra **29** (3) (1974), 555–566.

[Ta90] T. Tambour, *Connections between commutative and noncommutative rings of invariants*, Preprint of University of Lund and Lund Institute of Technology, 1990:6, 1990:8.

[Ta91] T, Tambour, *Noncommutative classical invariant theory*, Ark. Math. **29** (1) (1991), 127–182.

[Te88] Y. Teranishi, *Noncommutative classical invariant theory*, Nagoya Math. J., **112** (1988), 153–169.

[To73] H. Tominaga, *Note on Galois subrings of prime Goldie rings*, Math J. Okayama Univ. **16** (1) (1973), 115–116.

[Vo89] N. Vonessen, *Action of linearly reductive group on affine PI-algebra*, Mem. Amer. Math. Soc. **81** (414) (1989).

[Wo36] M.C. Wolf, *Symmetric functions of noncommuting elements*, Duke Math. J. **2** (3–4) (1936), 626–637.

[ZN77] A.E. Zalesskii and O.M. Neroslavskii, *There exists a simple Noetherian ring with divisors of zero but without idempotents*, Comm. Algebra **5** (3) (1977), 231–245.

Modules with Distributive Submodule Lattice

Askar A. Tuganbaev

Department of Mechanics and Mathematics, Moscow State University, Moscow, 119899 GSP Russia
E-mail: askar@tuganbaev.mccme.ru

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

Preface

All rings are assumed to be associative and (except for some explicitely indicated cases) to have a nonzero identity element. Expressions such as a "Noetherian ring" mean that the corresponding right and left conditions hold. We denote by Lat(M) the lattice of all submodules of a module M. A module is *distributive* if $F \cap (G + H) = F \cap G + F \cap H$ for all $F, G, H \in \text{Lat}(M)$. A module M is *uniserial* if any two submodules of M are comparable with respect to inclusion. All uniserial modules are distributive. Any quasi-cyclic Abelian p-group is a uniserial module over the ring of integers **Z**. The additive group of rational numbers is a distributive **Z**-module which is not uniserial. All valuation rings in division rings are uniserial.

The class of distributive rings includes all commutative Dedekind rings (for example, rings of integral algebraic numbers or commutative principal ideal rings). In particular, the ring of integers and the polynomial ring $F[x]$ over a field F are distributive rings. In addition, all Abelian regular rings (for example, all quotient-rings of direct products of skew fields and all commutative regular rings) are distributive. All commutative Dedekind rings and all serial right Noetherian rings possess the following property: each finitely generated A-module decomposes into a direct sum of distributive modules.

We point out that [23,13,112,113], and [14] are among the first works concerning distributive modules and rings in the noncommutative case. The systematic study of distributive modules over noncommutative rings was initiated in papers [1,146,32], and [24]. Distributive modules were considered in [16, Ch. 9], [38, §4.1], and [132, §2.2]. In [232], distributive modules are applied to complex analysis. In [211], distributive rings are used for a study of rings of continuous functions in topological spaces. In [229], distributive modules are applied to study of rings with a duality.

In [89,157,164], and [163], distributive rings are used to investigate rings of weak global dimension one and hereditary rings. In [161,163], and [170], some applications of distributive rings and modules to formal power series rings were obtained. Distributive group and semigroup rings were studied in [98,84,37,160,162,163,165,167], and [176]. Distributive quaternion algebras were studied in [177,178], and [179]. In [158,165,133], and [182], modules which are distributive over their endomorphism rings were studied. Topological aspects of properties of distributive modules and rings were considered in [212,210], and [165]. Distributive modules over incidence algebras were studied in [61]. Modules decomposing into a direct sum of distributive modules, and rings which are direct sums of distributive right ideals were studied in [67,71,148,230,95–97,31,19,229,225,193], and [194]. Distributive graded modules were considered in [34]. Rings possessing faithful distributive modules were studied in [13] and [14]. Conditions sufficient for the distributivity of some lattices of linear subspaces were considered in [85] and [91]. Noncommutative rings whose lattices of two-sided ideals are distributive were studied in [23,114,17,102], and [92]. Distributive modules and rings were addressed in surveys [18,106], and [107].

If F is a subset of a module M_A and $G \in \text{Lat}(M)$, then $(F : G)$ denotes the right ideal $\{a \in A \mid Fa \subseteq G\}$ of A. A module is called a *multiplication* module if $M(M : N) = N$ for all $N \in \text{Lat}(M)$). Distributive modules are closely related to multiplication modules. For example, a module M over a commutative ring A is distributive $\Leftrightarrow$ all finitely generated submodules of M are multiplication [5]. Multiplication modules over noncommutative rings were considered in [204,150,205,206,145], and [192]. Multiplication modules over commutative rings were studied in [5–7,10,30,35,48,49,52,55–57,73,82,104,115,116,125, 126,129–131,141], and [144].

In this chapter, as a rule, we consider distributive modules over noncommutative rings. Since the distributivity of a commutative ring A is equivalent to the fact that all localizations of A w.r.t. its maximal ideals are uniserial rings [89] (in particular, all Prüfer domains are distributive), the commutative case deserves a special consideration. Here, we just point out the papers [89,90,140,5,212], and [175]. Moreover, it is worth noting that uniserial and serial modules and rings (e.g., valuation rings), which are addressed in considerable number of papers, are little touched on here. We just mention the papers [41, 121,122,142], and [215]. Also, see [20–22,33,36,42–47,58–60,65,66,72,81,86–88,93,94, 99,100,105,117–120,134–136,140,143,201,216,203,207,209,213,217–224,226–228].

1. Characterizations of distributive modules

We denote by $J(M)$, $\mathrm{End}(M)$, and $\max(M)$ the Jacobson radical, the endomorphism ring, and the set of all maximal submodules of a module M, respectively. For a subset B of a ring A, we denote by $r(B)$ the right annihilator of B in A. A module M is a *Bezout* module if every finitely generated submodule of M is cyclic. A module M is *completely cyclic* if that all submodules of M are cyclic. A module M is *quasi-invariant* (*resp. invariant*) if all its maximal submodules (resp. all its submodules) are fully invariant in M. A ring A is right quasi-invariant (resp. right invariant) if all maximal right ideals (resp. all right ideals) of A are ideals. A module M is *hereditary* (resp. semihereditary, Rickartian) if all submodules (resp. all finitely generated submodules, all cyclic submodules) of M are projective. A module M_A is *countably injective* if for any countably generated right ideal B of A, every homomorphism $B_A \to M$ can be extended to a homomorphism $A_A \to M$.

A ring A is *regular* if for every $a \in A$, there is $b \in A$ with $a = aba$. A *semiregular* ring is any ring A such that $A/J(A)$ is a regular ring and idempotents of $A/J(A)$ can be lifted modulo $J(A)$. A ring A is an *exchange ring* if the following two equivalent conditions hold: (1) for any $a, b \in A$ with $a + b = 1$, there are idempotents $e \in aA$ and $f \in bA$ with $e + f = 1$; (2) for any $a, b \in A$ with $a + b = 1$, there are idempotents $e' \in Aa$ and $f' \in Ab$ with $e' + f' = 1$. All semiregular rings are exchange rings. A ring is *normal* if all its idempotents are central. A ring A is *Abelian regular* A is a regular normal ring. A ring without nonzero nilpotent elements is a *reduced* ring. A ring A is a *pf-ring* if given any $m, n \in A$ such that $mn = 0$, there are $a, b \in A$ such that $a + b = 1$, $ma = 0$, and $bn = 0$.

A ring A with the centre C is *integral* (resp. *algebraic*) over C if for every $a \in A$, there is a polynomial $f(x)$ with coefficients in C whose leading coefficient is invertible (resp. is not a zero-divisor) in A.

Let M be the rectangular $m \times n$-matrix with elements in a ring A, $k = \min(m, n)$. We say that M *admits a diagonal reduction* if there are two invertible matrices U and V such that $UMV = D$, where D is a diagonal $m \times n$-matrix and $Ad_{ii} \cap d_{ii}A \supseteq Ad_{i+1,i+1}$ for $i = 1, \ldots, k - 1$. If every rectangular matrix over A admits a diagonal reduction, A is an *elementary divisor ring*. A ring A is *right Hermitian* if given any $a, b \in A$, the row (a, b) admits a diagonal reduction.

A topological space is a T_0-*space* if for any two its points, at least one of these points has an open neighborhood which does not contain the second point.

1.1 [112,146,32]. Let M be a right module over a ring A. Then M is distributive $\Leftrightarrow$ for any $m, n \in M$, there is $a \in A$ such that $maA + n(1-a)A \subseteq mA \cap nA$ $\Leftrightarrow$ for any $m, n \in M$, there are $a, b, c, d \in A$ such that $1 = a + b$, $ma = nc$, and $nb = md$ $\Leftrightarrow$ $A = (m : nA) + (n : mA)$ for all $m, n \in M$ $\Leftrightarrow$ M has no subfactors $S \oplus T$, where T is a nonzero homomorphic image of S $\Leftrightarrow$ M has no subfactors which are isomorphic to $T \oplus T$, where T is a simple module $\Leftrightarrow$ M has no 2-generated submodules which have quotient modules which are isomorphic to $T \oplus T$, where T is a simple module.

1.2 [149,165]. Let M be a right module over a ring A. Then M is distributive $\Leftrightarrow$ for any $m, n \in M$, there is a right ideal B of A such that $(m+n)A = mB + nB$ $\Leftrightarrow$ for any subquotient $\overline{M}$ of M and for any $\overline{m}, \overline{n} \in \overline{M}$ such that $\overline{m}A \cap \overline{n}A = 0$, there are $a, b \in A$ such that $1 = a + b$ and $\overline{m}a = \overline{n}b = 0$.

1.3 [199]. M is a distributive module $\Leftrightarrow$ the lattice $\mathrm{Lat}(M)$ of all submodules of M is isomorphic to the lattice of all open subsets of a topological T_0-space.

In addition, the lattice of all submodules of a finitely generated distributive module is isomorphic to the lattice of all open subsets of a compact topological T_0-space.

1.4 [158,165,182]. M is a distributive right module over a ring A $\Leftrightarrow$ for any indecomposable quasi-injective module E_A, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is uniserial $\Leftrightarrow$ for any module E_A which is a direct product of indecomposable quasi-injective right A-modules, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is a distributive Bezout module $\Leftrightarrow$ for any injective module E_A, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is distributive $\Leftrightarrow$ for any quasi-injective module E_A, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is distributive $\Leftrightarrow$ for any module E_A which is the injective hull of a direct sum of all representatives of classes of isomorphic simple right A-modules, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is distributive $\Leftrightarrow$ for any direct summand of any direct product E_A of quasi-injective right A-modules, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is distributive $\Leftrightarrow$ for any module E_A which is the injective hull of an arbitrary simple subfactor T of M, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is distributive $\Leftrightarrow$ for any module E_A which is the injective hull of an arbitrary simple subfactor T of M, $\mathrm{Hom}(M, E)$ is a Bezout left module over the ring $\mathrm{End}(E)$ $\Leftrightarrow$ for any module E_A which is a direct product of injective hulls of nonisomorphic simple subfactors of M, $\mathrm{Hom}(M, E)$ is a left Bezout module over the ring $\mathrm{End}(E)$.

1.5 [168]. Let M be a right module over a semiperfect ring A, and let e be a basis idempotent of A.

(1) M_A is distributive $\Leftrightarrow$ Me_{eAe} is a Bezout module.

(2) If A is left perfect, then M_A is distributive $\Leftrightarrow$ Me_{eAe} is an Artinian Bezout module.

(3) If A is right perfect, then M_A is distributive $\Leftrightarrow$ Me_{eAe} is completely cyclic.

(4) If A is perfect, then M_A is distributive $\Leftrightarrow$ Me_{eAe} is a completely cyclic module with a composition series.

1.6 [68]. Let M be a right module over a semiperfect ring A, and let $1 = \sum_{i=1}^{n} e_i$ be a decomposition of the identity element of A into a sum of local orthogonal idempotents. Then M is distributive $\Leftrightarrow$ Me_i is a right uniserial e_iAe_i-module for any e_i.

1.7 [165]. Let M be a right module over a right quasi-invariant ring A. Then M is distributive $\Leftrightarrow$ any 2-generated semisimple quotient module H of an arbitrary 2-generated submodule N of M is cyclic $\Leftrightarrow$ every 2-generated semisimple quotient module H of any 2-generated submodule N of M is a cyclic $A/J(A)$-module.

1.8 [149,150,165,187]. Let M be a right module over a right invariant ring A. Then M is distributive $\Leftrightarrow$ all finitely generated submodules of M are multiplication modules $\Leftrightarrow$ for each finitely generated submodule N of M and for any finitely generated submodule F of N, there is a finitely generated ideal B of A such that $F = NB$ $\Leftrightarrow$ all 2-generated submodules of M are multiplication modules $\Leftrightarrow$ $A = (F : G) + (G : F)$ for any two finitely generated submodules F and G of M $\Leftrightarrow$ $L \cap N = L(L : N)$ for any finitely generated $L \in \mathrm{Lat}(M)$ and for each $N \in \mathrm{Lat}(M)$ $\Leftrightarrow$ $L(L : N) = N(N : L) = L \cap N$ for any finitely generated $L, N \in \mathrm{Lat}(M)$ $\Leftrightarrow$ $(L : (F + G)) = (L : F) + (L : G)$ for any finitely generated $F, G \in \mathrm{Lat}(M)$ and for each $L \in \mathrm{Lat}(M)$ $\Leftrightarrow$ $((F \cap G) : L) = (F : L) + (G : L)$ for any finitely generated $F, G \in \mathrm{Lat}(M)$ and for each $L \in \mathrm{Lat}(M)$ $\Leftrightarrow$ $M = \bigoplus_{i \in I} M_i$, all M_i are distributive, and $A = r(m_i) + r(m_j)$ for all distinct $i, j \in I$ and for all $m_i \in M_i, m_j \in M_j$.

1.9 [187]. If M is a Noetherian right module over a right invariant ring, then M is distributive $\Leftrightarrow$ all submodules of M are multiplication invariant modules.

1.10 [165]. Let M be a right module over an invariant ring A. Then

M is a distributive locally Noetherian module $\Leftrightarrow$.

M is the unique direct sum of uniform distributive modules M_i $(i \in I)$ such that $A = r(m_i) + r(m_j)$ for all distinct $i, j \in I$ and all $m_i \in M_i$, $m_j \in M_j$, and for every $i \in I$, the following condition holds:

either M_i is a uniserial Artinian module, and there is a maximal ideal B of A such that the annihilator of any nonzero element of M_i is a power of B,

or $A/r(M_i)$ is an invariant hereditary Noetherian domain with the division ring of quotients Q, and M is isomorphic to a submodule of Q_A.

1.11 [165,146]. (1) Let all simple subquotients of a module M_A be isomorphic (this is the case if A is a matrix-local ring). Then M is distributive $\Leftrightarrow$ M is uniserial.

(2) Let M be a module over a local ring A. Then M is distributive $\Leftrightarrow$ M is uniserial $\Leftrightarrow$ M is a Bezout module.

1.12 [199]. Let M be a right module over a ring A, and let the quotient ring $A/J(A)$ be a normal exchange ring (this is the case if A is Abelian regular). Then M is distributive $\Leftrightarrow$ M is a Bezout module.

A subset T in a ring A is a *right denominator set* in A if there are a ring A_T and a ring homomorphism $f_T \equiv f : A \to A_T$ such that $f(T) \subseteq U(A_T)$, $A_T = \{f(a)f(t)^{-1} \mid a \in A,\ t \in T\}$, and $\mathrm{Ker}(f) = K(T)$. In this case, A_T is called the *right ring of quotients* for A with respect to T, and f_T is the *natural homomorphism*. The right A_T-module $N \otimes_A A_T$ is the *module of quotients for N with respect to T* and is denoted by N_T. The A-module homomorphism $g : N \to N_T$ such that $g(n) = n \otimes 1$ is denoted by g_T. If $A \setminus T \equiv M$ is a right ideal of A, then we use f_M, A_M, and N_M instead of f_T, A_T, and N_T. A *right localizable* ring is any ring A such that the right ring of quotients A_M exists for each $M \in \max(A_A)$. Every commutative ring A is localizable.

1.13 [170]. Let N be a right module over a right localizable ring A. Then N is a distributive A-module $\Leftrightarrow$ N_M is a distributive A_M-module for all $M \in \max(A_A)$ $\Leftrightarrow$ N_M is a uniserial A_M-module for all $M \in \max(A_A)$.

1.14 [199]. Let A be a ring without 1, $\mathbf{Z}$ be the ring of integers, and let A^1 be the direct product of additive groups of A and $\mathbf{Z}$. Define multiplication of pairs (a_1, z_1), $(a_2, z_2) \in A^1$ by the rule $(a_1, z_1) \cdot (a_2, z_2) = (a_1a_2 + z_2a_1 + z_1a_2, z_1z_2)$. The group A^1 is a ring with the identity element $(0, 1)$.

(1) If M is a right A-module, then M is distributive $\Leftrightarrow$ for any $m_1, m_2 \in M$, there are $z, y_1, y_2 \in \mathbf{Z}$ and $a, b_1, b_2 \in A$ such that $m_1a + zm_1 = m_2b_2 + y_2m_2$ and $m_2 - m_2a - zm_2 = m_1b_1 + y_1m_1$.

(2) A^1 is right distributive $\Leftrightarrow$ A^1 is a distributive right A-module $\Leftrightarrow$ for any $a_1, a_2 \in A$ and for any $x_1, x_2 \in \mathbf{Z}$, there are $z, x_1, x_2, y_1, y_2 \in \mathbf{Z}$ and $a, b_1, b_2 \in A$ such that $a_1a + x_1a + za_1 = a_2b_2 + x_2b_2 + y_2a_2$, $a_2 - a_2a - x_2a - za_2 = a_1b_1 + x_1b_1 + y_1a_1$, $zx_1 = y_2x_2$, and $x_2 - zx_2 = y_1x_1$.

2. Characterizations of distributive rings

2.1 [158]. A is a left distributive ring $\Leftrightarrow$ all direct products of projective right A-modules are endodistributive.

2.2 [158,165,182]. A is a right distributive ring $\Leftrightarrow$ all direct summands of all direct sums or direct products of quasi-injective right A-modules are endodistributive $\Leftrightarrow$ all direct sums and all direct products of indecomposable quasi-injective right A-modules are endo-Bezout modules $\Leftrightarrow$ all indecomposable quasi-injective right A-modules are endouniserial $\Leftrightarrow$ the injective hull of the direct sum of all representatives of the classes of isomorphic simple right A-modules is endodistributive $\Leftrightarrow$ injective hulls of all simple right A-modules are endodistributive $\Leftrightarrow$ injective hulls of all simple right A-modules are endo-Bezout modules $\Leftrightarrow$ the direct sum or the direct product of injective hulls of every set of nonisomorphic simple right A-modules is an endo-Bezout module.

Let A be a ring, $B(A)$ be the set of all central idempotents of A, and let $S(A)$ be a (nonempty) set of all proper ideals of A generated by central idempotents. An ideal $P \in S(A)$ is said to be a *Pierce ideal* of A if P is a maximal (with respect to inclusion) element of the set $S(A)$. If P is a Pierce ideal of A, then the quotient ring A/P is said to be a *Pierce stalk* of A.

2.3 [199]. (1) A is a right distributive ring $\Leftrightarrow$ all Pierce stalks of A are right distributive $\Leftrightarrow$ all indecomposable quotient rings of A are right distributive.

(2) A is a right distributive exchange ring $\Leftrightarrow$ all Pierce stalks of A are right uniserial rings.

2.4 [157,188,189]. A is a right distributive right nonsingular ring $\Leftrightarrow$ A is a reduced right distributive ring $\Leftrightarrow$ A is a semiprime right distributive ring, and for any minimal prime ideal H of A, the ring A/H is right nonsingular $\Leftrightarrow$ for each $M \in \max(A_A)$, the right ring of quotients A_M exists and is a right uniserial domain. If A is a reduced pf-ring, all maximal right ideals of A are completely prime ideals, for any completely prime ideal N, the right ring of quotients A_N exists and is a right uniserial domain, the kernel H of the canonical homomorphism $f : A \to A_N$ is the unique minimal prime ideal H of A contained in N, and H coincides with the set $\{a \in A \mid ta = 0 \text{ for some } t \in A \setminus N\}$.

2.5 [146]. (1) A is a semiperfect right distributive ring $\Leftrightarrow$ A is a finite direct product of right uniserial rings.

(2) A is a right or left perfect right distributive ring $\Leftrightarrow$ A is a finite direct product of right uniserial right Artinian rings.

2.6 [189]. A is a semiprime right distributive right Noetherian left finite-dimensional ring $\Leftrightarrow$ A is a semiprime left distributive left Noetherian right finite-dimensional ring $\Leftrightarrow$ A is a finite direct product of invariant distributive Noetherian domains $\Leftrightarrow$ $A = A_1 \times \cdots \times A_n$, where every A_i is an invariant domain, and each proper quotient ring of A_i is a finite direct product of Artinian uniserial rings $\Leftrightarrow$ A is a finite direct product of invariant hereditary Noetherian domains.

2.7 [154,166,156]. (1) A is a Noetherian right distributive ring $\Leftrightarrow$ A is a finite direct product of Artinian right uniserial rings and invariant hereditary Noetherian domains.

(2) A is a distributive right Noetherian ring $\Leftrightarrow$ A is a distributive left Noetherian ring $\Leftrightarrow$ A is a finite direct product of uniserial Artinian rings and invariant hereditary Noetherian domains.

(3) A is a right distributive left Noetherian ring $\Leftrightarrow$ A is a finite direct product of Artinian right uniserial rings and right distributive left Noetherian domains.

2.8 [157,164,192,199]. Let A be an invariant ring, and let all square-zero elements of A be central. Then A is a distributive semiprime ring $\Leftrightarrow$ A is a distributive pf-ring $\Leftrightarrow$ all submodules of flat (right or left) A-modules are flat $\Leftrightarrow$ all 2-generated ideals of A are flat right A-modules $\Leftrightarrow$ all 2-generated ideals of A are flat left A-modules.

Let n be a positive integer. A module M is *n-integrally closed* if all endomorphisms of all n-generated submodules of M can be extended to endomorphisms of M. A module M is *integrally closed*, if M is n-integrally for all positive integers n.

2.9 [192]. (1) Let A be an invariant ring. Then A is distributive $\Leftrightarrow$ all quotient rings of A are right 2-integrally closed $\Leftrightarrow$ all quotient rings of A are right and left integrally closed.

(2) A is a distributive invariant semiprime ring $\Leftrightarrow$ A is reduced, and all quotient rings of A are right and left integrally closed $\Leftrightarrow$ A is reduced, and all quotient rings of A are right and left 2-integrally closed.

2.10 [164]. Let A be a right hereditary ring. Then A is right distributive $\Leftrightarrow$ A is right invariant.

2.11 [149,185,199]. (1) Let A be a right Bezout ring. Then A is right distributive $\Leftrightarrow$ A is right quasi-invariant.

(2) Let A be a Bezout ring. Then A is right distributive $\Leftrightarrow$ A is left distributive $\Leftrightarrow$ A is right quasi-invariant $\Leftrightarrow$ A is left quasi-invariant.

(3) Let A be a right Bezout ring A which is integral over its centre. Then A is right distributive $\Leftrightarrow$ A is right quasi-invariant $\Leftrightarrow$ A is left quasi-invariant $\Leftrightarrow$ A is right invariant $\Leftrightarrow$ all prime ideals of A are completely prime.

2.12 [199]. Let A be an elementary divisor ring. Then A is right quasi-invariant $\Leftrightarrow$ A is left quasi-invariant $\Leftrightarrow$ A is right distributive $\Leftrightarrow$ A is left distributive $\Leftrightarrow$ A is invariant.

2.13 [189]. Let A be a ring which is algebraic over its centre.

(1) A is a semiprime right distributive ring $\Leftrightarrow$ A is a semiprime left distributive ring $\Leftrightarrow$ A is a distributive reduced ring, for any $x \in A$, there is a positive integer m such that $x^m A = Ax^m$, and for any minimal prime ideal H of A, the ring A/H is a domain algebraic over its centre.

(2) A is a right distributive right Noetherian ring $\Leftrightarrow$ A is a finite direct product of right uniserial right Artinian rings and invariant hereditary Noetherian domains.

(3) A is a right distributive left Noetherian ring $\Leftrightarrow$ A is a finite direct product of Artinian right uniserial rings and invariant hereditary Noetherian domains.

2.14 [189,192]. Let A be a ring which is a finitely generated module over its unitary central subring R.

(1) A is right distributive $\Leftrightarrow$ for all $M \in \max(R)$, the ring of quotients A_M is a right Bezout ring, and $A_M/J(A_M)$ is a finite direct product of division rings.

(2) A is a semiprime right distributive ring $\Leftrightarrow$ A is a semiprime left distributive ring $\Leftrightarrow$ for each $M \in \max(R)$, the ring of quotients A_M is a finite direct product of semihereditary uniform Bezout domains $A_i(M)$, and all factor rings $A_i(M)/J(A_i(M))$ are finite direct products of division rings for all i $\Leftrightarrow$ A is a distributive reduced ring, all submodules of flat A-modules are flat, each prime factor ring of a ring A is a semihereditary order in a division ring, and there is a positive integer m such that $Aa^m = a^m A$ for all $a \in A$.

(3) A is a right Noetherian right or left distributive ring $\Leftrightarrow$ A is a left distributive right or left Noetherian ring $\Leftrightarrow$ A is a finite direct product of uniserial Artinian rings and invariant hereditary Noetherian domains.

(4) Let the prime radical N of A be a finitely generated right or left ideal. Then A is right distributive, and A, A/N are rings with the maximum condition on right annihilators $\Leftrightarrow$ A is left distributive, and A, A/N are rings with the maximum condition on left annihilators $\Leftrightarrow$ A is a finite direct product of distributive uniform domains and uniserial Artinian rings.

Let a and b be two units of a commutative ring A. We denote by $(a, b/A)$ the *generalized quaternion algebra* over A, i.e., the free A-module with the canonical basis $\{1, i, j, k\}$, in which multiplication is A-bilinear, and is defined on the canonical basis such that 1 is the common identity element of the rings A and $(a, b/A)$, and $i^2 = a$, $j^2 = b$, $k^2 = -ab$, $ij = -ji = k$, $ik = -ki = aj$, and $kj = -jk = bi$.

2.15 [178,179]. Let a and b be two units of a commutative ring A, and let $2^{-1} \in A$. Then $(a, b/A)$ is right distributive $\Leftrightarrow$ $(a, b/A)$ is right invariant $\Leftrightarrow$ for any $M \in \max(A)$, $(a_M, b_M/A_M)$ is right uniserial $\Leftrightarrow$ for any $M \in \max(A)$, $(a_M, b_M/A_M)$ is right invariant $\Leftrightarrow$ for any $M \in \max(A)$, $(a_M, b_M/A_M)$ is a uniserial invariant ring $\Leftrightarrow$ for any $M \in \max(A)$, A_M is uniserial, and for any $x, y, z \in A$, the inclusion $x^2 - ay^2 - bz^2 \in M$ is possible only for $x \in M$, $y \in M$, $z \in M$ $\Leftrightarrow$ A is distributive, and for all $M \in \max(A)$ and for any $x, y, z \in A/M$, the equality $x^2 - ay^2 - bz^2 = 0$ is possible only for $x = y = z = 0$.

If φ is an injective endomorphism of a ring A, then we denote by $A_\ell[[x,\varphi]]$ the *left skew (power) series ring* consisting of formal series $\sum_{i=0}^{\infty} a_i x^i$ of the variable x with canonical coefficients $a_i \in A$, where addition is defined naturally and multiplication is defined by the rule $x^i a = \varphi^i(a)x^i$. If φ is an automorphism of a ring A, then the *left skew Laurent series ring* $A_\ell((x,\varphi))$ consists of the series $f \equiv \sum_{i=m}^{\infty} a_i x^i$, where $m = m(f)$ is an integer, and $x^i a = \varphi^i(a)x^i$. Analogously, we define the *right skew (power) series ring* $A_r[[x,\varphi]]$ and the *right skew Laurent series ring* $A_r((x,\varphi))$.

2.16 [157,162]. (1) $A[x,x^{-1}]$ is right distributive $\Leftrightarrow$ $A[x,x^{-1}]$ is a right quasi-invariant right Bezout ring $\Leftrightarrow$ A is a commutative regular ring.

(2) If φ is an injective endomorphism of a ring A, then $A_\ell[x,\varphi]$ is right distributive $\Leftrightarrow$ A is a commutative regular ring, and $\varphi \equiv 1_A$.

2.17 [153,198]. Let φ be an injective endomorphism of a ring A. Then $A_\ell[[x,\varphi]]$ is right distributive $\Leftrightarrow$ $A_\ell[[x,\varphi]]$ is a right Bezout ring, and either A is right quasi-invariant, or right annihilators of all elements in A are ideals $\Leftrightarrow$ $A_\ell[[x,\varphi]]$ is a distributive reduced Bezout ring, and all submodules of flat A-modules are flat $\Leftrightarrow$ A is an Abelian regular countably injective ring, φ is an automorphism, and $\varphi(e) = e$ for any idempotent $e \in A$.

2.18 [157,163,170]. Let φ be an automorphism of a normal ring A such that $\varphi(e) = e$ for any idempotent $e \in A$. Then $A_\ell[[x,\varphi]]$ is right distributive $\Leftrightarrow$ all submodules of flat $A_\ell[[x,\varphi]]$-modules are flat $\Leftrightarrow$ all 2-generated right ideals of $A_\ell[[x,\varphi]]$ are flat $\Leftrightarrow$ all 2-generated left ideals of $A_\ell[[x,\varphi]]$ are flat $\Leftrightarrow$ A is an Abelian regular countably injective ring.

2.19 [202]. Let φ be an automorphism of a ring A. Then $A_\ell((x,\varphi))$ is a semilocal right distributive ring $\Leftrightarrow$ $A_\ell((x,\varphi))$ is a finite direct product of right uniserial rings $\Leftrightarrow$ $A_\ell((x,\varphi))$ is a finite direct product of right Artinian right uniserial rings $\Leftrightarrow$ A is a finite direct product of right Artinian right uniserial rings $A_1,\dots,A_n$, and $\varphi(A_i) = A_i$ for all i.

2.20 [199]. Let A be an exchange ring A. Then A is right distributive $\Leftrightarrow$ A is a normal right Bezout ring $\Leftrightarrow$ for every maximal right or left ideal M of A, there is the two-sided ring of quotients ${}_M A_M$ which is a right uniserial ring, and the canonical ring homomorphism $f: A \to {}_M A_M$ is surjective.

Let $\{G_i\}_{i\in I}$ be a set of nonintersecting monoids enumerated by elements of a semilattice I. A monoid G is a *strong semilattice* of the G_i $(i \in I)$ if $G = \bigcup_{i\in I} G_i$ and for any $i, j \in I$ such that $i \geqslant j$ there is a monoid homomorphism $f_{i,j}: G_i \to G_j$ such that the following conditions are satisfied: (1) $f_{i,i}$ is the identity mapping of G_i; (2) if $i > j > k$, then $f_{j,k} f_{i,j} = f_{i,k}$; (3) for any $a \in G_i$ and $b \in G_j$, their product ab in G coincides with $f_{i,ij}(a) f_{j,ij}(b) \in G_{ij}$.

2.21 [198]. (1) Let A be a ring and let G be a monoid which is a strong semilattice of monoids G_i. Then $A[G]$ is right distributive (resp. right Bezout ring) $\Leftrightarrow$ all rings $A[G_i]$ are right distributive (resp. right Bezout rings).

(2) Let A be a ring and let G be a regular monoid. Then $A[G]$ is right distributive $\Leftrightarrow$ G is a strong semilattice of groups G_i and all group rings $A[G_i]$ are right distributive.

2.22 [198]. Let A be a commutative ring, and let G be a regular monoid. Then $A[G]$ is right distributive $\Leftrightarrow$ G is a strong semilattice of groups G_i, and for any pair (A, G_i) one of the three conditions cited below holds.

$(*)$ A is regular, G_i is an Abelian group, and each finitely generated subgroup H of G_i is a direct product of a cyclic group and a finite group whose order is invertible in A.

$(**)$ A is a distributive algebra over the field of rational numbers Q, G_i is a Hamiltonian group, and for any odd number n which is order of an element of G_i, the ring $A \otimes_Q (-1, -1/Q(\varepsilon_n))$ is right distributive.

$(***)$ A is distributive, G_i is a torsion Abelian group, and if M is any maximal ideal of A such that $\operatorname{char}(A/M) = p > 0$ and the p-primary component $G_i(p)$ of G_i is not equal to 1, then the ring of quotients A_M is a field, and the group $G_i(p)$ is either cyclic or quasi-cyclic.

2.23 [160,167]. Let A be a commutative ring and let G be a cancellative monoid. Then $A[G]$ is right distributive $\Leftrightarrow$ one of the following three conditions holds:

$(*)$ A is a distributive algebra over the field of the rational numbers Q, G is a Hamiltonian group, and for any odd number n that is an order of some element of G, the ring $A \otimes_Q (-1, -1/Q(\varepsilon_n))$ is right distributive.

$(**)$ A is regular, G is a commutative cancellative monoid with the Abelian group of quotients Q, and each finitely generated subgroup H of Q is a direct product of a cyclic group and a finite group whose order is invertible in A.

$(***)$ A is distributive, G is a torsion Abelian group, and if M is any maximal ideal of A such that $\operatorname{char}(A/M) = p > 0$ and the p-primary component G_p of G is not equal to the identity, then the ring of quotients A_M is a field, and the group G_p is either cyclic or quasi-cyclic.

2.24 [167]. Let A be a commutative ring, and let G be a monoid which is a strong semilattice of cancellative monoids G_i $(i \in I)$. Then $A[G]$ is distributive if and only if for any pair (A, G_j) the conditions of 2.23 hold.

2.25 [167]. Let A be a ring, and let G be a cancellative monoid which is not a torsion Abelian group. Then $A[G]$ is right distributive $\Leftrightarrow$

either A is a right distributive algebra over the field of rational numbers Q, G is a Hamiltonian group, and for any odd number n which is an order of an element of G, the ring $A \otimes_Q (-1, -1/Q(\varepsilon_n))$ is right distributive

or A is a commutative regular ring, G is a commutative cancellative monoid with the Abelian group of quotients Q, and each finitely generated subgroup H of Q is a direct product of a cyclic group and a finite group whose order is invertible in A.

2.26 (Open questions).

(1) Is every right distributive ring with right and left Krull dimensions left distributive?

(2) Is every distributive ring localizable?

(3) Are all right ideals of a distributive reduced ring flat?

(4) Is every distributive module over a regular ring a Bezout module?

(5) If A is a right distributive ring which is module-finite over its centre, is A left distributive?

(6) If A is a right distributive ring which is integral over its centre, is A left distributive?

(7) If all right modules over a ring A are endo-distributive, are all left A-modules endo-distributive?

(8) Let all right modules over a ring A be endo-distributive. Is A right or left Noetherian? Is A a finite direct product of uniserial Artinian rings?

(9) When is every finitely generated (resp. finitely presented) right module over a ring A endo-distributive?

(10) When is a skew Laurent series ring right distributive?

References

[1] H. Achkar, *Sur les anneaux arithmétiques à gauche*, C. R. Acad. Sci. Paris **278** (5) (1974), A307–A309.

[2] H. Achkar, *Sur les propriétés des anneaux arithmétiques à gauche*, C. R. Acad. Sci. Paris **284** (17) (1977), A993–A995.

[3] H. Achkar, *Anneaux arithmétiques à gauche*, C. R. Acad. Sci. Paris **286** (20) (1978), A871–A873.

[4] H. Achkar, *Anneaux arithmétiques à gauche: propriétés de transfert*, C. R. Acad. Sci. Paris **288** (11) (1979), A583–A586.

[5] T. Albu and C. Năstăsescu, *Modules arithmétiques*, Acta. Math. Acad. Sci. Hungar. **25** (3–4) (1974) 299–311.

[6] D.D. Anderson, *Multiplication ideals, multiplication rings and the ring $R(X)$*, Canad. J. Math. **28** (4) (1976), 760–768.

[7] D.D. Anderson, *Some remarks on multiplication ideals*, Math. Japan. **25** (4) (1980), 463–469.

[8] D.D. Anderson, D.F. Anderson and R. Markanda, *The rings $R(X)$ and $R\langle X\rangle$*, J. Algebra **95** (1) (1985), 96–115.

[9] A. Barnard, *Distributive extensions of modules*, J. Algebra **70** (2) (1981), 303–315.

[10] A. Barnard, *Multiplication modules*, J. Algebra **71** (1) (1981), 174–178.

[11] E.A. Behrens, *Algebren mit vorgegebenem endlichem, distributivem Idealverband*, Math. Ann. **133** (1957), 79–90.

[12] E.A. Behrens, *Eine Charakteriesierung der T-Moduln mit distributivem Untermoduleverνand bei halbprimärem T*, Arch. Math. **8** (1957), 265–273.

[13] E.A. Behrens, *Distributive darstellbare Ringe I*, Math. Z. **73** (5) (1960), 409–432.

[14] E.A. Behrens, *Distributive darstellbare Ringe II*, Math. Z. **76** (4) (1961), 367–384.

[15] E.A. Behrens, *Die Halbgruppe der Ideale in Algebren mit distributivem Idealverband*, Arch. Math. **13** (1962), 251–266.

[16] E.A. Behrens, *Ring Theory*, Academic Press, New York (1972).

[17] E.A. Behrens, *Noncommutative arithmetic rings*, Rings, Modules and Radicals, Amsterdam, London (1973), 67–71.

[18] K.I. Beidar, V.N. Latyshev, V.T. Markov, A.V. Mikhalev, L.A. Skornyakov and A.A. Tuganbaev, *Associative rings*, Progress of Science and Technology: Series on Algebra, Topology, and Geometry, Vol. 22, Itogi Nauki i Tekhn. VINITI, All-Union Institute for Scientific and Technical Information, Akad. Nauk SSSR, Moscow (1985), 3–116.

[19] T. Belzner, *Towards self-duality of semidistributive Artinian rings*, J. Algebra **135** (1) (1990), 74–95.

[20] K. Bessenrodt, H.H. Brungs and G. Törner, *Right chain rings. Part 1*, Schriftenreihe des Fachbereich Mathematik, Universität Duisburg (1990).

[21] K. Bessenrodt, H.H. Brungs and G. Törner, *Right chain rings. Part 2a*, Schriftenreihe des Fachbereich Mathematik, Universität Duisburg (1992).

[22] K. Bessenrodt, H.H. Brungs and G. Törner, *Right chain rings. Part 2b*, Schriftenreihe des Fachbereich Mathematik, Universität Duisburg (1992).

[23] R.L. Blair, *Ideal lattice and the structure of rings*, Trans. Amer. Math. Soc. **75** (1) (1953), 136–153.

[24] H.H. Brungs, *Rings with distributive lattice of right ideals*, J. Algebra **40** (2) (1976), 392–400.

[25] H.H. Brungs, *Bezout domains and rings with a distributive lattice of right ideals*, Canad. J. Math. **38** (2) (1986), 286–303.

[26] H.H. Brungs, *Ideal theory in rings with a distributive lattice of right ideals*, Wiss. Beitr. M.-Luther-Univ. Halle. Wittenberg, No. 48 (1987), 73–80.
[27] H.H. Brungs and J. Gräter, *Valuation rings in finite-dimensional division algebras*, J. Algebra **120** (1989), 90–99.
[28] H.H. Brungs and J. Gräter, *Value groups and distributivity*, Canad. J. Math. **43** (6) (1991), 1150–1160.
[29] H.H. Brungs and J. Gräter, *Noncommutative Prüfer and valuation rings*, Proc. Int. Conf. Algebra Dedicat. Mem. A.I. Mal'cev, Novosibirsk (1989), Pt. 1, Providence, RI (1992), 253–269.
[30] H.S. Butts and R.C. Phillips, *Almost multiplication rings*, Canad. J. Math. **17** (1965), 267–277.
[31] C.E. Caballero, *Self-duality and ℓ-hereditary semidistributive rings*, Comm. Algebra **14** (10) (1986), 1821–1843.
[32] V. Camillo, *Distributive modules*, J. Algebra **36** (1) (1975), 16–25.
[33] R. Camps, A. Facchini and G.E. Puninski, *Serial rings that are endomorphism rings of Artinian modules*, Rings and Radicals, Pitman Series Sci. & Tech., Longman (1996).
[34] Q. Chen, *On graded distributive modules*, Northeast. Math. J. **10** (2) (1994), 273–278.
[35] C.W. Choi, *Multiplication modules and endomorphisms*, Math. J. Toyama Univ. **18** (1995), 1–8.
[36] A.W. Chatters, *Serial rings with Krull dimension*, Glasgow Math. J. **32** (1990), 71–78.
[37] L.G. Chouinard, B.R. Hardy and T.S. Shores, *Arithmetical semihereditary semigroup rings*, Comm. Algebra **8** (17) (1980), 1593–1652.
[38] P. Cohn, *Free Rings and Their Relations*, Academic Press, London (1971).
[39] R.R. Colby and K.R. Fuller, *Modules over diserial rings*, Comm. Algebra **9** (5) (1981), 511–532.
[40] T.M.K. Davison, *Distributive homomorphisms of rings and modules*, J. Reine Angew. Math. **271** (1974), 28–34.
[41] Yu.A. Drozd, *On generalized uniserial rings*, Mat. Zametki **18** (5) (1975), 707–710.
[42] N.I. Dubrovin, *Uniserial domains*, Vestn. MGU, Mat., Mekh., No. 2 (1980), 51–54.
[43] N.I. Dubrovin, *On uniserial rings*, Uspekhi Mat. Nauk **37** (4) (1982), 139–140.
[44] N.I. Dubrovin, *The rational closure of group rings of left-ordered groups*, Gerhard Mercator Universität Duisburg Gesamthochschule (1994).
[45] D. Eisenbud and P. Griffith, *Serial rings*, J. Algebra **17** (3) (1971), 268–287.
[46] D. Eisenbud and P. Griffith, *The structure of serial rings*, Pacific J. Math. **36** (1971), 109–121.
[47] P. Eklof and I. Herzog, *Model theory of modules over a serial ring*, Ann. Pure Appl. Logic **72** (1995), 145–176.
[48] Z.A. El-Bast and P.F. Smith, *Multiplication modules*, Comm. Algebra **16** (4) (1988), 755–779.
[49] Z.A. El-Bast and P.F. Smith, *Multiplication modules and theorems of Mori and Mott*, Comm. Algebra **16** (4) (1988), 781–796.
[50] V. Erdoğdu, *Distributive modules*, Canad. Math. Bull. **30** (2) (1987), 248–254.
[51] V. Erdoğdu, *Modules with locally linearly ordered distributive hulls*, J. Pure Appl. Algebra **47** (1987), 119–130.
[52] V. Erdoğdu, *Multiplication modules which are distributive*, J. Pure Appl. Algebra **54** (1988), 209–213.
[53] V. Erdoğdu, *Rings with principal distributive ideal*, Türk. Mat. Derg. **4** (2) (1990), 120–124.
[54] V. Erdoğdu, *Cyclically decomposable distributive modules*, Comm. Algebra **25** (5) (1997), 1635–1639.
[55] J. Escoriza and B. Torrecillas, *Multiplication modules relative to torsion theories*, Comm. Algebra **23** (11) (1995), 4315–4331.
[56] J. Escoriza and B. Torrecillas, *Relative multiplication and distributive modules*, Comment. Math. Univ. Carolin. **38** (1997), 201–225.
[57] J. Escoriza and B. Torrecillas, *Multiplication objects in commutative Grothendieck categories*, Comm. Algebra **26** (6) (1988), 1867–1883.
[58] A. Facchini and G.E. Puninski, *Σ-pure-injective modules over serial rings*, Abelian Groups and Modules, A. Facchini and C. Menini, eds, Kluwer Acad. Publishers, Dodrecht (1995), 145–162.
[59] A. Facchini and G.E. Puninski, *Classical localizations in serial rings*, Comm. Algebra **24** (11) (1996), 3537–3559.
[60] A. Facchini and L. Salce, *Uniserial modules: Sums and isomorphisms of subquotients*, Comm. Algebra **18** (2) (1990), 499–517.
[61] R.B. Feinberg, *Faithful distributive modules over incidence algebras*, Pacific J. Math. **65** (1) (1976), 35–45.

[62] M. Ferrero and G. Törner, *Rings with annihilator chain conditions and right distributive rings*, Proc. Amer. Math. Soc. **119** (1993), 401–405.
[63] M. Ferrero and G. Törner, *On the ideal structure of right distributive rings*, Comm. Algebra **21** (8) (1993), 2697–2713.
[64] M. Ferrero and G. Törner, *On waists of right distributive rings*, Forum Math. **7** (1995), 419–433.
[65] L. Fuchs and L. Salce, *Modules over valuation domains*, Lecture Notes in Pure and Appl. Math., Vol. 97 (1985).
[66] K.R. Fuller, *Generalized uniserial rings and their Kupisch series*, Math. Z. **106** (4) (1968), 248–260.
[67] K.R. Fuller, *Weakly symmetric rings of distributive module type*, Comm. Algebra **5** (9) (1977), 997–1008.
[68] K.R. Fuller, *Rings of left invariant module type*, Comm. Algebra **6** (1978), 153–167.
[69] K.R. Fuller, *On generalization of serial rings*, Proc. of the Filadelphia Conference on Rep. Thy, Lecture Notes in Pure and Appl. Math., Vol. 37, Dekker (1978), 359–368.
[70] K.R. Fuller, *Biserial rings*, Lecture Notes in Math., Vol. 734 (1979), 64–90.
[71] K.R. Fuller, *On generalization of serial rings II*, Comm. Algebra **8** (7) (1980), 635–661.
[72] K.R. Fuller, J. Haack and H. Hullinger, *Stable equivalence of uniserial rings*, Proc. Amer. Math. Soc. **68** (2) (1978), 153–158.
[73] R.W. Gilmer and J.L. Mott, *Multiplication rings as rings in whish ideals with prime radical are primary*, Trans. Amer. Math. Soc. **114** (1965), 40–52.
[74] K.R. Goodearl, *Artinian and Noetherian modules over regular rings*, Comm. Algebra **8** (5) (1980), 475–504.
[75] J. Gräter, *Zur Theorie Nicht Kommutativer Prüfer Ringer*, Arch. Math. **41** (1983), 30–36.
[76] J. Gräter, *Über die Distributivität des Idealverbandes Eines Kommutativem Ringes*, Monatsh. Math. **99** (4) (1985), 267–278.
[77] J. Gräter, *On noncommutative Prüfer rings*, Arch. Math. **46** (5) (1986), 402–407.
[78] J. Gräter, *Lokalinvariante Prüferringe*, Results Math. **9** (1986), 10–32.
[79] J. Gräter, *Ringe mit Distributivem Rechtidealverband*, Results Math. **12** (1987), 95–98.
[80] J. Gräter, *Strong right D-domains*, Monatsh. Math. **107** (3) (1989), 189–205.
[81] O.E. Gregul and V.V. Kirichenko, *On semihereditary serial rings*, Ukr. Mat. Zh. **39** (2) (1987), 156–161.
[82] M. Griffin, *Multiplication rings via their quotient rings*, Canad. J. Math. **26** (1974), 430–449.
[83] J.K. Haack, *Self-duality and serial rings*, J. Algebra **59** (2) (1979), 345–363.
[84] B.R. Hardy and T.S. Shores, *Arithmetical semigroup rings*, Canad. J. Math. **32** (6) (1980), 1361–1371.
[85] C. Herrmann, *On a condition sufficient for the distributivity of lattices of linear subspaces*, Arch. Math. **33** (3) (1979), 235–238.
[86] D.A. Hill, *Injective modules over non-Artinian serial rings*, J. Austral. Math. Soc. **44** (1988), 242–251.
[87] G. Ivanov, *Left generalized uniserial rings*, J. Algebra **31** (1974), 166–181.
[88] G. Ivanov, *Decomposition of modules over serial rings*, Comm. Algebra **3** (11) (1975), 1031–1036.
[89] C.U. Jensen, *A remark on arithmetical rings*, Proc. Amer. Math. Soc. **15** (6) (1964), 951–954.
[90] C.U. Jensen, *Arithmetical rings*, Acta Math. Acad. Sci. Hungar. **17** (1) (1966), 115–123.
[91] K.H. Kim and F.W. Roush, *Reflexive lattices of subspaces*, Proc. Amer. Math. Soc. **78** (1) (1980), 17–18.
[92] K.H. Kim and F.W. Roush, *Regular rings and distributive lattices*, Comm. Algebra **8** (13) (1980), 1283–1290.
[93] V.V. Kirichenko, *Generalized uniserial rings*, Mat. Sb. **99** (4) (1976), 559–581.
[94] V.V. Kirichenko, *On serial hereditary and semihereditary rings*, Zap. Nauchn. Sem. LOMI Akad. Nauk SSSR **114** (1982), 137–147.
[95] V.V. Kirichenko and O.Ya. Bernik, *Semiperfect rings of distributive module type*, Dokl. Akad. Nauk SSSR, -A (3) (1988), 15–17.
[96] V.V. Kirichenko and P.P. Kostyukevich, *Biserial rings*, Ukr. Mat. Zh. **38** (6) (1986), 718–723.
[97] V.V. Kirichenko and Yu.V. Yaremenko, *Noetherian biserial rings*, Ukr. Mat. Zh. **40** (4) (1988), 435–440.
[98] I.B. Kozhukhov, *Uniserial semigroup rings*, Uspekhi Mat. Nauk **29** (1) (1974), 169–170.
[99] H. Krause, *Maps between tree and band modules*, J. Algebra **137** (1991), 186–194.
[100] H. Kupisch, *A characterizaton of Nakayama rings*, Comm. Algebra **23** (2) (1995), 739–741.
[101] D. Latsis, *Localizaton dans les anneaux duos*, C. R. Acad. Sci. Paris **A282** (24) (1976), 1403–1406.
[102] Lee Sin Min, *On the constructions of local and arithmetical rings*, Acta. Math. Acad. Sci. Hungar. **32** (1–2) (1978), 31–34.

[103] L.R. Le Riche, *The ring* $R\langle X\rangle$, J. Algebra **67** (1980), 327–341.

[104] G.M. Low and P.F. Smith, *Multiplication modules and ideals*, Comm. Algebra **18** (12) (1990), 4353–4375.

[105] R. Makino, *QF-1 algebras with faithful direct sums of uniserial modules*, J. Algebra **136** (1991), 175–189.

[106] V.T. Markov, A.V. Mikhalev, L.A. Skornyakov and A.A. Tuganbaev, *Modules*, Progress of Science and Technology: Series on Algebra, Topology, and Geometry, Vol. 19, Itogi Nauki i Tekhn. VINITI, All-Union Institute for Scientific and Technical Information, Akad. Nauk SSSR, Moscow (1981), 31–134.

[107] V.T. Markov, A.V. Mikhalev, L.A. Skornyakov and A.A. Tuganbaev, *Endomorphism rings and submodule lattices*, Progress of Science and Technology: Series on Algebra, Topology, and Geometry, Vol. 21, Itogi Nauki i Tekhn. VINITI, All-Union Institute for Scientific and Technical Information, Akad. Nauk SSSR, Moscow (1983), 183–254.

[108] J. Martinez, *On commutative rings which are strongly Prüfer*, Comm. Algebra **2** (9) (1994), 3479–3488.

[109] R. Mazurek, *Distributive rings with Goldie dimension one*, Comm. Algebra **19** (3) (1991), 931–944.

[110] R. Mazurek and E.R. Puczylowski, *On nilpotent elements of distributive rings*, Comm. Algebra **18** (2) (1990), 463–471.

[111] R. Mazurek and E.R. Puczylowski, *On semidistributive rings*, Comm. Algebra **25** (11) (1997), 3463–3471.

[112] W. Menzel, *Über den Untergruppenverband einer Abelschen Operatorgruppe. Teil II. Distributive und M-Verbande von Untergruppen einer Abelschen Operatorgruppe*, Math. Z. **74** (1) (1960), 52–65.

[113] W. Menzel, *Ein Kriterium für Distributivität des Untergruppenverbands einer Abelschen Operatorgruppe*, Math. Z. **75** (3) (1961), 271–276.

[114] A.V. Mikhalev, *Special structural spaces of rings*, Dokl. Akad. Nauk SSSR **150** (7) (1963), 259–261.

[115] J.L. Mott, *Equivalent conditions for a ring to be a multiplication ring*, Canad. J. Math. **16** (1964), 429–434.

[116] J.L. Mott, *Multiplication rings containing only finitely many minimal primes*, J. Sci. Hiroshima Univ. Ser. A-I. **33** (1969), 73–83.

[117] B.J. Müller, *The structure of serial rings*, Methods in Module Theory, G. Abrams, J. Haefner and K.M. Rangaswamy, eds, Marcel Dekker, New York (1992), 249–270.

[118] B.J. Müller, *Goldie-prime serial rings*, Contemp. Math. **130** (1992), 301–310.

[119] B.J. Müller and S. Singh, *Uniform modules over serial rings II*, Lecture Notes in Math., Vol. 1448 (1989), 25–32.

[120] B.J. Müller and S. Singh, *Uniform modules over serial rings*, J. Algebra **144** (1991), 94–109.

[121] T. Nakayama, *Note on uniserial and generalized uniserial rings*, Proc. Imp. Acad. Tokyo **16** (1940), 285–289.

[122] T. Nakayama, *On Frobenius algebras. II*, Ann. Math. **42** (1) (1941), 1–21.

[123] A.G. Naoum, *The Ohm type properties for finitely generated multiplication ideals*, Period. Math. Hungar. **18** (4) (1986), 287–293.

[124] A.G. Naoum, *On the product of a module by an ideal*, Beiträge Algebra Geom. **27** (1988), 153–159.

[125] A.G. Naoum, *On the dual of a finitely generated multiplication module II*, Beiträge Algebra Geom. **27** (1988), 83–88.

[126] A.G. Naoum, *On the ring of endomorphisms of a finitely generated multiplication module*, Period. Math. Hungar. **21** (1990), 249–255.

[127] A.G. Naoum and F.H. Al-Alwan, *Dedekind modules*, Comm. Algebra **24** (2) (1996), 397–412.

[128] A.G. Naoum and F.H. Al-Alwan, *Dense submodules of multiplication modules*, Comm. Algebra **24** (2) (1996), 413–424.

[129] A.G. Naoum and M.M. Balboul, *On finitely generated multiplication ideals in commutative rings*, Beiträge Algebra Geom. **19** (1985), 75–82.

[130] A.G. Naoum and M.A.K. Hasan, *The residual of finitely generated multiplication modules*, Arch. Math. **46** (1986), 225–230.

[131] A.G. Naoum and K.R. Sharaf, *A note on the dual of a finitely generated multiplication module*, Beiträge Algebra Geom. **27** (1988), 5–11.

[132] R. Pierce, *Associative Algebras*, Springer, New York (1982).

[133] G.E. Puninskii, *Endodistributive and pure injective modules over uniserial rings*, Uspekhi Mat. Nauk **48** (3) (1993), 201–202.

[134] G.E. Puninskii, *Indecomposable pure-injective modules over uniserial rings*, Trudy Moskov. Mat. Obshch. **56** (1994), 1–13.

[135] G.E. Puninskii, *Pure-injective modules over right Noetherian serial rings*, Comm. Algebra **23** (4) (1995), 1579–1592.

[136] G.E. Puninskii, *Serial Krull–Schmidt rings and pure-injective modules*, Fund. Prikl. Mat. **1** (2) (1995), 471–490.

[137] G.E. Puninskii and R. Wisbauer, *Σ-injective modules over right duo and right distributive rings*, J. Pure Appl. Algebra **113** (1996), 55–66.

[138] G.E. Puninskii, R. Wisbauer and M. Yousif, *On p-injective rings*, Glasgow Math. J. **37** (1994), 373–378.

[139] V.B. Repnitskii, *Arithmetical modules*, Mat. Sb. **110** (1) (1979), 150–157.

[140] T. Shores and W.J. Lewis, *Serial modules and endomorphism rings*, Duke Math. J. **41** (4) (1974), 889–909.

[141] S. Singh and F. Mehdi, *Multiplication modules*, Canad. Math. Bull. **22** (1) (1979), 93–98.

[142] L.A. Skornyakov, *When are all modules serial?* Mat. Zametki **5** (2) (1969), 173–182.

[143] A. Skowronski and J. Waschbüsch, *Representation-finite biserial algebras*, J. Reine Angew. Math. **345** (1983), 172–181.

[144] P.F. Smith, *Some remarks on multiplication modules*, Arch. Math. **50** (3) (1988), 223–235.

[145] P.F. Smith, T. Ukegawa and N. Umaya, *On noncommutative Noetherian multiplication rings*, Math. Sem. Notes Kobe Univ. **11** (2) (1983), 259–276.

[146] W. Stephenson, *Modules whose lattice of submodules is distributive*, Proc. London Math. Soc. **28** (2) (1974), 291–310.

[147] G. Törner and J. Zima, *Some remarks on linear structures in right distributive rings*, Forum Math. **11** (1) (1999), 1–15.

[148] A.A. Tuganbaev, *Rings all modules over which are direct sums of distributive modules*, Vestn. MGU, Mat., Mekh., No. 1 (1980), 61–64.

[149] A.A. Tuganbaev, *Distributive Noetherian rings*, Vestn. MGU, Mat., Mekh., No. 2 (1980), 30–34.

[150] A.A. Tuganbaev, *Distributive modules*, Uspekhi Mat. Nauk **35** (5) (1980), 245–246.

[151] A.A. Tuganbaev, *Integrally closed rings*, Mat. Sb. **115** (4) (1981), 544–559.

[152] A.A. Tuganbaev, *Integrally closed Noetherian rings*, Uspekhi Mat. Nauk **36** (5) (1981), 195–196.

[153] A.A. Tuganbaev, *Distributive modules and rings*, Uspekhi Mat. Nauk **39** (1) (1984), 157–158.

[154] A.A. Tuganbaev, *Distributive rings*, Mat. Zametki **35** (3) (1984), 329–332.

[155] A.A. Tuganbaev, *Rings with distributive lattice of ideals*, Abelian Groups and Modules, No. 5, Tomsk (1985), 88–104.

[156] A.A. Tuganbaev, *Right distributive rings*, Izv. Vuzov, Mat., No. 1 (1985), 46–51.

[157] A.A. Tuganbaev, *Rings with flat right ideals and distributive rings*, Mat. Zametki **38** (2) (1985), 218–228.

[158] A.A. Tuganbaev, *Distributive rings and endodistributive modules*, Ukr. Mat. Zh. **38** (1) (1986), 63–67.

[159] A.A. Tuganbaev, *Rings whose lattice of right ideals is distributive*, Izv. Vuzov, Mat., No. 2 (1986), 44–49.

[160] A.A. Tuganbaev, *Rings with distributive lattice of right ideals*, Uspekhi Mat. Nauk **41** (3) (1986), 203–204.

[161] A.A. Tuganbaev, *Distributive series rings*, Mat. Zametki **39** (4) (1986), 518–528.

[162] A.A. Tuganbaev, *Distributive and semigroup rings*, Abelian Groups and Modules, No. 6, Tomsk (1986), 133–144.

[163] A.A. Tuganbaev, *Series rings and weak global dimension*, Izv. Vuzov, Mat., No. 11 (1987), 70–78.

[164] A.A. Tuganbaev, *Hereditary rings*, Mat. Zametki **41** (3) (1987), 303–312.

[165] A.A. Tuganbaev, *Distributive rings and modules*, Trudy Moskov. Mat. Obshch. **51** (1988), 95–113.

[166] A.A. Tuganbaev, *Distributive rings with the finiteness condition*, Izv. Vuzov, Mat., No. 10 (1988), 50–55.

[167] A.A. Tuganbaev, *On monoid rings*, Abelian Groups and Modules, No. 7, Tomsk (1988), 124–130.

[168] A.A. Tuganbaev, *Modules with distributive lattice of submodules*, Uspekhi Mat. Nauk **44** (1) (1989), 215–216.

[169] A.A. Tuganbaev, *Rings with distributive lattice of right ideals*, Abelian Groups and Modules, No. 8, Tomsk (1989), 124–127.

[170] A.A. Tuganbaev, *Distributive rings and modules*, Mat. Zametki **47** (2) (1990), 115–123.

[171] A.A. Tuganbaev, *Embeddings into serial modules*, Abelian Groups and Modules, No. 9, Tomsk (1990), 115–120.

[172] A.A. Tuganbaev, *Rings with projective principal right ideals*, Ukr. Mat. Zh. **42** (6) (1990), 861–863.

[173] A.A. Tuganbaev, *Serial rings and modules*, Mat. Zametki **48** (2) (1990), 99–106.

[174] A.A. Tuganbaev, *Elementary divisor rings and distributive rings*, Uspekhi Mat. Nauk **46** (6) (1991), 219–220.

[175] A.A. Tuganbaev, *Rings of quotients of distributive rings*, Abelian Groups and Modules, No. 10, Tomsk (1991), 148–152.

[176] A.A. Tuganbaev, *Distributive monoid algebras*, Mat. Zametki **51** (2) (1992), 101–108.

[177] A.A. Tuganbaev, *Quaternion algebras over local rings*, Uspekhi Mat. Nauk **47** (3) (1992), 179–180.

[178] A.A. Tuganbaev, *Generalized quaternion algebras*, Mat. Zametki **53** (5) (1993), 120–128.

[179] A.A. Tuganbaev, *Quaternion algebras over commutative rings*, Mat. Zametki **53** (2) (1993), 126–131.

[180] A.A. Tuganbaev, *Flat modules and distributive rings*, Ukr. Mat. Zh. **45** (5) (1993), 721–724.

[181] A.A. Tuganbaev, *On quaternion algebras*, Trudy Sem. im. I.G. Petrovskogo, No. 17, Moscow (1994), 209–219.

[182] A.A. Tuganbaev, *Endomorphism rings and distributivity*, Mat. Zametki **56** (4) (1994), 141–152.

[183] A.A. Tuganbaev, *On semigroup rings*, Abelian Groups and Modules, No. 11–12, Tomsk (1994), 218–225.

[184] A.A. Tuganbaev, *Endomorphism ring of a distributive module*, Uspekhi Mat. Nauk **50** (1) (1995), 215–216.

[185] A.A. Tuganbaev, *On left distributivity of right distributive rings*, Fund. Prikl. Mat. **1** (1) (1995), 289–300.

[186] A.A. Tuganbaev, *Endomorphisms of distributive modules*, Uspekhi Mat. Nauk **50** (3) (1995), 167–168.

[187] A.A. Tuganbaev, *Noetherian semiprime rings and distributivity*, Fund. Prikl. Mat. **1** (3) (1995), 767–779.

[188] A.A. Tuganbaev, *Right or left distributive rings*, Mat. Zametki **58** (4) (1995), 604–627.

[189] A.A. Tuganbaev, *Distributive semiprime rings*, Mat. Zametki **58** (5) (1995), 736–761.

[190] A.A. Tuganbaev, *Flat modules and distributivity*, Uspekhi Mat. Nauk **50** (6) (1995), 221–222.

[191] A.A. Tuganbaev, *Distributively decomposable rings*, Uspekhi Mat. Nauk **51** (3) (1996), 215–216.

[192] A.A. Tuganbaev, *Flat modules and rings that are finitely generated over its centre*, Mat. Zametki **60** (2) (1996), 254–277.

[193] A.A. Tuganbaev, *Piecewise domains and nonsingularly prime ideals*, Uspekhi Mat. Nauk **51** (4) (1996), 171–172.

[194] A.A. Tuganbaev, *Direct sums of distributive modules*, Mat. Sb. **187** (12) (1996), 137–156.

[195] A.A. Tuganbaev, *On distributive rings and modules*, First International Tainan–Moscow Algebra Workshop, Walter de Gruyter, Berlin–New York (1996), 323–328.

[196] A.A. Tuganbaev, *Semidistributive hereditary rings*, Uspekhi Mat. Nauk **52** (4) (1997), 215–216.

[197] A.A. Tuganbaev, *Ideals of distributive rings*, Fund. Prikl. Mat. **4** (2) (1998), 791–794.

[198] A.A. Tuganbaev, *Semidistributive Rings and Modules*, Kluwer Academic Publishers, Dordrecht (1998).

[199] A.A. Tuganbaev, *Rings with many idempotents*, Preprint (1999).

[200] A.A. Tuganbaev, *Semidistributive and distributively decomposable rings*, Mat. Zametki **65** (2) (1999), 307–313.

[201] D.A. Tuganbaev, *Uniserial Laurent series rings*, Fund. Prikl. Mat. **3** (3) (1997), 947–951.

[202] D.A. Tuganbaev, *Semilocal distributive Laurent skew series rings*, Preprint (1999).

[203] M. Uematsu and K. Yamagata, *On serial quasi-hereditary rings*, Hokkaido Math. J. **19** (1990), 165–174.

[204] T. Ukegawa, *Some properties of noncommutative multiplication rings*, Proc. Japan Acad. **54A** (9) (1978), 279–284.

[205] T. Ukegawa, *Left Noetherian multiplication rings*, Osaka J. Math. **17** (2) (1980), 449–453.

[206] T. Ukegawa, *Some remarks on M-rings*, Math. Japon. **28** (2) (1983), 195–203.

[207] M.H. Upham, *Serial rings with right Krull dimension one*, J. Algebra **109** (1987), 319–333.

[208] P. Vàmos, *Finitely generated Artinian and distributive modules are cyclic*, Bull. London Math. Soc. **10** (3) (1978), 287–288.

[209] N. Vanaja and V.N. Purav, *A note on generalized uniserial rings*, Comm. Algebra **21** (4) (1993), 1153–1159.

[210] E.M. Vechtomov, *Distributive rings of continuous functions*, Mat. Zametki **34** (34) (1983), 321–332.

[211] E.M. Vechtomov, *Rings of Continuous Functions and F-spaces, Selected Themes*, Moscow State Teachers' Training University Press, Moscow (1992).

[212] A. Vogel, *Spectrum of a locally cyclic modules over a commutative regular rings*, Comm. Algebra **9** (16) (1981), 1617–1637.

[213] B. Wald and J. Waschbüsch, *Tame biserial algebras*, J. Algebra **95** (1985), 480–500.

[214] R.B. Warfield, *Decomposability of finitely presented modules*, Proc. Amer. Math. Soc. **25** (1970), 167–172.

[215] R.B. Warfield, *Serial rings and finitely presented modules*, J. Algebra **37** (3) (1975), 187–222.

[216] R.B. Warfield, *Bezout rings and serial rings*, Comm. Algebra **7** (1979), 533–545.
[217] J. Waschbüsch, *Self-duality of serial rings*, Comm. Algebra **14** (4) (1986), 581–589.
[218] R. Wisbauer, *Modules locally of serial type*, Period. Math. Hungar. **18** (1987), 39–52.
[219] M.H. Wright, *Serial rings with right Krull dimension one, II*, J. Algebra **117** (1988), 99–116.
[220] M.H. Wright, *Krull dimension in serial rings*, J. Algebra **124** (1989), 317–328.
[221] M.H. Wright, *Links between right ideals of serial rings with Krull dimension*, Lecture Notes in Math., Vol. 1448 (1989), 33–40.
[222] M.H. Wright, *Certain uniform modules over serial rings are uniserial*, Comm. Algebra **17** (1989), 441–469.
[223] M.H. Wright, *Uniform modules over serial rings with Krull dimension*, Comm. Algebra **18** (8) (1990), 2541–2557.
[224] M.H. Wright, *Prime serial rings with Krull dimension*, Comm. Algebra **18** (8) (1990), 2559–2572.
[225] M.H. Wright, *Right locally distributive rings*, Ring Theory. Proc. Bien. Ohio State–Denison Conf., Granville, Ohio, May, 1992, Singapore (1993), 350–357.
[226] W. Xue, *Two examples of local Artinian rings*, Proc. Amer. Math. Soc. **107** (1) (1989), 63–65.
[227] W. Xue, *A note on three Artinian rings*, Comm. Algebra **18** (7) (1990), 2243–2247.
[228] W. Xue, *Exact modules and serial rings*, J. Algebra **134** (1990), 209–221.
[229] W. Xue, *Rings with Morita duality*, Lecture Notes in Math., Vol. 1523 (1992), 1–197.
[230] Y. Yukimoto, *Artinian rings whose projective indecomposables are distributive*, Osaka J. Math. **22** (1985), 339–344.
[231] B.V. Zabavskii and N.Ya. Komarnitskii, *Distributive elementary divisor domains*, Ukr. Mat. Zh. **42** (7) (1990), 1000–1004.
[232] V.V. Zharinov, *Distributive lattices and their applications in complex analysis*, Trudy Mat. Inst. Akad. Nauk SSSR **162** (1983), 3–116.

Serial and Semidistributive Modules and Rings

Askar A. Tuganbaev
Department of Mechanics and Mathematics, Moscow State University, Moscow, 119899 GSP Russia
E-mail: askar@tuganbaev.mccme.ru

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

Preface

All rings are assumed to be associative and (except for some explicitely indicated cases) to have a nonzero identity element. Expressions such as a "Noetherian ring" mean that the corresponding right and left conditions hold. A module M is *uniserial* if any two submodules of M are comparable with respect to inclusion. A direct sum of uniserial modules is a *serial* module. Any quasi-cyclic Abelian p-group is a uniserial module over the ring of integers $\mathbf{Z}$. All valuation rings in division rings are uniserial.

We point out [80,61,70–72,44,37], and [97] as among the first works concerning serial Artinian rings. The systematic study of serial non-Artinian rings was initiated in the papers [131] and [23].

A module is *distributive* if $F \cap (G + H) = F \cap G + F \cap H$ for any three submodules F, G, and H of M. A *semidistributive* module is any direct sum of distributive modules. All serial modules (in particular, all semisimple modules) are semidistributive.

Semidistributive rings and modules are considered in [16,20,38–43,56–58,60,69,102, 115–118,141,146], and [147]. Semidistributive and serial modules and rings were addressed in surveys [67,68], and [2].

Let F be a field, and let A be the 5-dimensional F-algebra generated by all (3×3)-matrices of the form

$$\begin{pmatrix} f_{11} & f_{12} & f_{13} \\ 0 & f_{22} & 0 \\ 0 & 0 & f_{33} \end{pmatrix},$$

where $f_{ij} \in F$.

Then A is an Artinian semidistributive left serial ring which is not right serial. (The right ideal of A generated by the local idempotent $e_{11}A$ is distributive and is not uniserial.)

A module M is *distributively generated* if M is a sum of distributive submodules. A module is *arithmetical* if the lattice of its fully invariant submodules is distributive. Every semidistributive module is arithmetical and distributively generated. Each Abelian group is a distributively generated over the ring of integers.

A semiperfect ring A is called a *right biserial* ring if for each local idempotent e of A, the right A-module eA contains uniserial submodules M and N (possibly equal to zero) such that $M + N$ is eA or the largest proper submodule of M and $M \cap N$ is zero or the smallest proper submodule of M. A semiperfect ring A is called a *right diserial* ring if A is right biserial and for each local idempotent e of A, the right A-module eA is distributive.

Diserial Artinian rings were studied in [20,38,41], and [42]. Diserial non-Artinian rings were studied in [56,57], and [58]. Biserial rings were considered in [60,82,83,92,93,98,99, 128], and [129].

We denote by $J(M)$, $\mathrm{End}(M)$, and $\max(M)$ the Jacobson radical, the endomorphism ring, and the set of all maximal submodules of a module M, respectively. For a subset B of a ring A, we denote by $r(B)$ the right annihilator of B in A.

1. Serial modules

1.1 [81,97]. All right modules over a ring A are serial $\Leftrightarrow$
all left A-modules are serial $\Leftrightarrow$
A is a serial Artinian ring.

1.2 [23,131]. A is a serial ring $\Leftrightarrow$
all finitely presented right A-modules are serial $\Leftrightarrow$
all finitely presented left A-modules are serial $\Leftrightarrow$
every finitely presented right A-module is a direct sum of uniserial modules of the form $e_jA/e_ja_{ji}e_iA$, where e_j are local idempotents of A $\Leftrightarrow$
every finitely presented left A-module is a direct sum of uniserial modules of the form $Ae_j/Ae_ia_{ji}e_j$, where e_j are local idempotents of A $\Leftrightarrow$
A is semiperfect, all finitely presented indecomposable right A-modules are local, and all finitely presented indecomposable left A-modules are local.

1.3 [53]. All finitely generated right modules over a ring A are serial $\Leftrightarrow$
A is serial, and all uniform right A-modules are uniserial $\Leftrightarrow$
all 2-generated right A-modules are serial $\Leftrightarrow$
A is right semiuniform, and all indecomposable injective right A-modules are uniserial $\Leftrightarrow$
A is right serial, and for every isomorphism $f : M \to N$ between arbitrary submodules M and N of any primitive cyclic A-modules xA and yA, respectively, either f or f^{-1} can be extended to a homomorphism between xA and yA.

1.4 [131]. A is a serial right Noetherian ring $\Leftrightarrow$
A is right Noetherian, and each finitely generated right A-module is serial $\Leftrightarrow$
A is right semiuniform, and each injective right A-module is serial.

A module M is *hereditary* (resp. *semihereditary*) if all submodules (resp. all finitely generated submodules) of M are projective.

1.5 [109]. Let A be a semiprime ring, and let Q be the maximal right ring of quotients of A. Then
each injective right A-module is serial $\Leftrightarrow$
A is right hereditary, and Q_A is a finite direct sum of uniserial right A-modules $\Leftrightarrow$
A is a serial right Noetherian ring.

1.6 [109]. If all injective right modules over a ring A are serial, then A is a semiperfect right Noetherian ring, and for the prime radical N of A, the factor ring A/N is a serial semiprime ring.

A ring is *right invariant* if all its right ideals are ideals.

1.7 [109]. A ring A is right invariant, and each injective right A-module is serial $\Leftrightarrow$
A is left invariant, and each injective left A-module is serial $\Leftrightarrow$
A is a finite direct product of uniserial Noetherian rings.

1.8 [96]. Let A be an Artinian ring such that the ring $A/J(A)$ is a finitely generated modules over its centre. Then

every injective left A-module is serial $\Leftrightarrow$
A is right serial.

1.9 [142]. (1) There exists a local Artinian ring A such that A is not right serial but all injective left A-modules are serial.

(2) There exists a local Artinian right serial ring R such that the indecomposable injective left A-module is not serial.

1.10 [51]. Let A be a left serial ring and suppose that for each primitive idempotent e of A, the ring eAe has its indecomposable injective left modules uniserial. Then
every finitely generated A-module is serial $\Leftrightarrow$
the injective hull of each simple left A-module is uniserial $\Leftrightarrow$
every indecomposable injective left A-module is uniserial.

1.11 [23]. For a semiperfect ring A, the following assertions hold.

(1) If all finitely presented indecomposable right A-modules have indecomposable projective hulls, then A is left serial.

(2) A is right serial $\Leftrightarrow$
for any three local projective right modules P, P_1, and P_2 and for any two homomorphisms $f_1 : P_1 \to P$ and $f_2 : P_2 \to P$, either there is $x \in \mathrm{Hom}(P_1, P_2)$ such that $f_1 = f_2 x$ or there is $y \in \mathrm{Hom}(P_2, P_1)$ such that $f_2 = f_1 y \Leftrightarrow$
for any three local projective left modules Q, Q_1, and Q_2 and for any two homomorphisms $f_1 : Q \to Q_1$ and $f_2 : Q \to Q_2$, either there is $x \in \mathrm{Hom}(Q_2, Q_1)$ such that $f_1 = x f_2$ or there is $y \in \mathrm{Hom}(Q_1, Q_2)$ such that $f_2 = y f_1$.

1.12 [45]. (1) There exists a unit-regular ring which has a uniserial Noetherian cyclic non-Artinian faithful module.

(2) There exists a unit-regular ring which has a uniserial Artinian cyclic non-Noetherian faithful module.

1.13 [30]. Let $n \geqslant 2$ be an integer. Then there exists a serial ring A and n^2 pairwise non-isomorphic finitely presented uniserial A-modules M_{ij} $(i, j = 1, \dots, n)$ such that for any two permutations σ, τ of $\{1, \dots, n\}$,

$$M_{11} \oplus M_{22} \oplus \cdots \oplus M_{nn} \cong M_{\sigma(1)\tau(1)} \oplus M_{\sigma(2)\tau(2)} \oplus \cdots \oplus M_{\sigma(n)\tau(n)}.$$

1.14 [89]. For every uniserial ring, a decomposition of a finitely presented module into a direct sum of uniserial modules is unique.

1.15 [30]. Let M be a nonzero uniserial right module, E be the endomorphism ring of M, S be the subset of E consisting of all the endomorphisms of M that are not monomorphisms, and let T be the subset of E consisting of all the endomorphisms of M that are not epimorphisms.

Then S and T are completely prime ideals of E, every right (or left) proper ideal of E is contained either in S or in T, and either

(a) the ideals S and T are comparable, E is a local ring, and $S \cup T$ is the Jacobson radical of E, or

(b) the ideals S and T are not comparable, $S \cap T$ is the Jacobson radical $J(E)$ of E, and $E/J(E)$ is canonically isomorphic to the direct product $E/S \times E/T$ of two division rings E/S and E/T.

1.16 [30]. Let M be a finite direct sum of uniserial modules, and let X, Y be two arbitrary modules.

Then $M \oplus X \cong M \oplus Y$ implies $X \cong Y$.

Let $1 = e_1 + \cdots + e_n$ be a decomposition of the identity element of a semiperfect ring A into a sum of local orthogonal idempotents e_i. The ring A is *semi-invariant* if either $A_i a \subseteq a A_j$ or $a A_j \subseteq A_i a$ for all i, j and any $a \in e_i A e_j$.

1.17 [87]. Let A be a serial ring. Then

A is semi-invariant $\Leftrightarrow$

$J(e_i A e_i)a \subseteq a e_j A e_j$ and $a J(e_j A e_j) \subseteq e_i A e_i a$ for any $a \in e_i A e_j$ $\Leftrightarrow$

$(e_i + J(e_i A e_i))a \subseteq a U(e_j A e_j)$ and $a(e_j + J(A_j)) \subseteq U(e_i A e_i)a$ for any nonzero $a \in e_i A e_j$.

1.18 [90]. Let A be a serial ring. Then

every finitely presented indecomposable R-module has a local endomorphism ring $\Leftrightarrow$

R is semi-invariant.

1.19 [131]. Every finitely generated nonsingular module over a serial ring is a serial projective module.

A module M is *pure-projective* if M is a direct summand of direct sum of finitely presented modules. A right module X over a ring A is a *pure-injective* module if for every module M_A and each pure submodule N of M, all homomorphisms $N \to X$ can be extended to homomorphisms $M \to X$. A module M is *Σ-pure-injective* if for any index set I, the direct sum $M^{(I)}$ is a pure-injective module.

1.20 [29, p. 153–154]. Let e be a local idempotent of a serial ring A.

(1) If M is a pure projective or pure-injective A-module, then Me is a distributive left End(M)-module.

(2) If M is a pure-injective indecomposable A-module, then Me is a uniserial left End(M)-module.

1.21 [88]. Every pure-injective or pure-projective right module over a uniserial ring is a distributive left module over its endomorphism ring.

Every indecomposable pure-injective right module over a uniserial ring is a uniserial left module over its endomorphism ring.

1.22 [32]. Every Σ-pure-injective module over a serial ring is serial.

1.23 [32]. Every indecomposable Σ-pure-injective right module M over a serial ring A is a Σ-injective uniserial $A/r(M)$-module and Artinian left End(M)-module.

1.24 [117]. (1) Every finitely generated module over a serial ring is a sum of a singular finitely generated module and a serial projective finitely generated module (the sum is not necessarily a direct sum).

(2) For each nonsingular module M over a serial ring, all submodules of M are flat.

1.25 [121]. Let M_A be a finitely generated module over a serial ring A, and let the singular submodule of M be a superfluous submodule in M.

Then M is a projective nonsingular serial module.

1.26 [104,105,111]. M is a distributive right module over a ring A $\Leftrightarrow$

for any indecomposable quasi-injective module E_A, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is uniserial $\Leftrightarrow$

for any module E which is the injective hull of an arbitrary simple subfactor of M, the left $\mathrm{End}(E)$-module $\mathrm{Hom}(M, E)$ is uniserial.

1.27 [104,105,111]. All indecomposable quasi-injective right modules over a ring A are uniserial left modules over their endomorphism ring $\Leftrightarrow$

the injective hull of each simple right A-module is a uniserial left module over its endomorphism ring $\Leftrightarrow$

A is right distributive.

1.28 [39]. Let M be a right module over a semiperfect ring A, and let $1 = \sum_{i=1}^{n} e_i$ be a decomposition of the identity element of A into a sum of local orthogonal idempotents. Then

Me_i is a uniserial right e_iAe_i-module for any e_i $\Leftrightarrow$

M is a distributive A-module.

A module is a *Bezout* module if every finitely generated submodule of it is cyclic.

1.29 [105,101]. (1) Let all simple subfactors of a module M_A be isomorphic (this is the case if A is a matrix-local ring). Then

M is uniserial $\Leftrightarrow$ M is distributive.

(2) Let M be a module over a local ring A. Then

M is uniserial $\Leftrightarrow$

M is a Bezout module $\Leftrightarrow$

M is distributive.

1.30 [117]. Let M be a distributive right module over a right serial ring A.

(1) M is a finite direct sum of uniform modules, and M is a completely finite-dimensional Bezout module.

(2) If all cyclic submodules of M are serial, then M is a finite direct sum of uniserial modules.

Therefore, the class of all semidistributive right A-modules coincides with the class of all serial right A-modules.

A subset T in a ring A is a *right denominator set* in A if there are a ring A_T and a ring homomorphism $f_T \equiv f : A \to A_T$ such that $f(T) \subseteq U(A_T)$, $A_T = \{f(a)f(t)^{-1} \mid a \in A,\ t \in T\}$, and $\mathrm{Ker}(f) = K(T)$. In this case, A_T is called the *right ring of quotients* for A with respect to T, and f_T is the *natural homomorphism*. The right A_T-module $N \otimes_A A_T$ is the *module of quotients for N with respect to T* and is denoted by N_T. The A-module homomorphism $g : N \to N_T$ such that $g(n) = n \otimes 1$ is denoted by g_T. If $A \setminus T \equiv M$ is a right ideal of A, then we use f_M, A_M, and N_M instead of f_T, A_T, and N_T. A *right localizable* ring is any ring A such that the right ring of quotients A_M exists for each $M \in \max(A_A)$. Every commutative ring A is localizable.

1.31 [107]. Let N be a right module over a right localizable ring A. Then
N_M is a uniserial A_M-module for all $M \in \max(A_A)$ $\Leftrightarrow$
N is a distributive A-module.

2. Serial rings

Let A be a semiperfect ring, $P = \bigoplus_{i=1,\dots,m} e_i A$ and $Q = \bigoplus_{j=1,\dots,n} e_j A$ be finitely generated projective right A-modules, where e_i and e_j are local idempotents, and let $f : P \to Q$ be a homomorphism. Denote by A_{ij} the Abelian group $e_i A e_j$. Since every homomorphism from $e_i A$ to $e_j A$ is given by left multiplication by an element of A_{ij}, f can be written as left multiplication by the $m \times n$ matrix (f_{ij}), where $f_{ij} \in A_{ij}$. From the case $P = Q = A_A$, we obtain that every $a \in A$ can be written in the form $a = \sum_{i,j} a_{ij}$, where $a_{ij} = e_i a e_j \in A_{ij}$. Hence a can be considered as the matrix $M(a) = (a_{ij})$, and $M(ab) = M(a) \cdot M(b)$ for all $a, b \in A$.

2.1 [23]. Let A be a serial ring, and let $P = \bigoplus_{i=1,\dots,n} e_i A$ and $Q = \bigoplus_{j=1,\dots,m} e_j A$ be two finitely generated projective right A-modules, where e_i and e_j are local idempotents of A, and $f = (f_{ij}) : P \to Q$ be a homomorphism with $f_{ij} \in A_{ij}$.

(1) There are automorphisms $\alpha : P \to P$ and $\beta : Q \to Q$ such that α and β are given by matrices which are products of elementary matrices, and $\beta f \alpha : P \to Q$ is given by a matrix $G = (g_{ij})$ such that G has at most one nonzero element in every row and every column.

(2) For every $a \in A$, there are $u, v \in U(A)$ such that all rows and columns of the matrix $M(uav)$ contain at most one nonzero element, and the matrices $M(u)$ and $M(v)$ are products of elementary matrices.

In addition, for some positive integer m, the matrix $M((uav)^m)$ is diagonal.

2.2 [46,117]. Let $1 = e_1 + \cdots + e_n$ be a decomposition of the identity element of a serial ring A into a sum of local idempotents e_i, T_i be the set of all regular elements in the ring $e_i A e_i$, and let T be the set of all elements t in A of the form $t = t_1 + \cdots + t_n$, where $t_i \in T_i$.

(1) T is an Ore set in A, the ring of quotients Q of A with respect to T is a serial ring, all $e_i Q$ are uniserial Q-modules, all e_i are local idempotents in Q, all the rings $e_i A e_i$ and $e_i Q e_i$ are uniserial, $e_i T e_i = T_i$ is an Ore set in $e_i A e_i$, $e_Q e_i$ is a classical ring of quotients of $e_i A e_i$, and $e_Q e_i$ is the ring of quotients of $e_i A e_i$ with respect to $e_i T e_i$ for all i.

(2) A has a classical ring of quotients $Q_{\mathrm{cl}}(A)$ which is a serial ring and coincides with the ring of quotients Q.

(3) Let each ring $e_i A e_i$ be either a ring with the maximum condition on right annihilators or a ring with the maximum condition on left annihilators.

Then A has the Artinian serial classical ring of quotients $Q_{\mathrm{cl}}(A)$ which coincides with the ring of quotients Q with respect to the set T.

It can happen that a serial ring has no Artinian classical ring of quotients. Let A be a uniserial prime ring with zero divisors (an example of such a ring is given, e.g., in [26]). Assume that A has a right or left Artinian classical ring of quotients Q. Then Q is a uniserial ring. In addition, Q is isomorphic to a matrix ring over a division ring. Therefore, Q is a division ring, and A is a domain; this is a contradiction.

2.3 [132]. Let A be a serial ring. Then
A is a ring with the maximum condition on right annihilators $\Leftrightarrow$
A is a ring with the maximum condition on left annihilators $\Leftrightarrow$
A has a serial Artinian classical ring of quotients.

2.4 [131]. Every indecomposable serial nonsingular ring is a two-sided order in a ring of block upper triangular matrices over a division ring.

2.5 [131,117]. (1) Let A be a right serial ring. Then
A is right nonsingular $\Leftrightarrow$
A is right semihereditary $\Leftrightarrow$
A is left semihereditary.
(2) Let A be a serial ring. Then
A is right nonsingular $\Leftrightarrow$
A is left nonsingular $\Leftrightarrow$
A is semihereditary.

2.6 [21]. Let A be a right serial left Noetherian ring. Then
A is right Noetherian $\Leftrightarrow$
$\bigcap_{n=1}^{\infty}(J(A))^n = 0$.

2.7 [21]. Let A be a right serial ring with $\bigcap_{n=1}^{\infty}(J(A))^n = 0$.
Then the $J(A)$-adic completion of A is a right serial right Noetherian ring.

2.8 [95]. If A is an indecomposable serial right Noetherian ring which is not left Noetherian, then A has only one non-maximal prime ideal, and there exists a unique simple projective right A-module.

2.9 [117]. A is a semiprime right serial left Noetherian ring $\Leftrightarrow$
A is a semiprime left serial right Noetherian ring $\Leftrightarrow$
A is a finite direct product of serial Noetherian hereditary prime rings.

2.10 [46]. Any indecomposable serial semiprime ring is a finite direct product of prime rings.

2.11 [131]. Every serial basis indecomposable Noetherian non-Artinian ring is isomorphic to a ring

$$H_n(V) = \begin{pmatrix} V & V & \dots & V \\ J & V & \dots & V \\ \vdots & \vdots & \ddots & \vdots \\ J & J & \dots & V \end{pmatrix},$$

where V is a Noetherian uniserial non-Artinian domain and $J = J(V)$.

Let J be an ideal of a ring A. For any ordinal α, let us define the ideal J_α of A as follows: $J_0 \equiv J$, $J_\alpha \equiv \bigcap_{\beta<\alpha} J_\beta$ if α is a limit ordinal, and $J_{\alpha+1} \equiv \bigcap_{n\in\omega} J_\alpha^n$.

2.12 [78,136]. Let A be a serial ring, $J \equiv J(A)$, and let α be an ordinal. Then
A has right Krull dimension $\alpha \Leftrightarrow$
A has left Krull dimension $\alpha \Leftrightarrow$

α be is least ordinal such that the ideal J_α is nilpotent.

2.13 [18]. Let A be an indecomposable serial ring with Krull dimension and suppose that A has a non-nilpotent ideal X such that $\bigcap_{n=1}^{\infty} X^n = 0$. Then A is a prime ring.

2.14 [18]. If A is a serial ring, then $A/\bigcap_{n=1}^{\infty}(J(A))^n$ is a finite direct product of Artinian rings and prime rings.

2.15 [131,126]. Let $1 = e_1 + \cdots + e_n$ be a decomposition of the identity element of a serial nonsingular ring A into a sum of local orthogonal idempotents e_i.

(1) A is a prime ring $\Leftrightarrow$ $e_i A e_j \neq 0$ for all i and j.

(2) A is an indecomposable ring $\Leftrightarrow$ for all i and j, we have that either $e_i A e_j \neq 0$ or $e_j A e_i \neq 0$.

2.16 [17]. Let A be a serial ring. Then

A is isomorphic to the endomorphism ring of an Artinian module $\Leftrightarrow$

A has finite Krull dimension and satisfies the maximum condition on right annihilators.

2.17 [31, Prop. 5.2]. Let A be a right uniserial ring.

(1) If B is a nonzero proper ideal of A and $B^2 = B$, then B is a completely prime ideal of A.

(2) If B is a proper ideal of A which is not nilpotent, then $\bigcap_{n=1}^{\infty} B^{\infty}$ is a completely prime ideal of A.

(3) If B is a completely prime ideal of A and $a \in A \setminus B$, then $B = Ba$.

(4) If B is a completely prime ideal which is right principal, then either $B = J(A)$ or $B = 0$.

2.18 [59]. Let A be a ring, and let G be a monoid. Then

the monoid ring $A[G]$ is right uniserial $\Leftrightarrow$

either A is right uniserial and $|G| = 1$ or A is a division ring of characteristic p, and G is a subgroup of the group $Z(p^{\infty})$.

2.19 [73]. Let k be a field of characteristic p, G be a finite group with the normal cyclic Sylow p-subgroup P.

Then the group ring $k[G]$ is a serial ring.

Let A be a ring, $B(A)$ be the set of all central idempotents of A, and let $S(A)$ be the (nonempty) set of all proper ideals of A generated by central idempotents. An ideal $P \in S(A)$ is said to be a *Pierce ideal* of A if P is a maximal (with respect to inclusion) element of the set $S(A)$. If P is a Pierce ideal of A, then the quotient ring A/P is said to be a *Pierce stalk* of A.

2.20 [119]. All Pierce stalks of a ring A are right uniserial rings $\Leftrightarrow$

A is a right distributive exchange ring.

Let a and b be two units of a commutative ring A. We denote by $(a, b/A)$ the *generalized quaternion algebra* over A, i.e., the free A-module with the canonical basis $\{1, i, j, k\}$, in which multiplication is A-bilinear, and is defined on the canonical basis such that 1 is the common identity element of the rings A and $(a, b/A)$, and $i^2 = a$, $j^2 = b$, $k^2 = -ab$, $ij = -ji = k$, $ik = -ki = aj$, and $kj = -jk = bi$.

2.21 [110]. Let a and b be two units of a commutative ring A, and let $2^{-1} \in A$. Then

$(a, b/A)$ is right uniserial $\Leftrightarrow$
$(a, b/A)$ is right invariant, and A is local $\Leftrightarrow$
$(a, b/A)$ is right distributive, and A is local $\Leftrightarrow$
$(a, b/A)$ is right quasi-invariant, and A is uniserial $\Leftrightarrow$
A is uniserial, and for any $x, y, z \in A$, the inclusion $x^2 - ay^2 - bz^2 \in J(A)$ is possible only for $x \in J(A), y \in J(A), z \in J(A)$.

2.22 [110]. Let a and b be two units of a commutative ring A, and let $2^{-1} \in A$. Then
$(a, b/A)$ is a right uniserial ring $\Leftrightarrow$
$(a, b/A)$ is a local right invariant ring $\Leftrightarrow$
$(a, b/A)$ is a uniserial invariant ring $\Leftrightarrow$
A is uniserial, and for any $x, y, z \in A$, the inclusion $x^2 - ay^2 - bz^2 \in M$ is possible only for $x, y, z \in J(A)$.

If φ is an injective endomorphism of a ring A, then we denote by $A_\ell[[x, \varphi]]$ the *left skew (power) series ring* consisting of formal series $\sum_{i=0}^{\infty} a_i x^i$ of the variable x with canonical coefficients $a_i \in A$, where addition is defined naturally and multiplication is defined by the rule $x^i a = \varphi^i(a)x^i$. If φ is an automorphism of a ring A, then the *left skew Laurent series ring* $A_\ell((x, \varphi))$ consists of the series $f \equiv \sum_{i=m}^{\infty} a_i x^i$, where $m = m(f)$ is an integer, and $x^i a = \varphi^i(a)x^i$. Analogously, we define the *right skew (power) series ring* $A_r[[x, \varphi]]$ and the *right skew Laurent series ring* $A_r((x, \varphi))$.

2.23 [107]. Let φ be an injective endomorphism of a ring A. Then
$A_r[[x, \varphi]]$ is a right uniserial ring $\Leftrightarrow$
A is right uniserial, and $\varphi(a)$ is a unit of A for any nonzero $a \in A$.

2.24 [122,123]. Let φ be an automorphism of a ring A. Then
$A_\ell((x, \varphi))$ is a right uniserial ring $\Leftrightarrow$
$A_\ell((x, \varphi))$ is a right uniserial right Artinian ring $\Leftrightarrow$
A is a right uniserial right Artinian ring.

2.25 [124]. Let φ be an automorphism of a ring A. Then
$A_\ell((x, \varphi))$ is a semilocal right distributive ring $\Leftrightarrow$
$A_\ell((x, \varphi))$ is a finite direct product of right uniserial rings $\Leftrightarrow$
$A_\ell((x, \varphi))$ is a finite direct product of right Artinian right uniserial rings $\Leftrightarrow$
A is a finite direct product of right Artinian right uniserial rings $A_1, \ldots, A_n$, and $\varphi(A_i) = A_i$ for all i.

2.26 [103,112]. A is a reduced right distributive ring $\Leftrightarrow$
for each $M \in \max(A_A)$, the right ring of quotients A_M exists and is a right uniserial domain.

3. Semidistributive modules and rings

3.1 [42,102]. All right A-modules are semidistributive $\Leftrightarrow$
all left A-modules are semidistributive $\Leftrightarrow$
A is an Artinian ring, and each right A-module is a direct sum of distributive completely cyclic modules with a composition series $\Leftrightarrow$

A is an Artinian ring with basis idempotent e, and for any right A-module M, the right eAe-module Me_{eAe} is a direct sum of completely cyclic modules with a composition series.

3.2 [42]. Let A be a diserial Artinian ring.
Then every indecomposable injective A-module is distributive, every local A-module is quasi-projective, and every uniform A-module is quasi-injective.

3.3 [41]. Let A be an Artinian right serial ring. Then
every right A-module is semidistributive $\Leftrightarrow$
every right A-module is a direct sum of distributive uniform modules $\Leftrightarrow$
every left A-module is a direct sum of distributive local modules $\Leftrightarrow$
A is a diserial ring.

3.4 [117]. All right modules over a ring A are semidistributive, and the ring $A/J(A)$ is normal $\Leftrightarrow$
A is an Artinian ring, and each right A-module is a direct sum of completely cyclic modules with a composition series.

3.5 [117]. Let A be a right invariant ring such that all injective right A-modules are semidistributive.
Then A is a right distributive right Noetherian ring.

3.6 [117]. A is a domain, and all injective right A-modules and all injective left A-modules are semidistributive $\Leftrightarrow$
A is a left finite-dimensional domain, and all injective right A-modules are semidistributive $\Leftrightarrow$
A is an invariant Noetherian hereditary domain.

3.7 [109]. Let A be an invariant ring. Then
each right A-module is isomorphic to a submodule of a semidistributive module $\Leftrightarrow$
all injective right A-modules are semidistributive $\Leftrightarrow$
A is a finite direct product of uniserial Artinian rings and hereditary Noetherian domains.

Serial and semidistributive rings with Morita duality were considered in [22,47,133, 144], and [146].

A module M is called *linearly compact* in the case any finitely solvable congruence $m \equiv m_i \bmod(M_i)$ is solvable, where M_i's are submodules and $m_i \in M$.

3.8 [146]. Every semidistributive linearly compact Noetherian ring A has a Morita duality.
Morever, the injective hull of each simple left (resp. right) A-module is distributive.

3.9 [39,105]. A is a semiperfect ring right semidistributive ring $\Leftrightarrow$
e_jAe_i is a right uniserial e_iAe_i-module for any e_i and any e_j.

3.10 [115]. Let M_A be a finite direct sum of n distributive modules $M_1, \ldots, M_n$, and let P be the prime radical of $\mathrm{End}(M)$.
(1) P contains all right or left nil-ideals of $\mathrm{End}(M)$.
(2) If M is nonsingular, then $P^n = 0$.
In addition, if $\mathrm{End}(M)$ is also a prime ring, then all the rings $\mathrm{End}(M_i)$ are domains.

(3) If $\operatorname{End}(M)$ is a right or left Rickartian ring, then $P^n = 0$.

3.11 [116]. Let A be a semilocal ring, and let A_A be an essential extension of a finite sum of distributive right ideals.

Then A is right finite-dimensional.

In addition, if A is right hereditary, then A is right Noetherian.

3.12 [116]. Let A be a semiperfect ring, and let all 2-generated right ideals of A be projective. Then

A is right semiuniform $\Leftrightarrow$

A is right serial.

3.13 [116]. Any semiperfect semiprime right semihereditary right finite-dimensional ring is a right serial ring.

3.14 [106,116]. Let A be a semiperfect semiprime right hereditary ring. Then

A is right Noetherian $\Leftrightarrow$

A is right serial $\Leftrightarrow$

A is right semidistributive.

Let A be a ring, and let $1 = \sum_{i=1}^{n} e_i$, where e_i are orthogonal idempotents. The ring A is called a *piecewise domain* (*with respect to* $\{e_1, \dots, e_n\}$) if the following three equivalent conditions hold.

(i) For every e_i and e_k, each nonzero homomorphism $e_iA \to e_kA$ is a monomorphism.

(ii) For every e_i, each nonzero homomorphism $e_iA \to A_A$ is a monomorphism.

(iii) $ab \neq 0$ for any nonzero elements $a \in e_jAe_i$, $b \in e_iAe_k$ and for all e_i, e_j, e_k.

Since condition (iii) is symmetric, conditions (i) and (ii) can be replaced by their left-side analogs.

A proper ideal B of A is called *piecewise integral* with respect to $\{e_1, \dots, e_n\}$ if the ring A/B is a piecewise domain with respect to the natural images of $\{e_1, \dots, e_n\}$ in A/B.

A proper ideal B of A is *nonsingularly prime* (resp. *nonsingularly semiprime*) if A/B is a prime right nonsingular (resp. semiprime right nonsingular) ring.

3.15 [115]. Let the identity element of a ring A be a sum of nonzero orthogonal idempotents $e_1, \dots, e_n$, and let all right A-modules e_iA be distributive.

(1) The prime radical P of A contains all right nil-ideals and all left nil-ideals of A.

(2) If A is right nonsingular, then $P^n = 0$, and all the rings e_iAe_i are reduced.

(3) If A is a prime right nonsingular ring, then all the rings e_iAe_i are domains.

(4) If A is prime, then all modules e_iA_A are uniform, and therefore A is right semiuniform, and the right Goldie dimension of A is equal to n.

(5) If F is a nonsingularly prime ideal of A, then for any $e_i \in A \setminus F$, the quotient ring e_iAe_i/e_iFe_i is a domain.

(6) If A is prime and all the rings e_iAe_i are division rings, then all e_iA are minimal right ideals of A, and therefore A is isomorphic to the ring of $n \times n$-matrices over a division ring.

(7) If the ring A is simple, then A is isomorphic to the ring of $n \times n$-matrices over a division ring.

(8) If A is a regular prime ring, then A is isomorphic to the ring of $n \times n$-matrices over a division ring.

(9) If B is an ideal of A and $h : A \to A/B$ is the natural epimorphism, then the quotient ring A/B is a direct sum of right distributive ideals $h(e_i A)$.

3.16 [115]. Let A be a right semidistributive ring which is a direct sum of n right distributive ideals $e_1 A, \ldots, e_n A$, where the e_i are orthogonal idempotents in A.

(1) If M is a maximal ideal of A, then A/M is isomorphic to the ring of $m \times m$-matrices over a division ring, where $m \leqslant n$.

(2) If B is the intersection of some set of maximal ideals of A, then the nilpotency index of any nilpotent element of the quotient ring A/B does not exceed n.

(3) Let F be a prime ideal of A. Then

F is a nonsingularly prime ideal $\Leftrightarrow$ F is piecewise integral with respect to $\{e_1, \ldots, e_n\}$ $\Leftrightarrow$ for any $e_i \in A \setminus F$, the ideal $e_i F e_i$ of $e_i A e_i$ is completely prime $\Leftrightarrow$ the ideal $e_i F e_i$ of $e_i A e_i$ is completely prime for one or more $e_i \in A \setminus F$.

(4) Let all left ideals Ae_i be distributive, and let F be a prime ideal of A. Then

F is nonsingularly prime $\Leftrightarrow$ A/F is left nonsingular.

(5) If F and G are two nonsingularly prime ideals of A, then either $F + G = A$ or the ideals F and G are comparable.

(6) Every proper nonsingularly prime ideal F of A contains precisely one minimal nonsingularly prime ideal.

3.17 [115]. Let A be a regular right semidistributive ring which is a direct sum of n right distributive ideals.

(1) If M is a prime ideal of A, then the factor ring A/M is isomorphic to the ring of $m \times m$-matrices over a division ring, where $m \leqslant n$.

In particular, the ideal M is maximal.

(2) The nilpotency index of any nilpotent element of every quotient ring of A does not exceed n.

3.18 [115]. Let the identity element of a ring A be the sum of orthogonal idempotents $e_1, \ldots, e_n$, and let any ring $e_i A e_i$ be right distributive.

Then the prime radical of A contains all right or left nil-ideals of A, and $F + G = A$ for any two incomparable nonsingularly prime ideals F, G of A.

3.19 [115]. Let the identity element of a ring A be a sum of nonzero orthogonal idempotents $e_1, \ldots, e_n$ such that the sum of any two incomparable prime ideals of any ring $e_i A e_i$ is equal to $e_i A e_i$. (This is the case if the ideal lattice of each ring $e_i A e_i$ is a chain.)

(1) $A = P + Q$ for any two incomparable prime ideals P and Q of A.

(2) If A is semiprime and A contains no infinite direct sums of nonzero ideals, then A is a finite direct product of prime rings.

(3) If A is a semiprime ring which contains no infinite direct sums of nonzero ideals, then

A is a piecewise domain with respect to $\{e_1, \ldots, e_n\}$ $\Leftrightarrow$ $e_i A e_i$ is a domain for any e_i.

(4) If F is a semiprime ideal of A such that the factor ring A/F contains no infinite direct sums of nonzero ideals, then

F is piecewise integral with respect to $\{e_1, \ldots, e_n\}$ $\Leftrightarrow$ $e_i F e_i$ is a completely prime ideal of $e_i A e_i$ for any $e_i \in A \setminus F$.

(5) Assume that F is a semiprime ideal of A, the quotient ring A/F contains no infinite direct sums of nonzero ideals, and all right ideals $(e_i A + F)/F$ of A/F are uniform. Then

F is nonsingularly semiprime $\Leftrightarrow$ F is piecewise integral with respect to $\{e_1, \ldots, e_n\}$ $\Leftrightarrow$ the ideal $e_i F e_i$ of $e_i A e_i$ is completely prime for any $e_i \in A \setminus F$.

3.20 [115]. Every semiprime right semidistributive ring which contains no infinite direct sums of nonzero ideals is right finite-dimensional.

3.21 [115]. Let A be a semiprime right semidistributive ring. Then
A is a ring with the maximum condition on right annihilators $\Leftrightarrow$
A is a ring with the maximum condition on left annihilators $\Leftrightarrow$
A is right nonsingular and does not contain infinite direct sums of nonzero ideals $\Leftrightarrow$
A is a right Goldie ring $\Leftrightarrow$
A is a finite direct product of prime right Goldie rings which are direct sums of uniform distributive right ideals.

3.22 [115]. Let A be a semiprime indecomposable right semidistributive ring with the maximum condition on right annihilators.
Then A is a prime right Goldie ring, A is a finite direct sum of distributive uniform right ideals $e_1 A, \ldots, e_m A$ such that $e_i^2 = e_i$, and all $e_i A_i e_i$ are right distributive domains.

3.23 [115]. Let A be a semilocal semiprime right semidistributive right nonsingular ring. Then A is a finite direct product of prime right semiuniform right Goldie rings.

3.24 [115]. Let A be a semiprime right hereditary right semidistributive ring such that either A contains no infinite direct sums of nonzero ideals or A is semilocal.
Then A is a finite direct product of prime right Noetherian rings.

3.25 [115]. Let A be a semiperfect right semidistributive ring, and let $1 = \sum_{i=1}^{n} e_i$ be a decomposition of the identity element of A into a sum of local orthogonal idempotents.
(1) The sum of any two incomparable prime ideals of A is equal to A.
(2) If A is semiprime, then A is a finite direct product of prime rings $A_1, \ldots, A_n$, and each A_i is a direct sum of distributive uniform right ideals.
(3) Let F be a semiprime ideal of A. Then
F is a nonsingularly semiprime ideal $\Leftrightarrow$
F is piecewise integral with respect to $\{e_1, \ldots, e_n\}$ $\Leftrightarrow$ the ideal $e_i F e_i$ of $e_i A e_i$ is completely prime for any $e_i \in A \setminus F$ $\Leftrightarrow$
A/F is a right Goldie ring.

3.26 [115]. Let A be a semiperfect right semidistributive ring, $1 = \sum_{i=1}^{n} e_i$ be a decomposition of the identity element of A into a sum of local orthogonal idempotents, and let N be the intersection of all nonsingularly semiprime ideals of A.
(1) N is the least nonsingularly semiprime ideal of A.
(2) N is the least semiprime ideal of A which is piecewise integral with respect to $\{e_1, \ldots, e_n\}$.
(3) N is the least ideal of A with the property that A/N is a semiprime right Goldie ring.

3.27 [115,117]. Let the identity element of a ring A be the sum of nonzero orthogonal idempotents $e_1, \ldots, e_n$ such that all right ideals $e_i A$ are uniform.
(1) $\mathrm{Sing}(A_A)$ coincides with the set of all elements $a \in A$ such that for any e_i, there is a nonzero element $b \in e_i A$ such that $ab = 0$.

(2) Let $i \in \{1, \ldots, n\}$, $a \in Ae_i$. Then
$a \in \mathrm{Sing}(A_A) \Leftrightarrow$
there is a nonzero element $b \in e_i A$ such that $ab = 0$.
(3) A is right nonsingular $\Leftrightarrow$
for any idempotent e_i, we have $ab \neq 0$ for each nonzero $a = ae_i \in Ae_i$ and for any nonzero $b = e_i b \in e_i A$.
(4) A is right nonsingular $\Leftrightarrow$
for all idempotents e_i, e_j, and e_k, we have $ab \neq 0$ for each nonzero $a = e_j a e_i \in e_j A e_i$ and for any nonzero $b = e_i b e_k \in e_i A e_k$.
(5) Assume that A is a ring with the maximum condition on right annihilators, $a \in Ae_i$, and there is a nonzero element $b \in e_i A$ such that $ab = 0$.
Then the ideal AaA of A is nilpotent.

3.28 [115,117]. Let the identity element of a ring A be the sum of nonzero orthogonal idempotents $e_1, \ldots, e_n$ such that all right ideals $e_i A$ and all left ideals Ae_i are uniform. Then the following conditions are equivalent.
(1) A is right nonsingular.
(2) A is left nonsingular.
(3) For all idempotents e_i, e_j, e_k, we have $ab \neq 0$ for each nonzero $a = e_j a e_i \in e_j A e_i$ and for any nonzero $b = e_i b e_k \in e_i A e_k$.

3.29 [117]. Let A be a Noetherian right semidistributive ring.
Then $\mathrm{Kdim}(A_A) \leqslant 1$ and $\bigcap_{i=1}^{\infty} (J(A))^i = 0$.

3.30 [114]. A is an arithmetical ring $\Leftrightarrow$
for any ideals $A_1, \ldots, A_n$ of A and for any $x_1, \ldots, x_n \in A$ such that $x_i - x_j \in A_i + A_j$ for all i, j, there is $x \in A$ such that $x - x_i \in A_i$ for $i = 1, \ldots, n \Leftrightarrow$
for any ideals A_1, A_2, and A_3 of A and for any $d \in (A_1 + A_2) \cap (A_1 + A_3)$, there is $x \in A_1$ such that $x - d \in A_2 \cap A_3$.

3.31 [117]. (1) If M is a semi-Noetherian or semi-Artinian module, then
M is semidistributive $\Leftrightarrow$
M is a direct sum of invariant modules, and M is arithmetical.
(2) If A is a semiperfect right semi-Noetherian ring, then
A is right semidistributive $\Leftrightarrow$
A is arithmetical and eA_A is an invariant module for each local idempotent e of A.
(3) All hereditary Noetherian semiprime rings are arithmetical.
(4) If A is a ring such that each idempotent endomorphism of any 2-generated right ideal of an arbitrary factor ring B of A can be extended to an endomorphism of B_B, then A is arithmetical.

3.32 [120]. Let M be a direct summand of a module $\bigoplus_{i=1}^{n} M_i$.
(1) If all modules $M_1, \ldots, M_n$ are nilpotently invariant, then the prime radical of $\mathrm{End}(M)$ contains all right or left nil-ideals of $\mathrm{End}(M)$.
(2) If all modules $M_1, \ldots, M_n$ are distributive, then the prime radical of $\mathrm{End}(M)$ contains all right or left nil-ideals of $\mathrm{End}(M)$.

3.33 [120]. Let M be a projective module which is a finite sum of its submodules $M_1, \ldots, M_n$, and let P be the prime radical of $\mathrm{End}(N)$.

(1) If all $M_1, \ldots, M_n$ are nilpotently invariant, then P contains all right or left nil-ideals of $\mathrm{End}(N)$.

(2) If all $M_1, \ldots, M_n$ are distributive, then P contains all right or left nil-ideals of $\mathrm{End}(N)$.

3.34 [120]. Let M be a finitely generated projective right module over a ring A, and let P be the prime radical of $\mathrm{End}(M)$.

(1) If M is a distributively generated module, then P contains all right nil-ideals and all left nil-ideals of $\mathrm{End}(M)$.

(2) If A is a right distributively generated ring, then P contains all right nil-ideals and all left nil-ideals of $\mathrm{End}(M)$.

References

[1] S. Barthwal, S.K. Jain, P. Kanwar and S.R. López-Permouth, *Nonsingular semiperfect CS-rings*, J. Algebra **203** (1998), 361–373.

[2] K.I. Beidar, V.N. Latyshev, V.T. Markov, A.V. Mikhalev, L.A. Skornyakov and A.A. Tuganbaev, *Associative rings*, Progress of Science and Technology: Series on Algebra, Topology, and Geometry, Vol. 22, Itogi Nauki i Tekhn. VINITI, All-Union Institute for Scientific and Technical Information, Akad. Nauk SSSR, Moscow (1985), 3–116.

[3] T. Belzner, *Towards self-duality of semidistributive artinian rings*, J. Algebra **135** (1) (1990), 74–95.

[4] K. Bessenrodt, H.H. Brungs and G. Törner, *Prime ideals in right chain rings*, Mitt. Math. Sem. Giessen **163** (1984), 141–167.

[5] K. Bessenrodt, H.H. Brungs and G. Törner, *Right chain rings*, Part 1, Schriftenreihe des Fachbereich Math. Universität Duisburg (1990).

[6] K. Bessenrodt, H.H. Brungs and G. Törner, *Right chain rings*, Part 2a, Schriftenreihe des Fachbereich Math. Universität Duisburg (1992).

[7] K. Bessenrodt, H.H. Brungs and G. Törner, *Right chain rings*, Part 2b, Schriftenreihe des Fachbereich Math. Universität Duisburg (1992).

[8] K. Bessenrodt and G. Törner, *Overrings of right chain domains*, Arch. Math. **48** (1987), 217–222.

[9] K. Bessenrodt and G. Törner, *On two-sided right ideals in chain domains*, Forum Math. **2** (1990), 307–322.

[10] H.H. Brungs, *Generalized discrete valuation rings*, Canad. J. Math. **21** (1969), 1404–1408.

[11] H.H. Brungs, *Noncommutative valuation rings*, Perspectives in Ring Theory, F. van Oystaeyen and L. Le Bruyn, eds, Kluwer Acad. Publishers, Dodrecht (1988), 105–115.

[12] H.H. Brungs and J. Gräter, *Valuation rings in finite-dimensional division algebras*, J. Algebra **120** (1989), 90–99.

[13] H.H. Brungs and J. Gräter, *Value groups and distributivity*, Canad. J. Math. **43** (6) (1991), 1150–1160.

[14] H.H. Brungs and J. Gräter, *Noncommutative Prüfer and valuation rings*, Proc. Int. Conf. Algebra Dedicat. Mem. A.I. Mal'cev, Novosibirsk (1989), Pt. 1, Providence, RI (1992), 253–269.

[15] H.H. Brungs and G. Törner, *Right chain domains and prescribed value holoid*, J. Algebra **176** (1995), 346–355.

[16] C.E. Caballero, *Self-duality and ℓ-hereditary semidistributive rings*, Comm. Algebra **14** (10) (1986), 1821–1843.

[17] R. Camps, A. Facchini and G.E. Puninski, *Serial rings that are endomorphism rings of Artinian modules*, Rings and Radicals, Pitman Series, Longman, Sci. & Tech. (1996).

[18] A.W. Chatters, *Serial rings with Krull dimension*, Glasgow Math. J. **32** (1990), 71–78.

[19] P. Cohn, *Free Rings and Their Relations*, Academic Press, London (1971).

[20] R.R. Colby and K.R. Fuller, *Modules over diserial rings*, Comm. Algebra **9** (5) (1981), 511–532.

[21] V.K. Deshpande, *A note on right serial rings*, Comm. Algebra **13** (8) (1985), 1737–1742.

[22] F. Dischinger and W. Mueller, *Einreihig zerlegbare artinsche Ringe sind selbstdual*, Arch. Math. **43** (1984), 132–1136.

[23] Yu.A. Drozd, *On generalized uniserial rings*, Mat. Zametki **18** (5) (1975), 707–710.
[24] N.I. Dubrovin, *Uniserial domains*, Vestn. MGU, Mat., Mekh., No. 2 (1980), 51–54.
[25] N.I. Dubrovin, *On uniserial rings*, Uspekhi Mat. Nauk **37** (4) (1982), 139–140.
[26] N.I. Dubrovin, *The rational closure of group rings of left-ordered groups*, Gerhard Mercator Universität Duisburg Gesamthochschule (1994).
[27] D. Eisenbud and P. Griffith, *Serial rings*, J. Algebra **17** (3) (1971), 268–287.
[28] D. Eisenbud and P. Griffith, *The structure of serial rings*, Pacif. J. Math. **36** (1971), 109–121.
[29] P. Eklof and I. Herzog, *Model theory of modules over a serial ring*, Ann. Pure Appl. Logic **72** (1995), 145–176.
[30] A. Facchini, *Krull–Schmidt fails for serial modules*, Trans. Amer. Math. Soc. **348** (11) (1996), 4561–4575.
[31] A. Facchini, *Module Theory: Endomorphism Rings and Direct Sum Decompositions in Some Classes of Modules*, Birkhäuser, Basel (1998).
[32] A. Facchini and G.E. Puninski, *Σ-pure-injective modules over serial rings*, Abelian Groups and Modules, A. Facchini and C. Menini, eds, Kluwer Acad. Publishers, Dordrecht (1995), 145–162.
[33] A. Facchini and G.E. Puninski, *Classical localizations in serial rings*, Comm. Algebra **24** (11) (1996), 3537–3559.
[34] A. Facchini and L. Salce, *Uniserial modules: Sums and isomorphisms of subquotients*, Comm. Algebra **18** (2) (1990), 499–517.
[35] L. Fuchs, *On polyserial modules over valuation domains*, Period. Math. Hungar. **18** (1987), 271–277.
[36] L. Fuchs and L. Salce, *Modules over valuation domains*, Lecture Notes in Pure and Appl. Math., Vol. 97 (1985).
[37] K.R. Fuller, *Generalized uniserial rings and their Kupisch series*, Math. Z. **106** (4) (1968), 248–260.
[38] K.R. Fuller, *Weakly symmetric rings of distributive module type*, Comm. Algebra **5** (9) (1977), 997–1008.
[39] K.R. Fuller, *Rings of left invariant module type*, Comm. Algebra **6** (1978), 153–167.
[40] K.R. Fuller, *On generalization of serial rings*, Proc. of the Filadelphia Conference on Rep. Thy., Lecture Notes in Pure and Appl. Math., Vol. 37, Dekker (1978), 359–368.
[41] K.R. Fuller, *Biserial rings*, Lecture Notes in Math., Vol. 734 (1979), 64–90.
[42] K.R. Fuller, *On generalization of serial rings II*, Comm. Algebra **8** (7) (1980), 635–661.
[43] K.R. Fuller, J. Haack and H. Hullinger, *Stable equivalence of uniserial rings*, Proc. Amer. Math. Soc. **68** (2) (1978), 153–158.
[44] A.W. Goldie, *Torsion-free modules and rings*, J. Algebra **1** (3) (1964), 268–287.
[45] K.R. Goodearl, *Artinian and Noetherian modules over regular rings*, Comm. Algebra **8** (5) (1980), 475–504.
[46] O.E. Gregul and V.V. Kirichenko, *On semihereditary serial rings*, Ukr. Mat. Zh. **39** (2) (1987), 156–161.
[47] J.K. Haack, *Self-duality and serial rings*, J. Algebra **59** (2) (1979), 345–363.
[48] J.K. Haack, *When do the transpose and dual agree?* Linear Algebra Appl. **56** (1984), 63–72.
[49] J.F. Habibi and W.H. Gustafson, *Reflexive serial algebras*, Linear Algebra Appl. **99** (1988), 217–223.
[50] D.A. Hill, *Rings whose indecomposable injective modules are uniserial*, Canad. J. Math. **34** (1982), 797–805.
[51] D.A. Hill, *Injective modules over non-Artinian serial rings*, J. Austral. Math. Soc. **44** (1988), 242–251.
[52] G. Ivanov, *Left generalized uniserial rings*, J. Algebra **31** (1974), 166–181.
[53] G. Ivanov, *Decomposition of modules over serial rings*, Comm. Algebra **3** (11) (1975), 1031–1036.
[54] V.V. Kirichenko, *Generalized uniserial rings*, Mat. Sb. **99** (4) (1976), 559–581.
[55] V.V. Kirichenko, *On serial hereditary and semihereditary rings*, Zap. Nauchn. Sem. LOMI Akad. Nauk SSSR **114** (1982), 137–147.
[56] V.V. Kirichenko and O.Ya. Bernik, *Semiperfect rings of distributive module type*, Dokl. Akad. Nauk USSR, -A, No. 3 (1988), 15–17.
[57] V.V. Kirichenko and P.P. Kostyukevich, *Biserial rings*, Ukr. Mat. Zh. **38** (6) (1986), 718–723.
[58] V.V. Kirichenko and Yu.V. Yaremenko, *Noetherian biserial rings*, Ukr. Mat. Zh. **40** (4) (1988), 435–440.
[59] I.B. Kozhukhov, *Uniserial semigroup rings*, Uspekhi Mat. Nauk **29** (1) (1974), 169–170.
[60] H. Krause, *Maps between tree and band modules*, J. Algebra **137** (1991), 186–194.
[61] H. Kupisch, *Beiträge zur Theorienichthalbeinfacher Ringen mit Minimalbedingung*, J. Reine Angew. Math. **201** (1959), 100–112.

[62] H. Kupisch, *Über eine Klasse von Ringen mit Minimalbedingung*, Arch. Math. **17** (1966), 20–35.
[63] H. Kupisch, *Einreihige Algebren über einem perfekten Körper*, J. Algebra **33** (1) (1975), 68–74.
[64] H. Kupisch, *A characterizaton of Nakayama rings*, Comm. Algebra **23** (2) (1995), 739–741.
[65] Lee Sin Min, *On the constructions of local and arithmetical rings*, Acta. Math. Acad. Sci. Hungar. **32** (1–2) (1978), 31–34.
[66] R. Makino, *QF-*1 *algebras with faithful direct sums of uniserial modules*, J. Algebra **136** (1991), 175–189.
[67] V.T. Markov, A.V. Mikhalev, L.A. Skornyakov and A.A. Tuganbaev, *Modules*, Progress of Science and Technology: Series on Algebra, Topology, and Geometry, Vol. 19, Itogi Nauki i Tekhn. VINITI, All-Union Institute for Scientific and Technical Information, Akad. Nauk SSSR, Moscow (1981), 31–134.
[68] V.T. Markov, A.V. Mikhalev, L.A. Skornyakov and A.A. Tuganbaev, *Endomorphism rings and submodule lattices*, Progress of Science and Technology: Series on Algebra, Topology, and Geometry, Vol. 21, Itogi Nauki i Tekhn. VINITI, All-Union Institute for Scientific and Technical Information, Akad. Nauk SSSR, Moscow (1983), 183–254.
[69] R. Mazurek and E.R. Puczylowski, *On semidistributive rings*, Comm. Algebra **25** (11) (1997), 3463–3471.
[70] I. Murase, *On the structure of generalized uniserial rings, I*, Sci. Papers College Gen. Ed. Univ. Tokyo **13** (1963), 1–22.
[71] I. Murase, *On the structure of generalized uniserial rings, II*, Sci. Papers College Gen. Ed. Univ. Tokyo **13** (1963), 131–158.
[72] I. Murase, *On the structure of generalized uniserial rings, III*, Sci. Papers College Gen. Ed. Univ. Tokyo **14** (1964), 11–25.
[73] I. Murase, *Generalized uniserial group rings, I*, Sci. Papers College Gen. Ed. Univ. Tokyo **15** (1965), 15–28.
[74] I. Murase, *Generalized uniserial group rings, II*, Sci. Papers College Gen. Ed. Univ. Tokyo **15** (1965), 111–128.
[75] B.J. Müller, *The structure of serial rings*, Methods in Module Theory, G. Abrams, J. Haefner and K.M. Rangaswamy, eds, Marcel Dekker, New York (1992), 249–270.
[76] B.J. Müller, *Goldie-prime serial rings*, Contemp. Math. **130** (1992), 301–310.
[77] B.J. Müller, *Ore sets in serial rings*, Comm. Algebra **25** (10) (1997), 3137–3146.
[78] B.J. Müller and S. Singh, *Uniform modules over serial rings II*, Lecture Notes in Math., Vol. 1448 (1989), 25–32.
[79] B.J. Müller and S. Singh, *Uniform modules over serial rings*, J. Algebra **144** (1991), 94–109.
[80] T. Nakayama, *Note on uniserial and generalized uniserial rings*, Proc. Imp. Acad. Tokyo **16** (1940), 285–289.
[81] T. Nakayama, *On Frobenius algebras. II*, Ann. Math. **42** (1) (1941), 1–21.
[82] Z. Pogorzaly *Representation-finite biserial algebras over a perfect field*, Bull. Pol. Acad. Sci. Math. **33** (7–8) (1985), 349–353.
[83] Z. Pogorzaly, *A characterization of representation-finite biserial algebras over a perfect field*, Rozpr. Mat. **277** (1989), 1–44.
[84] G.E. Puninskii, *Endodistributive and pure injective modules over uniserial rings*, Uspekhi Mat. Nauk **48** (3) (1993), 201–202.
[85] G.E. Puninskii, *A note on pure injective modules over serial rings*, Uspekhi Mat. Nauk **49** (5) (1994), 171–172.
[86] G.E. Puninskii, *Pure-injective modules over right Noetherian serial rings*, Comm. Algebra **23** (4) (1995), 1579–1592.
[87] G.E. Puninskii, *Serial Krull–Schmidt rings and pure-injective modules*, Fundam. Prikl. Mat. **1** (2) (1995), 471–490.
[88] G.E. Puninskii, *Indecomposable pure-injective modules over uniserial rings*, Trudy Moskov. Mat. Obshch. **56** (1995), 1–13.
[89] G.E. Puninskii, *Finitely presented modules over serial rings*, Fundam. Prikl. Mat. **3** (2) (1997), 631–633.
[90] G.E. Puninskii and A.A. Tuganbaev, *Rings and Modules*, Soyuz, Moscow (1998).
[91] G.E. Puninskii, R. Wisbauer and M. Yousif, *On p-injective rings*, Glasgow Math. J. **37** (1994), 373–378.
[92] M.I. Ramalho Martins, *Composition factors of indecomposable modules*, Trans. Amer. Math. Soc. **350** (5) (1998), 2009–2031.
[93] C.M. Ringel, *The repetitive algebra of a gentle algebra*, Bul. Soc. Mat. Mexicana **3** (1997), 235–253.

[94] T. Shores and W.J. Lewis, *Serial modules and endomorphism rings*, Duke Math. J. **41** (4) (1974), 889–909.
[95] S. Singh, *Serial right Noetherian rings*, Canad. J. Math. **36** (1) (1984), 22–37.
[96] S. Singh, *Artinian right serial rings*, Proc. Amer. Math. Soc. **125** (8) (1997), 2239–2240.
[97] L.A. Skornyakov, *When are all modules serial?* Mat. Zametki **5** (2) (1969), 173–182.
[98] A. Skowronski, *On nonexistence of oriented cycles in Auslander–Reiten quivers*, Acta Univ. Carolin. **25** (1) (1984), 45–52.
[99] A. Skowronski and J. Waschbüsch, *Representation-finite biserial algebras*, J. Reine Angew. Math. **345** (1983), 172–181.
[100] N. Snashall, *Reflexivity of left serial algebras*, Abelian Groups and Modules, A. Facchini and C. Menini, eds, Kluwer Acad. Publishers, Dordrecht (1995), 453–457.
[101] W. Stephenson, *Modules whose lattice of submodules is distributive*, Proc. London Math. Soc. **28** (2) (1974), 291–310.
[102] A.A. Tuganbaev, *Rings all modules over which are direct sums of distributive modules*, Vestnik Moskov. Univ., Mat. Mekh., No. 1 (1980), 61–64.
[103] A.A. Tuganbaev, *Rings with flat right ideals and distributive rings*, Mat. Zametki **38** (2) (1985), 218–228.
[104] A.A. Tuganbaev, *Distributive rings and endodistributive modules*, Ukr. Mat. Zh. **38** (1) (1986), 63–67.
[105] A.A. Tuganbaev, *Distributive rings and modules*, Trudy Moskov. Mat. Obshch. **51** (1988), 95–113.
[106] A.A. Tuganbaev, *Modules with distributive lattice of submodules*, Uspekhi Mat. Nauk **44** (1) (1989), 215–216.
[107] A.A. Tuganbaev, *Distributive rings and modules*, Mat. Zametki **47** (2) (1990), 115–123.
[108] A.A. Tuganbaev, *Embeddings into serial modules*, Abelian Groups and Modules, No. 9, Tomsk (1990), 115–120.
[109] A.A. Tuganbaev, *Serial rings and modules*, Mat. Zametki **48** (2) (1990), 99–106.
[110] A.A. Tuganbaev, *On quaternion algebras*, Trudy Sem. Petrovsk., No. 17, Moscow (1994), 209–219.
[111] A.A. Tuganbaev, *Endomorphism rings and distributivity*, Mat. Zametki **56** (4) (1994), 141–152.
[112] A.A. Tuganbaev, *Distributive semiprime rings*, Mat. Zametki **58** (5) (1995), 736–761.
[113] A.A. Tuganbaev, *Distributively decomposable rings*, Uspekhi Mat. Nauk **51** (3) (1996), 215–216.
[114] A.A. Tuganbaev, *Flat modules and rings that are finitely generated over its centre*, Mat. Zametki **60** (2) (1996), 254–277.
[115] A.A. Tuganbaev, *Direct sums of distributive modules*, Mat. Sb. **187** (12) (1996), 137–156.
[116] A.A. Tuganbaev, *Semidistributive hereditary rings*, Uspekhi Mat. Nauk **52** (4) (1997), 215–216.
[117] A.A. Tuganbaev, *Semidistributive Rings and Modules*, Kluwer Acad. Publishers, Dordrecht (1998).
[118] A.A. Tuganbaev, *Semidistributive and distributively decomposable rings*, Mat. Zametki **65** (2) (1999), 307–313.
[119] A.A. Tuganbaev, *Rings with many idempotents*, Preprint (1999).
[120] A.A. Tuganbaev, *Distributively generated rings and distributive modules*, Preprint (1999).
[121] A.A. Tuganbaev, *Modules over serial rings*, Mat. Zametki, to appear.
[122] D.A. Tuganbaev, *Uniserial Laurent series rings*, Fundam. Prikl. Mat. **3** (3) (1997), 947–951.
[123] D.A. Tuganbaev, *Uniserial Laurent skew series rings*, Vestnik Moscov. Univ., Mat. Mekh., to appear.
[124] D.A. Tuganbaev, *Semilocal distributive Laurent skew series rings*, Preprint (1999).
[125] M. Uematsu and K. Yamagata, *On serial quasi-hereditary rings*, Hokkaido Math. J. **19** (1990), 165–174.
[126] M.H. Upham, *Serial rings with right Krull dimension one*, J. Algebra **109** (1987), 319–333.
[127] N. Vanaja and V.N. Purav, *A note on generalized uniserial rings*, Comm. Algebra **21** (4) (1993), 1153–1159.
[128] R. Vila-Freyer and W. Crawley-Boevey, *The structure of biserial rings*, J. London Math. Soc. **57** (1998), 41–54.
[129] B. Wald and J. Waschbüsch, *Tame biserial algebras*, J. Algebra **95** (1985), 480–500.
[130] R.B. Warfield, *Decomposability of finitely presented modules*, Proc. Amer. Math. Soc. **25** (1970), 167–172.
[131] R.B. Warfield, *Serial rings and finitely presented modules*, J. Algebra **37** (3) (1975), 187–222.
[132] R.B. Warfield, *Bezout rings and serial rings*, Comm. Algebra **7** (1979), 533–545.
[133] J. Waschbüsch, *Self-duality of serial rings*, Comm. Algebra **14** (4) (1986), 581–589.
[134] R. Wisbauer, *Modules locally of serial type*, Period. Math. Hungar. **18** (1987), 39–52.
[135] M.H. Wright, *Serial rings with right Krull dimension one, II*, J. Algebra **117** (1988), 99–116.

[136] M.H. Wright, *Krull dimension in serial rings*, J. Algebra **124** (1989), 317–328.

[137] M.H. Wright, *Links between right ideals of serial rings with Krull dimension*, Lecture Notes in Math., Vol. 1448 (1989), 33–40.

[138] M.H. Wright, *Certain uniform modules over serial rings are uniserial*, Comm. Algebra **17** (1989), 441–469.

[139] M.H. Wright, *Uniform modules over serial rings with Krull dimension*, Comm. Algebra **18** (8) (1990), 2541–2557.

[140] M.H. Wright, *Prime serial rings with Krull dimension*, Comm. Algebra **18** (8) (1990), 2559–2572.

[141] M.H. Wright, *Right locally distributive rings*, Ring Theory. Proc. Bien. Ohio State–Denison Conf., Granville, Ohio, May, 1992, Singapore (1993), 350–357.

[142] W. Xue, *Two examples of local Artinian rings*, Proc. Amer. Math. Soc. **107** (1) (1989), 63–65.

[143] W. Xue, *A note on three Artinian rings*, Comm. Algebra **18** (7) (1990), 2243–2247.

[144] W. Xue, *Exact modules and serial rings*, J. Algebra **134** (1990), 209–221.

[145] W. Xue, *Rings with Morita duality*, Lecture Notes in Math., Vol. 1523 (1992), 1–197.

[146] W. Xue, *On a generalization of Azumaya's exact rings*, J. Algebra **172** (1995), 78–89.

[147] Y. Yukimoto, *Artinian rings whose projective indecomposables are distributive*, Osaka J. Math. **22** (1985), 339–344.

[148] B. Huisgen-Zimmermann, *The geometry of uniserial representations of finite dimensional algebras. III: Finite uniserial type*, Trans. Amer. Math. Soc. **348** (12) (1996), 4775–4812.

This Page Intentionally Left Blank

Modules with the Exchange Property and Exchange Rings

Askar A. Tuganbaev

Department of Mechanics and Mathematics, Moscow State University, Moscow, 119899 GSP Russia
E-mail: askar@tuganbaev.mccme.ru

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

Preface

All rings are assumed to be associative and (except for Section 5 and for some specially mentioned cases) unital. A ring A is an *exchange ring* if the following two equivalent conditions hold: (1) for any $a \in A$, there is $e = e^2 \in aA$ with $1 - e \in (1 - a)A$; (2) for any $a \in A$, there is $f = f^2 \in Aa$ with $1 - f \in A(1 - f)$. (The fact that (1)$\Leftrightarrow$(2) is proved in [113].)

Let $\aleph$ be a cardinal number. A module M is a module with the *$\aleph$-exchange property* if for every module X and for every direct decomposition $X = M' \oplus Y = \bigoplus_{i \in I} N_i$, where $M' \cong M$ and $\mathrm{card}(I) \leqslant \aleph$, there are submodules $N_i' \subseteq N_i$ $(i \in I)$ such that $X = M' \oplus (\bigoplus_{i \in I} N_i')$.

A module M is a module with the *finite exchange property* if M has the n-exchange property for every positive integer n. A module M has the finite exchange property if and only if the ring $\mathrm{End}(M)$ ring is an exchange ring [113].

A module M is a module with the *exchange property* if M has the $\aleph$-exchange property for every cardinal number $\aleph$. It is unknown whether every module with the finite exchange property has the exchange property. In [75], Matlis raised the following open question: If a module M is a direct sum of indecomposable injective modules, is every direct summand of M also a direct sum of indecomposable injective modules. The answer to the Matlis problem is affirmative whenever M has the finite exchange property [131]. (Also, see [36, 39,65,111,121,122].)

The systematic study of modules with the finite exchange property and exchange rings was initiated in [31,111–113,83]. Also, see [3,4,15,26,30,34,48–55,62,64,66,68,69,71,79, 80,83,84,88,89,91,98,104,109,110,116–118,121–129,131–133,135].

A ring A is *regular* or von Neumann regular if for every $a \in A$, there is $b \in A$ with $a = aba$. A ring A is *π-regular* if for every $a \in A$, there is $b \in A$ such that $a^n = a^n b a^n$ for some positive integer n. Every π-regular ring is an exchange ring [98]. Consequently, all regular rings are exchange rings. All semiregular rings[1] are exchange rings [113].

A ring A is *strongly π-regular* if for every $x \in A$, there is a positive integer m such that $x^m \in x^{m+1}A$. It is proved in [32] that a ring A is strongly π-regular $\Leftrightarrow$ for every $x \in A$, there is a positive integer n such that $x^n \in Ax^{n+1}$. Regular rings are well investigated. For example, the books [42] and [95] are devoted to regular rings. A ring A is *semiregular* if $A/J(A)$ is regular, and all idempotents of $A/J(A)$ can be lifted to idempotents of A. Every regular ring is a semiregular π-regular ring. In addition, all right or left Artinian rings are semiregular π-regular rings, and the endomorphism ring of any injective module is semiregular. Semiregular rings were considered in many sources. Here, we just mention [12,72,82,96,108,115], and [120].

Every strongly π-regular ring is π-regular [11]. If F is a division ring and M is a right vector F-space with an infinite basis $\{e_i\}_{i=1}^{\infty}$, then the ring $\mathrm{End}(M_F)$ is a regular (and π-regular) ring which is not strongly π-regular. The quotient ring of the ring of integers with respect to the ideal generated by the integer 4 is a strongly π-regular ring which is not regular. The papers [7,67,73,11], and [102] are among the first works concerned with π-regular and strongly π-regular rings. Also, see [2,9,13,14,22,27,29,32,43,56–60,74,76–78,85–87,97,105,107,108,119], and [130].

---

[1] For the definition of 'semiregular' see below.

A nonnecessarily unital ring A is an I_0-*ring* if the following equivalent conditions hold: (1) for every $a \in A \setminus J(A)$, there is a nonzero $x \in A$ with $x = xax$; (2) every right ideal of A which is not contained in $J(A)$ contains a nonzero idempotent; (3) every left ideal of A which is not contained in $J(A)$ contains a nonzero idempotent. ((1)$\Leftrightarrow$ (2) $\Leftrightarrow$ (3) is proved in [81].) A nonnecessarily unital ring A is a *Zorn* ring if A is an I_0-ring, and $J(A)$ is a nil-ideal. A nonnecessarily unital ring A is *potent* if A is an I_0-ring, and all idempotents of $A/J(A)$ can be lifted to idempotents of A.

All exchange rings are potent rings. Every π-regular ring is a Zorn ring, and all Zorn rings are potent. If A is the ring of all rational numbers with odd denominators, then A is an exchange ring which is not a Zorn ring (in particular, A is not π-regular).

Let B be a unitary subring of a ring A, let $\{A_i\}_{i=1}^{\infty}$ be a countable set of copies of A, let D be the direct product of the rings A_i, and let R be the subring in D generated by the ideal $\bigoplus_{i=1}^{\infty} A_i$ and by the subring $\{(b, b, b, \ldots) \mid b \in B\}$.

(1) If A is the field of rational numbers and B is the ring of all rational numbers with odd denominators, then R is a commutative reduced semiprimitive exchange ring which is not semiregular.

(2) If A is the field of rational numbers and B is the ring of integers, then R is a commutative reduced semiprimitive potent ring which is not an exchange ring, and the quotient ring $R/(\bigoplus_{i=1}^{\infty} A_i)$ is not potent.

(3) If A is the ring of all 2×2 matrices over a field F and B is the ring of upper triangular 2×2 matrices over F, then R is a strongly π-regular semiprimitive ring which has a non-semiprime ring and satisfies the polynomial identity $[[X_1, X_2]^2, X_3] = 0$ (where $[x, y]$ denotes the commutator $xy - yx$) (in particular, R is not semiregular).

(4) Let W be the first Weyl algebra over a field F of zero characteristic (i.e. W is the F-algebra with two generators x, y and one relation $xy - yx = 1$), let A be the ring of all 2×2 matrices over the ring W, and let B be the ring of all upper triangular 2×2 matrices over the ring W. Then R is semiprime, $x^2 = 0$ for every nilpotent element $x \in R(A, B)$, all prime ideals of R are maximal ideals, and R is not π-regular.

If φ be an injective endomorphism of a ring A, then the skew power series ring $A_\ell[[x, \varphi]]$ is an exchange ring (resp. a potent ring) $\Leftrightarrow$ A is an exchange ring (resp. a potent ring).

If A is the ring of all rational numbers with odd denominators, then A is a commutative local domain, $A[[x]]$ is a local domain, and the Laurent series ring $A((x))$ is not potent. In particular, A is an exchange ring, and $A((x))$ is not an exchange ring.

If $\{F_i\}_{i=1}^{\infty}$ is an infinite countable set of fields and A is the direct product of all fields F_i, then A is a commutative regular ring, and $A((x))$ is a commutative potent semiprimitive ring which is not semiregular.

1. Exchange rings

If A is a ring, then $J(A)$, $C(A)$, and $U(A)$ denote the Jacobson radical, the *centre* and the *group of invertible elements* of A, respectively. For a module M, the endomorphism ring and the lattice of all submodules are denoted by $\mathrm{End}(M)$ and $\mathrm{Lat}(M)$, respectively.

1.1. Every finitely generated module with the finite exchange property has the exchange property. (1.1 is directly verified.)

1.2 [31]. If $\aleph$ is a cardinal number, then every direct summand of a module with the $\aleph$-exchange property has the $\aleph$-exchange property, and every finite direct sum of modules with the $\aleph$-exchange property has the $\aleph$-exchange property.

1.3 [135]. Let M be a module, $S \equiv \mathrm{End}(M)$, and let $\aleph$ be a cardinal number. Then
M is a module with the $\aleph$-exchange property $\Leftrightarrow$
for every direct decomposition $M \oplus B = \bigoplus_{i\in I} X_i$ such that $\mathrm{card}(I) \leqslant \aleph$ and $X_i \cong M$ for all $i \in I$, there are submodules $C_i \subseteq X_i$ such that $M \oplus (\bigoplus_{i\in I} C_i) = \bigoplus_{i\in I} X_i$ $\Leftrightarrow$
for every summable set $\{f_i\}_{i\in I}$ of endomorphisms of M such that $\mathrm{card}(I) \leqslant \aleph$ and $\sum_{i\in I} f_i = 1_M$, there is a summable set $\{e_i\}_{i\in I}$ of orthogonal idempotents of S such that $\sum_{i\in I} e_i = 1_M$ and $e_i \in Sf_i$ for all $i \in I$.

1.4 [31,113]. Let M be a right module, and let $R \equiv \mathrm{End}(M)$. Then
M is a module with the finite exchange property $\Leftrightarrow$
M is a module with the 2-exchange property $\Leftrightarrow$
R is an exchange ring $\Leftrightarrow$
there are orthogonal idempotents $e_1, \dots, e_n$ of R such that $1_M = e_1 + \cdots + e_n$, and all rings $e_i \cdot R \cdot e_i$ are exchange rings $\Leftrightarrow$
for any $f_1, \dots, f_n \in R$ with $\sum_{i=1}^n f_i = 1_M$, there are orthogonal idempotents $e_i \in Rf_i$ with $\sum_{i=1}^n e_i = 1_M$ $\Leftrightarrow$
for any $f_1, \dots, f_n \in R$ with $\sum_{i=1}^n f_i = 1_M$, there are orthogonal idempotents $e_i \in f_iR$ with $\sum_{i=1}^n e_i = 1_M$.

1.5 [113,83]. A is an exchange ring $\Leftrightarrow$
A_A is a module with the finite exchange property $\Leftrightarrow$
${}_AA$ is a module with the finite exchange property $\Leftrightarrow$
A_A is a module with the exchange property $\Leftrightarrow$
${}_AA$ is a module with the exchange property $\Leftrightarrow$
there are orthogonal idempotents $e_1, \dots, e_n$ of A such that $1 = e_1 + \cdots + e_n$, and all e_iAe_i are exchange rings $\Leftrightarrow$
for all elements $f_1, \dots, f_n$ of A with $\sum_{i=1}^n f_i = 1$, there are orthogonal idempotents $e_i \in Af_i$ such that $\sum_{i=1}^n e_i = 1$ $\Leftrightarrow$
for all elements $f_1, \dots, f_n$ of A with $\sum_{i=1}^n f_i = 1$, there are orthogonal idempotents $e_i \in f_iA$ such that $\sum_{i=1}^n e_i = 1$.

1.6 [83]. A is an exchange ring $\Leftrightarrow$
for every $x \in A$, there is $e = e^2 \in A$ with $e - x \in (x - x^2)A$ $\Leftrightarrow$
all idempotents of A can be lifted modulo every right ideal $\Leftrightarrow$
$A/J(A)$ is an exchange ring, and all idempotents of $A/J(A)$ can be lifted to idempotents of A $\Leftrightarrow$
there is an ideal B of A such that $B \subseteq J(A)$, A/B is an exchange ring, and all idempotents of A/B can be lifted to idempotents of A $\Leftrightarrow$
for every $x \in A$, there is $e = e^2 \in xA$ such that $A = eA + (1 - x)A$ $\Leftrightarrow$
for every $e = e^2 \in A$, the ring eAe is an exchange ring.

1.7 [113]. Every semiregular ring is an exchange ring. Consequently, if M is a module, and $\mathrm{End}(M)$ is semiregular, then M has the finite exchange property.

1.8. (1) If A is a ring such that for every $x \in A$, there is a unitary subring S of A such that S is an exchange ring and $x \in S$, then A is an exchange ring.

(2) A is an exchange ring without nontrivial idempotents $\Leftrightarrow$

A is a potent ring without nontrivial idempotents $\Leftrightarrow$

A is an I_0-ring without nontrivial idempotents $\Leftrightarrow$

A is a local ring.

1.8 is directly verified.

1.9 [83]. Every central element of an exchange ring A is a sum of an invertible element and an idempotent.

1.10 [31,112]. M is an indecomposable module with the finite exchange property $\Leftrightarrow$

M is an indecomposable module with the exchange property $\Leftrightarrow$

$\mathrm{End}(M)$ is an exchange ring without nontrivial idempotents $\Leftrightarrow$

$\mathrm{End}(M)$ is a local ring.

1.11 [135]. Let $\{M_i\}_{i \in I}$ be a set of indecomposable modules. Then

$\bigoplus_{i \in I} M_i$ has the finite exchange property $\Leftrightarrow$

$\bigoplus_{i \in I} M_i$ has the exchange property $\Leftrightarrow$

all M_i have local endomorphism rings, and the set $\{M_i\}_{i \in I}$ is semi-T-nilpotent.[2]

A ring A is *orthogonally finite* if A does not contain infinite sets of nonzero orthogonal idempotents.

1.12 [83,26]. The endomorphism ring of a module M is semiperfect $\Leftrightarrow$

$\mathrm{End}(M)$ is an orthogonally finite I_0-ring $\Leftrightarrow$

$\mathrm{End}(M)$ is an exchange ring, and the identity element of $\mathrm{End}(M)$ is a finite sum of primitive orthogonal idempotents $\Leftrightarrow$

there are orthogonal idempotents $e_1, \dots, e_n$ of $\mathrm{End}(M)$ such that $1_M = e_1 + \cdots + e_n$, and all $e_i \mathrm{End}(M) e_i$ are local rings $\Leftrightarrow$

$M = \bigoplus_{i=1}^n M_i$, where all $\mathrm{End}(M_i)$ are local rings $\Leftrightarrow$

M is a finite direct sum of indecomposable modules with the finite exchange property $\Leftrightarrow$

M has the finite exchange property and is a finite direct sum of indecomposable modules $\Leftrightarrow$

M has the exchange property and is a finite direct sum of indecomposable modules.

1.13 [135]. M is a module with the exchange property $\Leftrightarrow$

for every direct decomposition $M \oplus G = \bigoplus_{i \in I} F_i$ such that $F_i \cong M$ for all $i \in I$, there are submodules $H_i \subseteq F_i$ such that $M \oplus (\bigoplus_{i \in I} H_i) = \bigoplus_{i \in I} F_i$ $\Leftrightarrow$

M is a module with the finite exchange property, and for every direct decomposition $M \oplus G = F = \bigoplus_{i \in I} F_i$ with $F_i \cong M$ for all $i \in I$, there is a submodule H of F such that H is a maximal element of the set of all submodules of F that have the following properties: $H = \bigoplus_{i \in I} H_i$, $H_i \subseteq F_i$, $H \cap M = 0$, and the natural monomorphism $M \to F/H$ is split.

1.14 [80]. M has the finite exchange property $\Leftrightarrow$

for every $f \in \mathrm{End}(M)$, there are $g, h \in \mathrm{End}(M)$ such that $gfg = g$ and $h(1-f)(1-gf) = 1 - gf$.

---

[2] For the definition of semi-T-nilpotent see just above 1.18 below.

1.15 [125]. Let the identity element of a ring A be a sum of primitive orthogonal idempotents. Then

all right A-modules and all left A-modules have the finite exchange property $\Leftrightarrow$

A is Artinian, and there are only finitely many non-isomorphic finitely generated right (left) A-modules.

1.16 [131]. Let M be a direct sum of modules with local endomorphism rings. Then

$\mathrm{End}(M)$ is an exchange ring $\Leftrightarrow$

every ideal of $\mathrm{End}(M)$ which does not belong to $J(\mathrm{End}(M))$ contains a nonzero idempotent.

1.17 [2]. Let a be an element of an exchange ring A such that there exist a direct decomposition $A_A = aA \oplus E$ and a positive integer n such that $A_A = a^n A \oplus r(a^n)$ and ${}_A A = Aa^n \oplus \ell(a^n)$.

Then $a = hv$, where $h = h^2$ and $v \in U(A)$.

A set $\{M_i\}_{i \in I}$ of modules M_i is *semi-T-nilpotent* if for each sequence $M_{i_1} \xrightarrow{f_1} M_{i_2} \xrightarrow{f_2} M_{i_3} \xrightarrow{f_3} \cdots$ of nonisomorphisms, where the subscripts $i_n \in I$ are pairwise different, and for each $x \in M_{i_1}$, there exists a positive integer n such that $f_n f_{n_1} \cdots f_2 f_1(x) = 0$.

1.18 [135]. Let $\{M_i\}_{i \in I}$ be a semi-T-nilpotent set of modules M_i with the exchange property.

(1) If M_i and M_j have no nontrivial isomorphic direct summands for $i \neq j$, then $\bigoplus_{i \in I} M_i$ has the exchange property.

(2) If all modules M_i are isomorphic, then $\bigoplus_{i \in I} M_i$ has the exchange property.

1.19 [135]. Let $\{M_i\}_{i \in I}$ be a set of modules, and let the lengths of all M_i be uniformly bounded by some integer. Then $\bigoplus_{i \in I} M_i$ has the finite exchange property.

1.20 [126]. Let M be a module with the finite exchange property and $\mathrm{End}(M)$ is a normal ring. Then M has the $\aleph_0$-exchange property. In addition, if M is projective, then M has the exchange property.

2. Homological properties

A module M is *projective with respect to* a module N (or N*-projective*) if for every epimorphism $h : N \to \overline{N}$ and for every $\overline{f} \in \mathrm{Hom}(M, \overline{N})$, there is $f \in \mathrm{Hom}(M, N)$ with $\overline{f} = hf$. A module which is projective with respect to itself is a *quasi-projective* module.

2.1 [109]. Let M be a quasi-projective module. Then

M has the finite exchange property $\Leftrightarrow$

for every decomposition $M = N_1 + \cdots + N_n$ into a sum of submodules N_i of M, there is a direct decomposition $M = M_1 \oplus \cdots \oplus M_n$ such that $M_i \subseteq N_i$ for all i $\Leftrightarrow$

for every decomposition $M = L + N$ into a sum of submodules L and N of M, there is a direct summand M_1 of M such that $M_1 \subseteq L$ and $M = M_1 + N$.

2.2 [83]. If M is a projective module such that every homomorphic image of M has a projective cover, then M has the finite exchange property.

2.3 [113]. Every projective module over an exchange ring A is isomorphic to a direct sum of cyclic direct summands of A_A.

A module M is *regular* if every finitely generated submodule of M is a direct summand of M. A ring A is regular $\Leftrightarrow$ A_A is a regular module $\Leftrightarrow$ ${}_AA$ is a regular module (see [42, Theorem 1.1]). A ring A is a *ring with the right P-exchange property* if every projective right A-module has the exchange property. A ring A is a *ring of index at most n* if there is a positive integer n such that $a^n = 0$ for every nilpotent element $a \in A$. A ring A is a ring *of bounded index* if A is a ring of index at most n for some positive integer n.

2.4 [98]. (1) Every regular projective module has the exchange property.

(2) Every projective module over a regular ring has the exchange property.

(3) If M is a regular projective module, then End(M) is an exchange ring.

(4) Let M be a projective module. Then $M^{(J)}$ has the exchange property for every set J $\Leftrightarrow$

for every set J and for each set $\{N_i\}_{i\in I}$ of submodules of $M^{(J)}$ such that $M^{(J)} = \sum_{i\in I} N_i$, there are submodules N_i' of N_i such that $M^{(J)} = \bigoplus_{i\in I} N_i'$.

(5) Every quotient ring of any ring with the right P-exchange property is a ring with the right P-exchange property.

(6) A is has the right P-exchange property $\Leftrightarrow$

$A/J(A)$ is a ring with the right P-exchange property, and $J(A)$ is right T-nilpotent.

(7) If $A/J(A)$ is a regular ring and $J(A)$ is right T-nilpotent, then A is has the right P-exchange property.

(8) Let M_A be a finitely generated right module. Then

$M^{(I)}$ has the exchange property for every set I $\Leftrightarrow$

End(M) has the right P-exchange property.

(9) If A has the right P-exchange property, then the endomorphism ring of every finitely generated right A-module has the right P-exchange property.

(10) Let the identity element of a ring A be a sum of nonzero orthogonal idempotents $e_1, \dots, e_n$. Then

A has the right P-exchange property $\Leftrightarrow$

all e_iAe_i have the right P-exchange property.

(11) Let A be a normal ring. Then

A has the right P-exchange property $\Leftrightarrow$

the direct sum of a countable set of copies of A_A has the finite exchange property $\Leftrightarrow$

$A/J(A)$ is a regular ring, and $J(A)$ is right T-nilpotent.

(12) Let A be a ring of index at most n such that all projective right modules have the finite exchange property. Then A is strongly π-regular, and every prime quotient ring of A is isomorphic to a matrix ring D_k for some division ring D and some $k \leqslant n$.

(13) Let A be a P.I. ring such that all projective right modules have the finite exchange property. Then A is strongly π-regular.

2.5 [124,54]. Let the identity element of a ring A be a sum of primitive orthogonal idempotents. Then

every projective right A-module has the exchange property $\Leftrightarrow$

A is a ring with the minimum condition on principal left ideals.

A module M is *quasi-continuous* if for any two submodules N_1, N_2 of M with $N_1 \cap N_2 = 0$, there is a direct decomposition $M = M_1 \oplus M_2$ such that $M_1 \supseteq N_1$ and $M_2 \supseteq N_2$. A module M is *continuous* if M is quasi-continuous, and every submodule of M which is isomorphic to a direct summand of M is a direct summand of M.

2.6 [89]. Every quasi-continuous module with the finite exchange property has the exchange property.

2.7 [79]. Every continuous module M has the exchange property, and $\mathrm{End}(M)$ is a semiregular ring.

2.8 [135]. Let X be a pure-injective module, and let M be a submodule of X such that $f(M) \subseteq M$ for each $f \in \mathrm{Hom}(M, X)$. Then M has the exchange property.

2.9 [134]. Every pure-injective module M has the exchange property, and $\mathrm{End}(M)$ is a semiregular ring.

2.10 [123]. (1) Let M be a direct sum of indecomposable injective modules. Then
M has the exchange property $\Leftrightarrow$
M has the finite exchange property $\Leftrightarrow$
$J(\mathrm{End}(M))$ coincides with the set $\{f \in \mathrm{End}(M) \mid \mathrm{Ker}(f) \text{ is essential in } M\}$.

(2) Every direct sum of indecomposable injective right modules over a ring A has the exchange property $\Leftrightarrow$

every direct sum of indecomposable injective right A-modules has the finite exchange property $\Leftrightarrow$

A is a ring with the maximum condition on right ideals B such that A/B is a uniform right A-module.

(3) There is a module M such that M is a direct sum of indecomposable injective modules, M has the exchange property, and M is not quasi-injective.

A ring A has *stable range* 1 if for any $f, v, x, y \in A$ with $fx + vy = 1$, there is $h \in A$ with $f + vh \in U(A)$. For any ring A, denote by A^n and ${}^nA$ the set of all rows of length n and the set of all columns of height n, respectively. A row $b \in A^n$ is *unimodular* if $b({}^nA) = A$ as a right A-module. A ring is said to *have stable range at most n* if for any unimodular row $b = (b_1, b') \in A \times A^n$, there is $x \in A^n$ such that $b_1x + b'$ is a unimodular row in A^n. A ring *has stable range* ∞ if for every positive integer n, A is not a ring of stable range at most n.

2.11 [129]. If A is an exchange ring such that every right primitive quotient ring of A is Artinian, then A has stable range 1.

2.12 [129]. Let A be an exchange ring A. Then
A has stable range 1 $\Leftrightarrow$
for every $a \in A$ such that aA is a direct summand of A_A, there are $g = g^2 \in A$ and $u \in U(A)$ such that $a = gu$ $\Leftrightarrow$

for any $a, x \in A$ and for every idempotent $e \in A$ with $ax + e = 1$, there is $y \in A$ such that $a + ey \in U(A)$.

2.13 [30]. Let A be an exchange ring A. Then
A has stable range 1 $\Leftrightarrow$

A is directly finite, and for any idempotents $e, f \in A$ with $e = 1 + ab$ and $f = 1 + ba$ for some $a, b \in A$, there is a central idempotent $u \in A$ such that ueA is isomorphic to a direct summand of ufA and $(1-u)fA$ is isomorphic to a direct summand of $(1-u)eA$.

2.14 [84]. If A is an exchange ring with stable range 1, then for any $a \in A$ there are idempotents $e, f \in A$ such that $e \in aA$, $1 - e \in (1-a)A$, $f \in Aa$, $1 - f \in A(1-a)$, and $f = u^{-1}eu$ for some $u \in U(A)$.

2.15 [118]. Let A be an exchange ring. Then

A has stable range at most n $\Leftrightarrow$

for any idempotent $e \in A$ and any $b \in A^n$, the equality $eA + b({}^nA) = A$ implies that there is $y \in A^n$ such that $ey + b$ is a unimodular row in A^n $\Leftrightarrow$

$A^n \oplus X \cong A \oplus Y$ implies that X is isomorphic to a direct summand of Y $\Leftrightarrow$

for any idempotent $e \in A$ and any A-module decomposition $A^n = X \oplus Y$ such that $X \cong eA$, the A-module $(1-e)A$ is isomorphic to a direct summand of Y.

2.16 [118]. Let M be a module with the finite exchange property, and let $S \equiv \mathrm{End}(M)$. Then

S has stable range at most n $\Leftrightarrow$

for any two modules X and Y, $M^n \oplus X \cong M \oplus Y$ implies that X is isomorphic to a direct summand of Y $\Leftrightarrow$

for any two decompositions $M = X_1 \oplus X_2$ and $M^n = Y_1 \oplus Y_2$, where $X_1 \cong Y_1$, the module X_2 is isomorphic to a direct summand of Y_2.

Let A be a ring, s be a positive integer, and let $V(A)$ be the monoid of isomorphism classes of finitely generated projective right A-modules, with the operation induced by the direct sum. We say that A *satifies s-comparability* if for any $p, q \in V(M)$, either p is a divisor of q^s (in $V(M)$) or q is a divisor of p^s.

2.17 [91]. Every exchange ring satifying s-separability has stable range 1, 2 or ∞.

2.18. A description of the closure of the natural affine continuous function representation of $K_0(A)$ for any exchange ring A is given in [92].

3. Clean rings and normal exchange rings

A ring A is a *clean* ring if every element of A is a sum of an invertible element and an idempotent. A ring is *normal* if all its idempotents are central. A ring A is *right quasi-invariant* (*resp. right invariant*), if all maximal right ideals (resp. all right ideals) of A are ideals. A ring A is *unit-regular* if for every $a \in A$, there is $u \in U(A)$ with $a = aua$.

3.1 [83]. (1) Every clean ring A is an exchange ring.

(2) Every normal exchange ring A is a clean ring.

(3) Every unit-regular ring A is a clean ring.

3.2 [83]. A is a clean ring $\Leftrightarrow$

$A/J(A)$ is a clean ring, and all idempotents of $A/J(A)$ can be lifted to idempotents of A.

3.3 [83,26]. (1) If $A/J(A)$ is a unit-regular ring and all idempotents of $A/J(A)$ can be lifted to idempotents of A, then A is a clean ring.

(2) Every semiperfect ring is a clean ring.

3.4 [26]. (1) If e is a nontrivial idempotent of a ring A and eAe, $(1-e)A(1-e)$ are clean rings, then $eAe + (1-e)A(1-e) + eA(1-e)$ is a clean ring.

(2) If A is a clean ring, then the ring $UT_n(A)$ of upper triangular $n \times n$ matrices over A is a clean ring.

3.5 [26]. Let A be a ring with $2^{-1} \in A$. Then
A is a clean ring $\Leftrightarrow$
for every $a \in A$, there are $v, w \in U(A)$ such that $a = v + w$ and $w^2 = 1$.

A module M is *distributive* if $F \cap (G + H) = F \cap G + F \cap H$ for any three submodules F, G, and H of M. A module M is a *Bezout module* if every finitely generated submodule of the module M is a cyclic module. A module M is *uniserial* if any two submodules of M are comparable with respect to inclusion. Every right or left uniserial ring is local. A module is *distributively generated* if it is generated by distributive submodules. A direct sum of distributive modules is a *semidistributive* module.

3.6 [110]. (1) Every normal exchange ring A is quasi-invariant.

(2) Let M be a module over a ring A, and let $A/J(A)$ be a normal exchange ring. Then
M is a distributive module $\Leftrightarrow$ M is a Bezout module.

3.7 [110]. (1) If M is a quasi-projective module with the finite exchange property and M is a finite sum of distributive submodules, then M is a finite direct sum of distributive modules.

(2) If M is a quasi-projective finitely generated distributively generated module and End(M) is an exchange ring, then M is semidistributive.

(3) Every right distributively generated exchange ring is right semidistributive.

Let A be a ring, and let $S(A)$ be the non-empty set of all proper ideals of A generated by central idempotents. An ideal $P \in S(A)$ is a *Pierce ideal* of A if P is a maximal element of the set $S(A)$ (with respect to inclusion). If P is a Pierce ideal of A, then the quotient ring A/P is called a *Pierce stalk* of A.

3.8 [110]. A is an exchange ring $\Leftrightarrow$
all Pierce stalks of A are exchange rings $\Leftrightarrow$
all indecomposable quotient rings of A are exchange rings.

3.9 [25]. A is a normal exchange ring $\Leftrightarrow$
every element of A is a sum of an invertible element and a central idempotent $\Leftrightarrow$
all Pierce stalks of A are local rings.

3.10 [104]. A is a right distributive exchange ring $\Leftrightarrow$
A is a normal right Bezout exchange ring $\Leftrightarrow$
all Pierce stalks of A are right uniserial rings.

3.11 [110]. Let A be a normal exchange ring, let M be a maximal right or left ideal, and let $T \equiv A \setminus M$.

Then there exist a local ring Q and a surjective ring homomorphism $f : A \to Q$ such that $f(T) \subseteq U(Q)$, $Q = \{f(a)f(t)^{-1} \mid a \in A,\ t \in T\} = \{f(t)^{-1}f(a) \mid a \in A,\ t \in$

$T\}$, $\mathrm{Ker}(f) = \{a \in A \mid \exists t \in T\colon\ at = 0\} = \{a \in A \mid \exists t \in T\colon\ ta = 0\}$, and $J(Q) = \{f(m)f(t)^{-1} \mid m \in M,\ t \in T\} = \{f(t)^{-1}f(m) \mid m \in M,\ t \in T\}$.

A ring without nonzero nilpotent elements is called a *reduced* ring. A ring A is a *pf-ring* if for any $a, b \in A$ with $ab = 0$, there are $x, y \in A$ such that $x + y = 1$, $ax = 0$, and $yb = 0$.

3.12 [110]. A is a normal exchange pf-ring $\Leftrightarrow$
A is a reduced exchange pf-ring $\Leftrightarrow$
all Pierce stalks of A are local domains.

3.13 [127]. Let A be a right or left quasi-invariant ring. Then
A is an exchange ring $\Leftrightarrow$ A is a clean ring.

4. Potent and Zorn rings

4.1. (1) A is a Zorn ring $\Leftrightarrow$
A is an I_0-ring, and $J(A)$ is a nil-ideal $\Leftrightarrow$
A is a potent ring, and $J(A)$ is a nil-ideal.

(2) Every π-regular ring A is a Zorn ring. In particular, A is a potent ring, and $J(A)$ is a nil-ideal.

(4.1 is directly verified.)

4.2 [83]. Every exchange ring is potent.

4.3 [81]. (1) A is a potent ring $\Leftrightarrow$
$A/J(A)$ is a potent ring, and all idempotents of $A/J(A)$ can be lifted to idempotents of A.

(2) A is a potent semiprimitive ring $\Leftrightarrow$
A is a semiprimitive I_0-ring $\Leftrightarrow$
every nonzero right ideal of A contains a nonzero idempotent $\Leftrightarrow$
every nonzero left ideal of A contains a nonzero idempotent $\Leftrightarrow$
for every nonzero $a \in A$, there is a nonzero $x \in A$ with $x = xax$.

4.4 [81,26]. A is a semiperfect ring $\Leftrightarrow$
A is an orthogonally finite potent ring $\Leftrightarrow$
A is an orthogonally finite I_0-ring $\Leftrightarrow$
for every right ideal L of A, there are $e = e^2 \in L$ and a right ideal M of A such that $M \subseteq J(A)$ and $L = eA + M$ $\Leftrightarrow$
A is a semilocal I_0-ring $\Leftrightarrow$
A is an I_0-ring which is isomorphic to a subring of a semiperfect ring $\Leftrightarrow$
A is a homomorphic image of a semiperfect ring $\Leftrightarrow$
the matrix ring A_n is semiperfect for every positive integer n $\Leftrightarrow$
the matrix ring A_n is semiperfect for some positive integer n $\Leftrightarrow$
A is an exchange ring, and the identity element of A is a sum of primitive orthogonal idempotents $\Leftrightarrow$
A is semiregular, and the identity element of A is a sum of primitive orthogonal idempotents $\Leftrightarrow$

all quotient rings of A are orthogonally finite semiperfect exchange rings.

A set $\{e_{ij}\}_{i,j=1}^{n}$ of nonzero elements of a ring A is said to be a *system of n^2 matrix units* if $e_{ij}e_{st} = \delta_{js}e_{it}$, where $\delta_{jj} = 1$ and $\delta_{js} = 0$ for $j \neq s$. In this sitiuation, set $e \equiv \sum_{i=1}^{n} e_{ii}$. Then e is an idempotent of A, the set $\{e_{ij}\}_{i,j=1}^{n}$ is contained in the subring eAe of A, and the ring eAe can be identified with the matrix ring B_n, where B is the subring of eAe formed by all the elements which commute with all the elements e_{ij}. An ideal B of a ring A is a *matrix ideal* if B is a unital ring which is isomorphic to a total matrix ring over a reduced Zorn ring. If N is the maximal nil-ideal of a ring A and each nonzero ideal of A/N contains a matrix ideal, then A is called a *weakly reducible* ring.

4.5 [73]. Let A be a Zorn ring, n be a positive integer, and let a be an element of A such that $a^n = 0$ and $a^{n-1} \notin J(A)$. Then the ideal AaA of A contains a system of n^2 matrix units.

4.6 [73]. Let B be a nonzero ideal of a Zorn semiprimitive ring A. Then either B contains a matrix ideal of A, or B contains an infinite sequence of matrix unit systems $\{e_{ik}^{(j)}\}$ $(i, k = 1, \dots, j;\ j = 1, 2, 3, \dots)$ such that $e_{11}^{(j+1)} \in e_{11}^{(j)} A e_{11}^{(j)}$.

4.7 [73]. Let A be a weakly reducible semiprimitive Zorn ring. Then

A is a subdirect product of matrix rings over reduced Zorn rings.

In addition, A is a primitive ring $\Leftrightarrow$ A is isomorphic to a matrix ring over a division ring.

4.8 [73]. Let A be a semiprimitive Zorn ring. If either A is a ring of bounded index or all primitive quotient rings of A are rings of bounded index, then A is weakly reducible.

4.9 [63]. Let A be a semiprimitive Zorn ring.

(1) If e is a nonzero idempotent of A, then the ring eAe is a potent semiprimitive ring.

(2) A is a nonsingular ring.

(3) Let A have a system $\{e_{ij}\}_{i,j=1}^{n}$ of n^2 matrix units and supper that either $e_{11}Ae_{11}$ is not a reduced ring, or the idempotent $\sum_{i=1}^{n} e_{ii}$ is not central in A. Then A has a system $\{f_{st}\}_{s,t=1}^{m}$ of matrix units such that $m > n$ and $f_{11} \in e_{11}Ae_{11}$.

(4) If every right primitive quotient ring of A is a ring of bounded index, then any nonzero ideal B of A contains a nonzero ideal C of A such that $C = D_n$, where D is a reduced ring, and the identity element u of D is a nonzero central idempotent of A.

4.10. If A is an indecomposable semiprimitive Zorn ring and every right primitive quotient ring of A is a ring of bounded index, then A is a simple ring which is isomorphic to the matrix ring over a simple domain D.

(4.9 follows from 4.8(4).)

4.11 [76]. Let A be an indecomposable semiprimitive Zorn ring such that every right primitive quotient ring of A is Artinian. Then A is a simple Artinian ring.

4.12 [110]. (1) If A is a right distributively generated Zorn ring, then $J(A)$ coincides with the prime radical of A.

(2) If A is a semiprime right distributively generated Zorn ring, then A is semiprimitive.

(3) Let n be a positive integer, and let A be a Zorn prime ring which is a sum of n distributive right ideals.

Then A is a simple Artinian ring of index at most n which is isomorphic to a ring of all $k \times k$ matrices over a division ring with $k \leqslant n$.

4.13 [81]. Let A be an I_0-ring. Then
every nonzero idempotent of A contains a primitive idempotent, and any two primitive idempotents of A are equivalent $\Leftrightarrow$
A has a local module, and $J(A)$ is a prime ideal.

4.14 [25]. A is a normal ring, and all quotient rings of A are Zorn rings $\Leftrightarrow$
A is a normal strongly π-regular ring $\Leftrightarrow$
all Pierce stalks of A are local strongly π-regular rings.
A commutative ring is *zero-dimensional* if all prime ideals of the ring are maximal. Every commutative zero-dimensional ring is a semiregular Zorn ring.

4.15 [45]. If A is a module-finite algebra over a commutative zero-dimensional ring, then A is semiregular, and $J(A)$ is a nil-ideal. In particular, A is a Zorn ring.

4.16 [45]. Let A be an algebra over a commutative zero-dimensional ring R such that for every maximal ideal M of R, the ring of quotients A_M is a finitely generated R_M-module. Then A is an exchange ring.

5. Strongly π-regular rings

5.1 [9]. The endomorphism ring of a module M is strongly π-regular $\Leftrightarrow$
for every $f \in \mathrm{End}(M)$, there is a positive integer n such that $M = \mathrm{Ker}(f^n) \oplus f^n(M)$.
In this case, every injective or surjective endomorphism of M is an automorphism.

5.2 [37,57,9]. A is a strongly π-regular ring $\Leftrightarrow$
the quotient ring A/N with respect to the prime radical N of A is strongly π-regular $\Leftrightarrow$
every prime quotient ring of A is strongly π-regular $\Leftrightarrow$
each injective endomorphism of every cyclic right A-module is an automorphism $\Leftrightarrow$
for every $a \in A$, there is a positive integer n such that $A_A = r(a^n) \oplus a^n A$ $\Leftrightarrow$
for every $a \in A$, there is a positive integer n such that ${}_A A = Aa^n \oplus \ell(a^n)$ $\Leftrightarrow$
for every $a \in A$, there is $x \in A$ such that $ax = xa$ and $a^n = a^{n+1}x = xa^{n+1}$ $\Leftrightarrow$
for every $a \in A$, there is a positive integer n such that system of equations $\{ay = ya,\ a^n = a^{2n}y,\ y^2a^n = y\}$ has exactly one solution.

5.3 [9]. All matrix rings A_n over a ring A are strongly π-regular rings $\Leftrightarrow$
for every positive integer n, all injective endomorphisms of every cyclic right A_n-module are automorphisms $\Leftrightarrow$
all injective endomorphisms of every finitely generated right A-module are automorphisms $\Leftrightarrow$
all injective endomorphisms of every finitely generated left A-module are automorphisms.

5.4 [102]. A is a π-regular reduced ring $\Leftrightarrow$
A is a strongly π-regular reduced ring $\Leftrightarrow$
A is an Abelian regular ring $\Leftrightarrow$.

5.5 [2]. Every strongly π-regular ring has stable range 1.

5.6 [22]. Let R be a ring, $\{A_i\}_{i\in I}$ be a set of strongly π-regular unitary subrings of R, and let $A \equiv \bigcap_{i\in I} A_i$.

Then A is strongly π-regular.

5.7 [11]. Let A be a strongly π-regular ring, and let n be a positive integer. Then
$r(a^n) = r(a^{n+1})$ for every $a \in A \Leftrightarrow$
$\ell(a^n) = \ell(a^{n+1})$ for every $a \in A \Leftrightarrow$
$a^n A = a^{n+1} A$ for every $a \in A \Leftrightarrow$
$Aa^n = Aa^{n+1}$ for every $a \in A \Leftrightarrow$
A is a ring of index at most n.

5.8 [59]. Let A be a ring, N be the prime radical of A, n be a positive integer, and let A/N be a ring of index at most n. Then
A is π-regular $\Leftrightarrow$
A is strongly π-regular $\Leftrightarrow$
every prime quotient ring of A is a π-regular ring of bounded index $\Leftrightarrow$
every prime quotient ring of A is a simple Artinian ring of index at most n $\Leftrightarrow$
$a^n A + N = a^{n+1} A + N$ for every $a \in A \Leftrightarrow$
$Aa^n + N = Aa^{n+1} + N$ for every $a \in A$.

5.9 [11,59]. Let A be a ring of index at most n. Then
A is π-regular $\Leftrightarrow$
A is strongly π-regular $\Leftrightarrow$
every prime quotient ring of A is a π-regular ring of bounded index $\Leftrightarrow$
every prime quotient ring of A is a simple Artinian ring of index at most n $\Leftrightarrow$
$a^n A + N = a^{n+1} A + N$ for every $a \in A \Leftrightarrow$
$Aa^n + N = Aa^{n+1} + N$ for every $a \in A \Leftrightarrow$
$a^n A = a^{n+1} A$ for every $a \in A \Leftrightarrow$
$Aa^n = Aa^{n+1}$ for every $a \in A$.

5.10 [110]. A is a π-regular ring, and every right primitive quotient ring of A is a ring of bounded index $\Leftrightarrow$

A is strongly π-regular, and every indecomposable semiprimitive quotient ring $\overline{A}$ of A is a simple Artinian ring.

5.11 [76,101]. Let A be a π-regular ring. For every $a \in A$, denote by $h(a)$ the least positive integer n such that $a^n \in a^n A a^n$. Then

all right primitive quotient rings of A are rings of bounded index $\Leftrightarrow$ for every countable set $\{e_i\}_{i=1}^{\infty}$ of orthogonal idempotents of A and for every countable set $\{x_i\}_{i=1}^{\infty}$ of elements of A, there is a positive integer n such that $h(\cdots h(h(e_1x_1)e_2x_2)\cdots e_nx_n) = 0 \Leftrightarrow$ for every π-regular subring S of A, all indecomposable semiprimitive quotient rings of S are Artinian rings.

5.12 [110]. Let A be a π-regular semiprime ring which is a sum of n distributive right ideals.

Then A is a strongly π-regular semiprimitive ring of index at most n, every prime quotient ring of A is isomorphic to a ring of all $k \times k$ matrices over a division ring with $k \leqslant n$, and $a^n A = a^{n+1} A$, $Aa^n = Aa^{n+1}$ for every $a \in A$.

5.13 [110]. Let A be a π-regular right distributively generated ring, and let N be the prime radical of A. Then A is a strongly π-regular ring, and A/N is a semiprimitive ring of bounded index.

6. Nonnecessarily unital rings

The rings considered in this section are not necessarily unital. A ring A is an *exchange ring* if the following two equivalent conditions hold: (1) for every $a \in A$, there are $e = e^2 \in A$ and $r, s \in A$ such that $e = ar = s + a - as$; (2) for every $a \in A$, there are $e = e^2 \in A$ and $r, s \in A$ such that $e = ra = s + a - sa$ (see [3]).

For an element a in a ring A, we define the *local ring of A at a* as the additive subgroup aAa, endowed with the product $(axa) \cdot (aya) := axaya$.

6.1 [4]. A ring A is an exchange ring $\Leftrightarrow$ every left ideal of A is an exchange ring $\Leftrightarrow$ every principal left ideal of A is an exchange ring $\Leftrightarrow$ for every $a \in A$, the local ring of A at a is an exchange ring.

6.2 [3]. Let A be a ring without unit, and let R be a unital ring containing A as an ideal. Then

A is an exchange ring $\Leftrightarrow$ for every $a \in A$, there is $e = e^2 \in A$ with $e - a \in A(a - a^2)$ $\Leftrightarrow$ for every $a \in A$, there is $e = e^2 \in Aa$ and $c \in R$ such that $(1 - e) - c(1 - a) \in J(A)$ $\Leftrightarrow$ for every $a \in A$, there is $e = e^2 \in Aa$ such that $R = Ae + R(1 - a)$ $\Leftrightarrow$ for every $a \in A$, there is $e = e^2 \in Aa$ such that $1 - e \in R(1 - a)$.

6.3 [3]. Let A be an ideal of a (possibly non-unital) ring R. Then

(1) R is an exchange ring $\Leftrightarrow$ A and R/A are exchange rings, and all idempotents of R/A can be lifted to idempotents of R $\Leftrightarrow$ A and R/A are exchange rings, and the canonical homomorphism $K_0(R) \to K_0(R/A)$ is surjective.

(2) If A is a π-regular ring, then R is an exchange ring $\Leftrightarrow$ R/A is an exchange ring.

(3) If A is an exchange ring and R/A is a radical ring, then R is an exchange ring.

6.4 [3]. If A is an exchange ring, then for every positive integer n, the matrix ring A_n is an exchange ring.

6.5 [3]. Let R be a unital ring, and let A be a non-unital exchange ring which is an ideal of R.

(1) If $e = e^2 \in R$, then eAe is an exchange ring.

(2) The canonical homomorphism $K_0(R) \to K_0(R/A)$ is surjective $\Leftrightarrow$

for every positive integer n, all idempotents of the matrix ring $(R/A)_n$ can be lifted to idempotents of R_n.

(3) If every finitely generated projective R/A-module is isomorphic to a direct sum of cylic modules, then the canonical homomorphism $K_0(R) \to K_0(R/A)$ is surjective $\Leftrightarrow$

all idempotents of R/A can be lifted to idempotents of R.

(4) If R/A is a purely infinite right self-injective ring, then R is an exchange ring.

6.6 [3]. A is a semiperfect ring $\Leftrightarrow$ A is a semilocal ring, and the canonical homomorphism $K_0(R) \to K_0(R/J(R))$ is an isomorphism.

6.7 [3]. The C^*-algebras of real rank zero are precisely those C^*-algebras which are exchange rings.

6.8 [4]. Let A be a ring, and let $\{e_i\}_{i\in I}$ be a nonempty family of idempotents of A such that e_iAe_i is an exchange ring for all $i \in I$. Then the ideal of A generated by $\{e_i\}_{i\in I}$ is an exchange ring.

6.9 [4]. Let A be a ring, and let B be the ideal of A generated by all of the local idempotents of A. Then B is an exchange ring.

6.10 [4]. Let A be a ring, and let B be the ideal of A generated by all of the idempotents of A. Then A is an exchange ring $\Leftrightarrow$ A/B is a radical ring, and eAe is an exchange ring for all idempotents e of A.

6.11 [4]. Every ring contains a largest exchange ideal.

6.12 [81]. If A is a potent ring (resp. an I_0-ring), then every one-sided ideal of A is a potent ring (resp. an I_0-ring), and for any $a, b \in A$, the ring aAb is potent (resp. an I_0-ring).

6.13 [81]. Let A be an I_0-ring.

(1) For every positive integer n, the matrix ring A_n is an I_0-ring.

(2) For every nonzero $e = e^2 \in A$, the ring eAe is an I_0-ring.

(3) If e is a nonzero idempotent of A, then e is a primitive idempotent $\Leftrightarrow$ e is a local idempotent $\Leftrightarrow$ $L = Ae$ for every left ideal $L \subseteq Ae$ such that $L \not\subseteq J(A)$.

(4) All primitive idempotents of $A/J(A)$ can be lifted to local idempotents of A.

(5) If e, f are two primitive idempotents of A, then e and f are equivalent $eAf \not\subseteq J(A)$.

(6) If $J(A) = 0$, then A has a primitive idempotent $\Leftrightarrow$ A has a maximal left (right) annihilator.

(7) If $J(A) = 0$, then A is a primitive ring with nonzero socle $\Leftrightarrow$ A is a prime ring with a maximal left (right) annihilator.

6.14 [81]. A is an I_0-ring, and every nonzero idempotent of A contains a primitive idempotent $\Leftrightarrow$ every left (right) ideal $L \not\subseteq J(A)$ contains a primitive idempotent $\Leftrightarrow$ A is an I_0-ring, and every nonzero left (right) ideal L of $A/J(A)$ contains a primitive idempotent.

6.15 [81]. Let A be an I_0-ring. Then every nonzero idempotent of A contains a primitive idempotent, and any two primitive idempotents of A are equivalent $\Leftrightarrow$ A contains a primitive idempotent, and $J(A)$ is a prime radical $\Leftrightarrow$ $A/J(A)$ is a primitive ring with nonzero socle.

References

[1] H. Al-Ezeh, *Exchange PF-rings and almost PP-rings*, Mat. Vestnik **42** (2) (1990), 77–80.

[2] P. Ara, *Strongly π-regular rings have stable range one*, Proc. Amer. Math. Soc. **124** (11) (1996), 3293–3298.

[3] P. Ara, *Extensions of exchange rings*, J. Algebra **197** (2) (1997), 409–423.
[4] P. Ara, M. Gómez Lozano and M. Siles Molina, *Local rings of exchange rings*, Comm. Algebra **26** (12) (1998), 4191–4205.
[5] P. Ara, K.R. Goodearl, K.C. O'Meara and E. Pardo, *Separative cancellation for projective modules over exchange rings*, Israel J. Math., to appear.
[6] P. Ara, K.C. O'Meara and D.V. Tyukavkin, *Cancellation of projective modules over regular rings with comparability*, J. Pure Appl. Algebra **107** (1) (1996), 19–38.
[7] R. Arens and I. Kaplansky, *Topological representation of algebras*, Trans. Amer. Math. Soc. **63** (1948), 457–481.
[8] E.P. Armendariz and J.W. Fisher, *Regular P.I.-rings*, Proc. Amer. Math. Soc. **39** (2) (1973), 247–251.
[9] E.P. Armendariz, J.W. Fisher and R.L. Snider, *On injective and surjective endomorphisms of finitely generated modules*, Comm. Algebra **6** (7) (1978), 659–672.
[10] E.P. Armendariz, J.W. Fisher and S.A. Steinberg, *Central localization of regular rings*, Proc. Amer. Math. Soc. **46** (3) (1974), 315–321.
[11] G. Azumaya, *Strongly π-regular rings*, J. Fac. Sci. Hokkaido Univ. **13** (1954), 659–672.
[12] G. Azumaya, *F-semi-perfect modules*, J. Algebra **136** (1991), 73–85.
[13] A. Badawi, *On semicommutative π-regular rings*, Comm. Algebra **22** (1) (1993), 151–157.
[14] A. Badawi, *On Abelian π-regular rings*, Comm. Algebra **25** (4) (1997), 1009–1021.
[15] I. Beck, *On modules whose endomorphism ring is local*, Israel J. Math. **29** (4) (1978), 393–407.
[16] K.I. Beidar and A.V. Mikhalev, *Semiprime rings with bounded index of nilpotent elements*, Trudy Sem. Petrovsk., No. 13 (1988), 237–249.
[17] K.I. Beidar and R. Wisbauer, *Strongly and properly semiprime rings and modules*, Ring Theory, Proc. Ohio State–Denison Conf. 1992, World Sci. Pub., Singapore (1993).
[18] K.I. Beidar and R. Wisbauer, *Properly semiprime self-pp-modules*, Comm. Algebra **23** (3) (1995), 841–861.
[19] G.F. Birkenmeier, *A survey of regularity conditions and the simplicity of prime factor rings*, Vietnam J. Math. **23** (1) (1995), 29–38.
[20] G.F. Birkenmeier, J.Y. Kim and J.K. Park, *A connection between weak regularity and the simplicity of prime factor rings*, Proc. Amer. Math. Soc. **122** (1) (1994), 53–58.
[21] G.F. Birkenmeier, J.Y. Kim and J.K. Park, *Regularity conditions and the simplicity of prime factor rings*, J. Pure Appl. Algebra **115** (1997), 213–230.
[22] W.D. Burgess and P. Menal, *On strongly π-regular rings and homomorphisms into them*, Comm. Algebra **16** (1988), 1701–1725.
[23] W.D. Burgess and W. Stephenson, *Pierce sheaves of non-commutative rings*, Comm. Algebra **4** (1) (1976), 51–75.
[24] W.D. Burgess and W. Stephenson, *An analogue of the Pierce sheaf for non-commutative rings*, Comm. Algebra **6** (9) (1978), 863–886.
[25] W.D. Burgess and W. Stephenson, *Rings all of whose Pierce stalks are local*, Canad. Math. Bull. **22** (2) (1979), 159–164.
[26] V. Camillo and H.-P. Yu, *Exchange rings, units and idempotents*, Comm. Algebra **22** (12) (1994), 4737–4749.
[27] R. Camps and P. Menal, *Power cancellation for Artinian modules*, Comm. Algebra **19** (7) (1991), 2081–2095.
[28] R. Camps and P. Menal, *The power substitution property for rings of continuous functions*, J. Algebra **161** (2) (1993), 455–466.
[29] M. Chacron, *On algebraic rings*, Bull. Austral. Math. Soc. **1** (1969), 385–389.
[30] H. Chen, *Exchange rings, related comparability and power substitution*, Comm. Algebra **26** (10) (1998), 3383–3401.
[31] P. Crawley and B. Jónsson, *Refinements for infinite direct decompositions of algebraic systems*, Pacific J. Math. **14** (1964), 797–855.
[32] F. Dischinger, *Sur les anneaux fortement π-réguliers*, C. R. Acad. Sci. Paris **283** (8) (1976), A571–A573.
[33] N.I. Dubrovin, *Lifting of idempotents in algebraic algebra over a Henselian valuation ring*, Uspekhi Mat. Nauk **41** (5) (1986), 173–174.
[34] S. Elliger, *Zu dem Satz von Krull–Remak–Schmidt–Azumaya*, Math. Z. **115** (1970), 227–230.

[35] A. Facchini, *Module Theory: Endomorphism Rings and Direct Sum Decompositions in Some Classes of Modules*, Birkhäuser, Basel (1998).

[36] C. Faith and E. Walker, *Direct sum representations of injective modules*, J. Algebra **5** (1967), 203–221.

[37] J.W. Fisher and R.L. Snider, *On the Von Neumann regularity of rings with regular prime factor rings*, Pacific J. Math. **54** (1) (1974), 135–144.

[38] J.W. Fisher and R.L. Snider, *Rings generated by their units*, J. Algebra **42** (1976), 363–368.

[39] L. Fuchs, *On quasi-injective modules*, Ann. Scuola Norm. Sup. Pisa **23** (1969), 541–546.

[40] L. Fuchs, *On a substitution property of modules*, Monatsh. Math. **75** (1971), 198–204.

[41] L. Fuchs and K.M. Rangaswamy, *On generalized regular rings*, Math. Z. **107** (1968), 71–81.

[42] K.R. Goodearl, *Von Neumann Regular Rings*, Pitman, London (1979).

[43] K.R. Goodearl, *Surjective endomorphisms of finitely generated modules*, Comm. Algebra **15** (3) (1987), 589–609.

[44] K.R. Goodearl and D.E. Handelman, *Homogenization of regular rings of bounded index*, Pacific J. Math. **84** (1) (1979), 63–78.

[45] K.R. Goodearl and R.R. Warfield, *Algebras over zero-dimensional rings*, Math. Ann. **223** (2) (1976), 157–168.

[46] V. Gupta, *Weakly π-regular rings and group rings*, Math. J. Okayama Univ. **19** (2) (1977), 123–127.

[47] V. Gupta, *A generalization of strongly regular rings*, Acta Math. Hungar. **43** (1984), 57–61.

[48] M. Harada, *On categories of indecomposable modules. II*, Osaka J. Math. **8** (1971), 309–321.

[49] M. Harada, *Supplementary remarks on categories of indecomposable modules*, Osaka J. Math. **9** (1972), 49–55.

[50] M. Harada, *Note on categories of indecomposable modules*, Publ. Dép. Math. (Lyon) **9** (4) (1972), 11–25.

[51] M. Harada, *Applications of factor categories to completely indecomposable modules*, Publ. Dép. Math. (Lyon) **11** (2) (1974), 19–104.

[52] M. Harada, *On the exchange property in a direct sum of indecomposable modules*, Osaka J. Math. **12** (3) (1975), 719–736.

[53] M. Harada, *A ring theoretical proof in a factor category of indecomposable modules*, J. Math. Soc. Japan **28** (1) (1976), 160–167.

[54] M. Harada and T. Ishii, *On perfect rings and the exchange property*, Osaka J. Math. **12** (2) (1975), 481–491.

[55] M. Harada and Y. Sai, *On categories of indecomposable modules. I*, Osaka J. Math. **7** (1970), 323–344.

[56] R.E. Hartwig, *On a theorem of Flanders*, Proc. Amer. Math. Soc. **85** (3) (1982), 613–616.

[57] Y. Hirano, *Some studies on strongly π-regular rings*, Math. J. Okayama Univ. **20** (2) (1978), 141–149.

[58] Y. Hirano, *On π-regular rings with involution*, Math. J. Okayama Univ. **27** (1985), 45–47.

[59] Y. Hirano, *Some characterizations of π-regular rings of bounded index*, Math. J. Okayama Univ. **32** (1990), 97–101.

[60] Y. Hirano and J.P. Park, *On self-injective strongly π-regular rings*, Comm. Algebra **21** (1) (1993), 85–91.

[61] C.Y. Hong, H.K. Kim and J.K. Park, *Rings with restricted descending chain conditions*, Comm. Algebra **25** (8) (1997), 2579–2584.

[62] T. Ishii, *On locally direct summands of modules*, Osaka J. Math. **12** (2) (1975), 473–482.

[63] N. Jacobson, *Structure of Rings*, AMS, Providence (1968).

[64] J. Kado, *Note on exchange property*, Math. J. Okayama Univ. **18** (2) (1975/1976), 153–157.

[65] U.S. Kahlon, *Problem of Krull–Schmidt–Remak–Azumaya–Matlis*, J. Indian Math. Soc. **35** (1971), 255–261.

[66] H. Kanbara, *Note on Krull–Remak–Schmidt–Azumaya's theorem*, Osaka J. Math. **9** (1972), 409–413.

[67] I. Kaplansky, *Topological representation of algebras*, Trans. Amer. Math. Soc. **68** (1950), 62–75.

[68] M. Kutami and K. Oshiro, *Direct sums of nonsingular indecomposable injective modules*, Math. J. Okayama Univ. **20** (2) (1978), 91–99.

[69] M. Kutami and K. Oshiro, *An example of a ring whose projective modules have the exchange property*, Osaka J. Math. **17** (2) (1980), 415–420.

[70] N. Lang, *On centroids of generalized regular rings*, Ann. Scuola Norm. Sup. Pisa (3) **26** (1972), 573–585.

[71] U. Leron, *Matrix methods in decompositions of modules*, Linear Algebra Appl. **31** (1980), 159–171.

[72] L. Lesieur, *Sur quelques classes d'anneaux non commutatifs*, Comm. Algebra **14** (8) (1986), 1481–1488.

[73] J. Levitzki, *On the structure of algebraic algebras and related rings*, Trans. Amer. Math. Soc. **74** (1953), 384–409.

[74] V.T. Markov, *On B-rings with polynomial identity*, Trudy Sem. Petrovsk., No. 7 (1981), 232–238.

[75] E. Matlis, *Injective modules over Noetherian rings*, Pacific J. Math. **8** (1958), 511–528.

[76] P. Menal, *On π-regular rings whose primitive factor rings are Artinian*, J. Pure Appl. Algebra **20** (1) (1981), 71–78.

[77] P. Menal, *Spectral Banach algebras of bounded index*, J. Algebra **154** (1) (1993), 27–66.

[78] K. Meyberg and B. Zimmermann-Huisgen, *Rings with descending chain condition on certain principal ideals*, Indag. Math. **39** (3) (1977), 225–229.

[79] S.H. Mohamed and B.J. Muller, *Continuous and Discrete Modules*, London Math. Soc. Lecture Notes Series 147, Cambridge Univ. Press (1990).

[80] G.S. Monk, *A characterization of exchange rings*, Proc. Amer. Math. Soc. **35** (2) (1972), 349–353.

[81] W.K. Nicholson, *I-rings*, Trans. Amer. Math. Soc. **207** (1975), 361–373.

[82] W.K. Nicholson, *Semiregular modules and rings*, Canad. J. Math. **28** (5) (1976), 1105–1120.

[83] W.K. Nicholson, *Lifting idempotents and exchange rings*, Trans. Amer. Math. Soc. **229** (1977), 269–278.

[84] W.K. Nicholson, *On exchange rings*, Comm. Algebra **25** (6) (1997), 1917–1918.

[85] C. Niţă, *Anneaux N-réguliers*, C. R. Acad. Sci. Paris **268** (1969), A1241–A1243.

[86] C. Niţă, *Remarques sur les anneaux N-réguliers*, C. R. Acad. Sci. Paris **271** (1970), A345–A348.

[87] M. Ohori, *On strongly π-regular rings and periodic rings*, Math. J. Okayama Univ. **19** (2) (1977), 123–127.

[88] K. Oshiro, *Projective modules over von Neumann regular rings have the finite exchange property*, Osaka J. Math. **20** (1983), 695–699.

[89] K. Oshiro and S.T. Rizvi, *The exchange property of quasi-continuous modules with the finite exchange property*, Osaka J. Math. **33** (1996), 217–234.

[90] F. van Oystaeyen and J. van Geel, *Local-global results for regular rings*, Comm. Algebra **4** (9) (1976), 811–821.

[91] E. Pardo, *Comparability, separativity, and exchange rings*, Comm. Algebra **24** (9) (1996), 2915–2929.

[92] E. Pardo, *Metric completions of ordered groups and K_0 of exchange rings*, Trans. Amer. Math. Soc. **350** (3) (1998), 913–933.

[93] R.S. Pierce, *Modules over Commutative Regular Rings*, Mem. Amer. Math. Soc., No. 70, Providence, RI (1967).

[94] K.M. Rangaswamy, *Endomorphism representations of Zorn rings and subclasses of rings with enough idempotents*, J. Reine Angew. Math. **276** (1975), 44–55.

[95] L.A. Skornyakov, *Complemented Modular Lattices and Regular Rings*, Oliver and Boyd, Edinburgh and London (1964).

[96] J.B. Srivastava and S.K. Shah, *Semilocal and semiregular group rings*, Nederl. Akad. Wetensch. Indag. Math. **42** (3) (1980), 347–352.

[97] A. Steger, *Elementary factorization in π-regular rings*, Canad. J. Math. **18** (1966), 307–313.

[98] J. Stock, *On rings whose projective modules have the exchange property*, J. Algebra **103** (1986), 437–453.

[99] K. Tanabe, *On rings whose Artinian modules are precisely Noetherian modules*, Comm. Algebra **22** (10) (1994), 4023–4032.

[100] D.V. Tjukavkin, *Rings all of whose one-sided ideals are generated by idempotents*, Comm. Algebra **17** (5) (1989), 1193–1198.

[101] D.V. Tjukavkin, *On regular rings of bounded index*, Mat. Zametki **54** (1) (1993), 87–93.

[102] H. Tominaga, *Some results on π-regular rings of bounded index*, Math. J. Okayama Univ. **4** (1955), 135–141.

[103] J. Trlifaj, *Steady rings may contain large sets of orthogonal idempotents*, Abelian Groups and Modules, A. Facchini and C. Menini, eds, Kluwer Acad. Publishers, Dordrecht (1995), 467–473.

[104] A.A. Tuganbaev, *Distributive semiprime rings*, Mat. Zametki **58** (5) (1995), 736–761.

[105] A.A. Tuganbaev, *Bass rings and perfect rings*, Uspekhi Mat. Nauk **51** (1) (1996), 173–174.

[106] A.A. Tuganbaev, *Rings over which every module has a maximal submodule*, Mat. Zametki **61** (3) (1997), 407–415.

[107] A.A. Tuganbaev, *Maximal submodules and locally perfect rings*, Mat. Zametki **64** (1) (1998), 136–142.

[108] A.A. Tuganbaev, *Semidistributive Modules and Rings*, Kluwer Acad. Publishers, Dordrecht (1998).

[109] A.A. Tuganbaev, *Quasi-projective modules with the finite exchange property*, Uspekhi Mat. Nauk **54** (2) (1999), 191–192.
[110] A.A. Tuganbaev, *Rings with many idempotents*, Preprint (1999).
[111] R.B. Warfield, *Decompositions of injective modules*, Pacific J. Math. **31** (1969), 263–276.
[112] R.B. Warfield, *A Krull–Schmidt theorem for infinite sums of modules*, Proc. Amer. Math. Soc. **22** (1969), 460–465.
[113] R.B. Warfield, *Exchange rings and decompositions of modules*, Math. Ann. **199** (1972), 31–36.
[114] R. Wisbauer, *F-semiperfekte und perfekte Moduln in $\sigma[M]$*, Math. Z. **173** (1980), 229–234.
[115] R. Wisbauer, *Foundations of Module and Ring Theory*, Gordon and Breach, Philadelphia (1991).
[116] T.S. Wu, *Exchange rings with primitive quotient rings Artinian*, Algebra Colloq. **3** (3) (1996), 225–230.
[117] T.S. Wu and W.T. Tong, *Finitely generated projective modules over exchange rings*, Manuscripta Math. **86** (2) (1995), 149–157.
[118] T.S. Wu and Y. Xu, *On the stable range condition of exchange rings*, Comm. Algebra **25** (7) (1997), 2355–2363.
[119] W. Xue, *Three questions on strongly π-regular rings and regular rings*, Comm. Algebra **21** (2) (1993), 699–704.
[120] W. Xue, *Semiregular modules and F-semiperfect modules*, Comm. Algebra **23** (3) (1995), 1035–1046.
[121] K. Yamagata, *Non-singular rings and Matlis' proplem*, Sci. Rep. Tokyo Kyoiku Daigaku Sect. A **11** (1972), 186–192.
[122] K. Yamagata, *A note on a proplem of Matlis*, Proc. Japan Acad. **49** (1973), 145–147.
[123] K. Yamagata, *The exchange property and direct sums of indecomposable injective modules*, Pacific J. Math. **55** (1) (1974), 301–317.
[124] K. Yamagata, *On projective modules with the exchange property*, Sci. Rep. Tokyo Kyoiku Daigaku Sect. A **12** (1974), 149–158.
[125] K. Yamagata, *On rings of finite representation type and modules with the finite exchange property*, Sci. Rep. Tokyo Kyoiku Daigaku Sect. A **13** (1975), 1–6.
[126] H.-P. Yu, *On modules for which the finite exchange property implies the finite exchange property*, Comm. Algebra **22** (10) (1994), 3887–3901.
[127] H.-P. Yu, *On quasi-duo rings*, Glasgow Math. J. **37** (1995), 21–31.
[128] H.-P. Yu, *On $\aleph_0$-quasi-continuous exchange rings*, Comm. Algebra **23** (6) (1995), 2187–2197.
[129] H.-P. Yu, *Stable range one for exchange rings*, J. Pure Appl. Algebra **98** (1) (1995), 105–109.
[130] H.-P. Yu, *On strongly π-regular rings of stable range one*, Bull. Austral. Math. Soc. **51** (3) (1995), 433–437.
[131] H.-P. Yu, *On Krull–Schmidt and a problem of Matlis*, Comm. Algebra **24** (9) (1996), 2851–2858.
[132] H.-P. Yu, *On the structure of exchange rings*, Comm. Algebra **25** (2) (1997), 661–670.
[133] B. Zimmermann-Huisgen, *Rings whose right modules are direct sums of indecomposable modules*, Proc. Amer. Math. Soc. **77** (2) (1979), 191–197.
[134] B. Zimmermann-Huisgen and W. Zimmermann, *Algebraically compact rings and modules*, Math. Z. **161** (1978), 81–93.
[135] B. Zimmermann-Huisgen and W. Zimmermann, *Classes of modules with the exchange property*, J. Algebra **88** (1984), 416–434.

This Page Intentionally Left Blank

Separable Algebras

Freddy Van Oystaeyen

Department of Mathematics, University of Antwerp, B-2610 Wilrijk, Belgium

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

0. Introduction

The obvious starting point for a theory of separable algebras should be the classical Galois theory of finite dimensional field extensions. In this theory separability appears as a relative property in the sense that it describes a property of an extension of fields closely related to properties of polynomials determining the field extension. That a more general Galois theory of extensions of commutative rings could be a tool of interest was suggested by the consideration of rings of integers in number field extensions. On the other, noncommutative, hand the theory of separable algebra is rooted in the representation theory of finite groups; indeed, the central simple algebras representing elements of the Brauer group are connected to irreducible characters of a finite group, or in other words, to direct summands of the group ring of a finite group over the field considered. Central simple algebras are separable algebras over their centre which is a field in this case.

Just as the arithmetic of a field may sometimes be studied in terms of the valuation rings in it, a similar idea may be applied to the central simple algebras, e.g. skewfields, replacing valuation theory by the theory of maximal orders in the central simple algebra defined over a valuation ring in the central field. Originally, this idea was not explicitly present when G. Azumaya introduced the so-called maximally central algebras in [54] stressing the property that ideals of these algebras are generated by their central parts. The more general notion, nowadays known as "Azumaya algebra" is in fact due to M. Auslander [50] and these algebras are exactly those that are separable over their centres. This definition may be phrased in terms of elementary homological algebra and is particularly suitable for cohomological interpretations as well as for geometric considerations focussing on Brauer groups of coordinate rings of algebraic varieties. In fact a geometrically defined Brauer group of a scheme reduces to the Brauer group of a ring when the scheme is affine. For projective schemes (Proj) the Brauer group is not that easy to describe and in order to present an algebraic equivalent it is necessary to introduce graded Brauer groups as well as Brauer groups relative to a Serre quotient category corresponding to a "global section" functor (localization). As it turns out, a very general theory of separability is obtained by viewing separability as a property of a functor, e.g., the restriction of scalars functor, in the sense of [401]. We develop the theory from general to particular; this may be less didactical perhaps but it is the most elegant approach avoiding eventually boring step by step generalization. Along the way we provide information concerning very particular separable algebras, e.g., those representing specific subgroups of the Brauer group like the Schur subgroup.

1. Separable functors

Let $\mathcal{C}$ and $\mathcal{D}$ be arbitrary categories and consider a covariant functor $F : \mathcal{C} \to \mathcal{D}$.

The functor F is said to be a *separable functor* if for all objects M, N in $\mathcal{C}$ there are given maps $\varphi^F_{M,N}$,

$$\varphi^F_{M,N} : \mathrm{Hom}_{\mathcal{D}}\big(F(M), F(N)\big) \to \mathrm{Hom}_{\mathcal{C}}(M, N),$$

satisfying the following properties:

SF.1 For $\alpha \in \mathrm{Hom}_{\mathcal{C}}(M, N)$, $\varphi^F_{M,N}(F(\alpha)) = \alpha$,

SF.2 If we have objects M', N' in $\mathcal{C}$ and $\alpha \in \mathrm{Hom}_{\mathcal{C}}(M, M')$, $\beta \in \mathrm{Hom}_{\mathcal{C}}(N, N')$, $f \in \mathrm{Hom}_{\mathcal{D}}(F(M), F(N))$, $g \in \mathrm{Hom}_{\mathcal{D}}(F(M'), F(N'))$ such that the following diagram is commutative:

$$\begin{tikzcd} F(M) \arrow[r, "f"] \arrow[d, "F(\alpha)"] & F(N) \arrow[d, "F(\beta)"] \\ F(M') \arrow[r, "g"] & F(N') \end{tikzcd}$$

then the following diagram is commutative as well:

$$\begin{tikzcd} M \arrow[r, "\varphi^F_{M,N}(f)"'] \arrow[d, "\alpha"] & N \arrow[d, "\beta"] \\ M' \arrow[r, "\varphi^F_{M',N'}(g)"'] & N' \end{tikzcd}$$

Note that SF.1 holds in case F is a full faithful functor, that is, whenever for M, N in $\mathcal{C}$ the map $\mathrm{Hom}_{\mathcal{C}}(M, N) \to \mathrm{Hom}_{\mathcal{D}}(F(M), F(N))$ defined by $f \mapsto F(f)$ is bijective. In the following lemma we list a few elementary observations on separability as defined above.

LEMMA 1.1. *An equivalence of categories is separable. If $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{E}$ are separable functors then the composition GF is separable; conversely if GF is separable then is F separable.*

PROPOSITION 1.2. *Given a separable functor $F : \mathcal{C} \to \mathcal{D}$ and objects M, N in $\mathcal{C}$ then the following properties hold:*

(1) *If $f \in \mathrm{Hom}_{\mathcal{C}}(M, N)$ is such that $F(f)$ is a split map then f is split. On the other hand if $F(f)$ is co-split then f is also co-split.*

(2) *Assume that $\mathcal{C}, \mathcal{D}$ are Abelian categories. If $F(M)$ is quasi-simple (i.e. every subobject splits off) and F preserves monomorphisms then M is quasi-simple in $\mathcal{C}$. On the other hand if F preserves epimorphisms, resp. monomorphisms, and $F(M)$ is projective, resp. injective, in $\mathcal{D}$, then M is projective, resp. injective in $\mathcal{C}$.*

All rings considered will be associative rings with unit. Let R be a ring. An Abelian group M that is a left and a right R-module such that $(am)b = a(mb)$ for $m \in M$, $a, b \in R$, is said to be an *R-bimodule*. In case R is a C-algebra we may look at those R-bimodules satisfying the extra condition $cm = mc$ for all $c \in C$, $m \in M$ and we say that such an R-bimodule is an *(R, C)-bimodule*. If S is another ring like R then it is clear what is meant by an *R-S-bimodule*. We will sometimes write ${}_RM_S$ to indicate that M is an R-S-bimodule with R acting on the left, S acting on the right. An additive map $f : {}_RM_S \to {}_RM_S$ is said to be an *R-S-bimodule morphism* when it is left R-linear and right S-linear. When $S = R$ we have defined *R-bimodule morphisms* and an R-bimodule morphism between (R, C)-bimodules is an *(R, C)-bimodule morphism*.

To a ring morphism $\varphi : R \to S$ there corresponds a restriction of scalars functor $\varphi_* : S\text{-mod} \to R\text{-mod}$ and also an induction functor $S \otimes_R : R\text{-mod} \to S\text{-mod}$, $M \mapsto S \otimes_R M$. In general φ_* is exact but the functor $S \otimes_R -$ is only right exact.

THEOREM 1.3. *Given any ring morphism $\varphi : R \to S$ then:*

(1) *φ_* is separable if and only if $\psi : S \otimes_R S \to S$, $s \otimes s' \mapsto ss'$ splits as an S-bimodule map.*

(2) *$S \otimes_R -$ is separable if and only if φ splits as an R-bimodule map.*

COROLLARY 1.4. *If $\varphi : R \to S$ is an epimorphism in the category of rings then φ_* is a separable functor.*

Instead of considering S-bimodules we may reduce everything to a class of left modules replacing $S \otimes_R S$ by $S^e = S \otimes_R S^0$ where S^0 is the opposite ring of S and we may rephrase Theorem 1.3 as follows:

THEOREM 1.5. *φ_* is separable if and only if S is a projective left $S \otimes_R S^0$-module. If the foregoing holds then the categories (R, S)-bimod and S^e-mod are equivalent.*

The theory of separable functors has obtained some interesting applications in the study of group graded rings: of course the forgetful functor forgetting the gradation of the modules is a separable functor, but a more interesting fact is that the left adjoint of the forgetful functor is also separable. We refer to [400] for these applications. In the next section we focus attention on separable algebras, that is separability over a commutative ring, but here we finish with an observation about separability of an R-algebra A over an R-subalgebra B.

PROPOSITION 1.6. *Given an inclusion $\varphi : B \to A$ of R-algebras, then the following statements are equivalent:*

(1) *φ_* is separable (we may also say: A is separable over B).*

(2) *There is an $e \in A \otimes_B A$ such that $\mu(e) = 1$ and $ae = ea$ for all $a \in A$, where $\mu : A \otimes_B A \to A$ is the A-bimodule map given by $a \otimes a' \mapsto aa'$.*

(3) *For every commutative R-algebra C and given (C, A)-bimodule N we have that $\mu_N : \varphi_*(N) \otimes_B A \to N$, $n \otimes a \mapsto na$ is a split surjective morphism of (C, A)-bimodules.*

2. Separable algebras and Azumaya algebras

In this section C is a commutative ring with unit. Let A be a C-algebra and A^0 the opposite algebra, that is the C-module A but with multiplication now defined by $x \circ y = yx$ for $x, y \in A$. In Section 1 we agreed to call A a separable C-algebra when A is a projective A^e-module. When A is separable over its centre $C = Z(A)$ then A is said to be an *Azumaya algebra*. An idempotent e as in Proposition 1.6 (2) is called a *separability idempotent* for A. A separability idempotent can be described by the following somewhat technical lemma.

LEMMA 2.1. *Given the exact sequence* $0 \to \operatorname{Ker}\mu \to A \otimes_C A^0 \to A \to 0$ *and a set of* C*-generators* $\{u_\alpha,\ \alpha \in \mathcal{A}\}$ *for* A.

(1) $\operatorname{Ker}\mu$ *is the left* A^e*-module generated by* $\{u_\alpha \otimes 1 - 1 \otimes u_\alpha,\ \alpha \in \mathcal{A}\}$.

(2) *If some* $x \in A^e$ *satisfies* $\mu(x) = 1_A$ *then* x *is a separability idempotent for* A *if and only if* $(\operatorname{Ker}\mu)x = 0$.

(3) *If* e *is a separability idempotent for* A *then* $e^2 = e$.

(4) *The element* $\sum_{j=1}^n x_j \otimes y_i$ *is a separability idempotent for* A *if and only if*: $\sum_{j=1}^n x_j y_j = 1_A$, *and for all* $\alpha \in \mathcal{A}$ $\sum_{j=1}^n u_\alpha x_j \otimes y_j = \sum_{j=1}^n x_j \otimes y_j u_\alpha$.

EXAMPLE 2.2. $M_n(C)$ is an Azumaya algebra.

Let G be a finite group and consider the group algebra CG. If $|G| = n$ is invertible in C then CG is separable over C. However if $|G'|^{-1} \in C$, where G' is the commutator subgroup of G, then CG is an Azumaya algebra but not necessarily separable over C. For example let $\mathbb{F}_p$ be a field of characteristic $p \neq 0$ and consider a Sylow p-group H in G, then $\mathbb{F}_P G$ is separable over $\mathbb{F}_p H$ but $\mathbb{F}_P G$ is only $\mathbb{F}_P$-separable when $H = 1$. At this point it is interesting to mention Higman's theorem concerning the finite representation type property. $\mathbb{F}_P G$ has finite representation type if and only if the Sylow p-groups of G are cyclic.

PROPOSITION 2.3 (Villamayor, Zelinsky). *A separable* C*-algebra that is projective as a* C*-module is finitely generated over* C. *When* k *is a field then a separable* k*-algebra is finite-dimensional.*

Using the Morita-theorems (cf. F. De Meyer and E. Ingraham, [168], p. 16) we obtain:

PROPOSITION 2.4. *For a* C*-algebra* A, *the following are equivalent*

(1) A *is an Azumaya algebra over* C.

(2) A *is a finitely generated* C*-module that is projective and a generator for* C-mod. *Moreover* $A^e \cong \operatorname{End}_C(A)$.

To an (A, C)-bimodule M we may associate the A-commutator module $M^A = \{m \in M,\ am = ma$ for all $a \in A\}$ and thus define a functor $(-)^A : A^e\text{-mod} \to R\text{-mod}$.

PROPOSITION 2.5. $\operatorname{Hom}_{A^e}(A, -) \cong (-)^A$ *is a natural equivalence of functors. Consequently* A *is a separable* C*-algebra if and only if* $(-)^A$ *is a right exact functor.*

For A-modules M and N we may define a structure of an (A, C)-bimodule on $\operatorname{Hom}_R(M, N)$ as follows. For $f \in \operatorname{Hom}_R(M, N)$ and $a \in A$ we put $(fa)(m) = f(am)$, $(af)(m) = af(m)$; this definition implies: $\operatorname{Hom}_R(M, N)^A = \operatorname{Hom}_A(M, N)$ and this leads to the following.

COROLLARY 2.6. *Let* A *be a separable* C*-algebra and consider an* A*-module* P *that is projective as a* C*-module then* P *is projective as an* A*-module. In particular a separable algebra over a field is semi-simple, therefore it is classically separable in the sense that* $J(A \otimes_k l) = 0$ *for every field extension* l *of* k.

The theory of Azumaya algebras over fields is now reduced to the study of central simple algebras, i.e. the Brauer group of a field. Further development of the theory of Azumaya algebras will show a few similarities but even more essential points of difference. First we consider several change of ring results and product theorems.

PROPOSITION 2.7. *Let C_1 and C_2 be commutative R-algebras and A_i an algebra over C_i, $i = 1, 2$. If A_1, A_2 resp., is separable over C_1, C_2 resp. then $A_1 \otimes_R A_2$ is $C_1 \otimes_R C_2$-separable. Conversely if $A_1 \otimes_R A_2$ is $C_1 \otimes_R C_2$-separable and A_2 is a faithful R-module containing R as a direct summand then A_1 is a separable C_1-algebra.*

COROLLARY 2.8. (1) *Let A be a separable C-algebra and D any commutative C-algebra, then $A \otimes_C D$ is a separable D-algebra.*

(2) *If A_1, A_2 are separable R-algebras such that A_2 contains R as a direct summand as an R-module then the separability of the R-algebra $A_1 \otimes_1 A_2$ yields the R-separability of A_1.*

(3) *Let A be a C-algebra and D a commutative C-algebra containing C as a direct summand as a C-module then A is C-separable if $A \otimes_C D$ is D-separable. When $1 \otimes D$ is the centre of $A \otimes_C D$ then C in the centre of A.*

PROPOSITION 2.9. *Let A be R-separable and I an ideal of A; then A/I is a separable R-algebra, hence a separable $R/I \cap R$-algebra. Moreover $Z(A/I) = Z(A) + I/I$. Consequently the separability of a direct sum of algebras is determined by summand-wise separability.*

PROPOSITION 2.10. *Let C be a commutative separable R-algebra and let A be a separable C-algebra then A is a separable R-algebra. On the other hand, if the C-algebra A is a separable R-algebra then A is separable over C; moreover if A is faithfully projective as a C-module then C is separable over R.*

Separability properties enjoy a satisfactory local-global behaviour with respect to prime or maximal ideals.

PROPOSITION 2.11. *Let A be a C-algebra such that A is a finitely generated C-module. The following statements are equivalent:*

(1) *A is a separable C-algebra.*
(2) *For a prime ideal p of C the localization A_p is separable over C_p.*
(3) *For a maximal ideal w of C the localization A_w is separable over C_w.*
(4) *For any prime ideal p of C, A/pA is separable over C/p.*
(5) *For every maximal ideal w of R, A/wA is separable over c/w.*

For Azumaya algebras we obtain the following specification of the foregoing proposition.

THEOREM 2.12. *For an R-algebra A the following statements are equivalent:*

(1) *A is an Azumaya algebra over R.*
(2) *A is faithfully projective as an R-module and $A^e \cong \mathrm{End}_R(A)$ as R-algebras.*
(3) *The functors $N \mapsto A \otimes_R N$ and $M \to M^A$ define an equivalence between the categories R-mod and A^e-mod.*

(4) *A is a faithful finitely generated R-module and A/wA is a central simple algebra over R/w for all maximal ideals w of R.*
(5) *There is an R-algebra B and a faithfully projective R-module P such that $A \otimes_R B \simeq \mathrm{End}_R(P)$ as R-algebras.*
(6) *A satisfies any one of the conditions of Proposition* 2.11 *but with $C = Z(A)$.*

COROLLARY 2.13. (1) *For an ideal I of an Azumaya algebra A we have that $I = A(I \cap A)$.*

(2) *For every $(A, Z(A))$-bimodule M, $M \cong A \otimes_{Z(A)} M^A$.*

(3) *Every $Z(A)$-endomorphism of A is an automorphism.*

(4) *If $R \subset Z(A)$ then A is a separable R-algebra if and only if A is an Azumaya algebra and $Z(A)$ is separable over R.*

Two Azumaya algebras A and B over C are said to be similar, written $A \sim B$, if $A \otimes_C B^0 \cong \mathrm{End}_C(P)$ for some faithfully projective C-module P, that is if and only if $A \otimes_C \mathrm{End}_C(P_1) \cong B \otimes_C \mathrm{End}_C(P_2)$ for some faithfully projective C-modules P_1 and P_2. The operation $\otimes_C$ induces a multiplication in the set similarity classes of Azumaya algebras over C. We obtain an Abelian group $\mathrm{Br}(C)$ with as neutral element the class of C, say $[C] = 1$, and the inverse of a class $[A]$ is the class $[A^0]$. The group $\mathrm{Br}(C)$ is called the Brauer group of C. When $C = k$ is a field then $\mathrm{Br}(k)$ is the classical Brauer group of a field consisting of classes of k-central simple algebras.

In the latter case $[A] = [B]$ means that $M_m(A) \cong M_n(B)$ for some n and $m \in \mathbb{N}$. Over a ring C the role of the matrix rings over fields is now played by endomorphism rings of projective modules. When D is any commutative C-algebra then the correspondence $A \mapsto A \otimes_C D$ induces a group morphism $\mathrm{Br}(C) \to \mathrm{Br}(D)$ and so we may view $\mathrm{Br}(-)$ as a covariant functor from the category of commutative rings to Abelian groups. In order to understand the Brauer group of a commutative ring in depth it is necessary to develop the algebraic structure theory of Azumaya algebras on the one hand and the cohomological theory yielding information on the group structure of the Brauer group on the other hand. Moreover, as can be inferred from the local-global results intrinsic in Proposition 2.11 and Theorem 2.12, the theory of the Brauer group of a commutative ring is closely linked to a geometrically defined version of the Brauer group of a scheme and in this way the topic places itself between algebraic geometry, étale cohomology and non-commutative algebra. Before paying more attention to the Brauer group, let us finish this section with two well-known commutator theorems.

THEOREM 2.14. *Let A be an Azumaya algebra over C and let B be subalgebra of A such that B is an Azumaya algebra and B contains C. Put $C_A(B) = \{a \in A,\ ab = ba$ for $b \in B\}$; then $C_A(B)$ is an Azumaya algebra and $C_A(C_A(B)) = B$. If moreover $Z(B) = C$ then $Z(C_A(B)) = C$ and $A \cong A \otimes_C C_A(B)$ as C-algebras.*

THEOREM 2.15. *Let A be an Azumaya algebra over C and let B_1 and B_2 be subalgebras such that $A \cong B_1 \otimes_C B_2$ is a C-algebra isomorphism (defined by: $b_1 \otimes b_2 \mapsto b_1 b_2$) then B_1 and B_2 are Azumaya algebras over C such that $C_A(B_1) = B_2$, $C_A(B_2) = B_1$.*

3. The Brauer group of a commutative ring

In this section C will be a commutative ring, however it may sometimes be useful to have a finiteness condition on C. The following proposition shows that for several properties it will not be restrictive to assume that C is a Noetherian ring.

PROPOSITION 3.1. *If A is an Azumaya algebra over C then there exists a Noetherian subring C' of C and an Azumaya algebra A' over C' contained in A such that $A = C \otimes_{C'} A'$.*

For a C-Azumaya algebra A we say that a commutative C-algebra D is a splitting ring if $D \otimes_C A = \mathrm{End}_D(P)$ for some faithfully projective D-module P. The morphism $C \to D$ yields a group morphism $\mathrm{Br}(C) \to \mathrm{Br}(D)$ and we write $\mathrm{Br}(D/C)$ for the kernel of the latter map, in other words $\mathrm{Br}(D/C)$ consists of the classes of C-Azumaya algebras split by the extension D of C. The existence of splitting extensions is not a trivial problem, over fields this may be settled by considering maximal commutative subfields in the skewfield representing the class of the Brauer group under consideration. Over a commutative ring the situation is more complicated but we may still define a similar concept: a maximal commutative C-subalgebra of the Azumaya algebra A over C is a commutative subalgebra of A not contained in a larger commutative subalgebra. If D is a maximal commutative subalgebra of A then obviously D contains C and $C_A(D) = D$. For separable extensions D of C the situation resembles the existing theory in case $C = k$, a field.

THEOREM 3.2. *Let D be a separable extension of C* (in the sense of Section 1) *such that D is faithfully projective as a C-module. Then D is a splitting ring for the Azumaya algebra A over C if and only if there is a representative B in the Brauer class* $[A]$ *such that B contains a maximal commutative subalgebra isomorphic to D.*

Splitting rings with the properties mentioned in the theorem do exist when C is a local ring, cf. Auslander and Goldman [50] and Endo and Watanabe [194], but its existence for general C is an unsolved problem. In fact it is also open whether a splitting extension D/C (for an Azumaya algebra A over C) such that D/C is faithfully flat can always be found.

A C-module I is said to be invertible if there exists a C-module J such that $I \otimes_C J = C$, i.e. is projective of rank one. The set $\mathrm{Pic}(C)$ obtained by taking isomorphism classes of invertible modules (or classes of invertible ones modulo free ones) is a group with respect to the operation induced by $\otimes_C$. The commutativity of C entails that $\mathrm{Pic}(C)$ is an Abelian group.

PROPOSITION 3.3. *If* $\mathrm{Pic}(C)$ *is torsionfree then every automorphism of an Azumaya algebra A over C is an inner automorphism* (*that is conjugation by a unit of A*).

The elements of $\mathrm{Pic}(C)$ appear in composing faithfully projective Δ-modules where A is an Azumaya over C.

PROPOSITION 3.4. *Let M and N be faithfully projective Δ-modules, then* $\mathrm{End}_A(M) \cong \mathrm{Hom}_A(N)$ *as C-algebras, if and only if $N \cong M \otimes_C I$ in A-*mod *for some* $[I] \in \mathrm{Pic}(C)$.

Let us write $Q_A(M)$ for the set of isomorphism classes as (left) A-modules of faithfully projective A-modules N such that $\mathrm{End}_A(M) = \mathrm{End}_A(N)$ as C-algebras. In view of the proposition we have that each such N is of the form $M \otimes_C I_{(N)}$ with $[I_{(N)}] \in \mathrm{Pic}(C)$. We may define on $Q_A(M)$ the structure of an Abelian group by putting $(M \otimes_C I_{(N_1)})(M \otimes_C I_{(N_2)}) = M \otimes_C (I_{(N_1)} \otimes_C I_{(N_2)})$.

PROPOSITION 3.5. *Let A be an Azumaya algebra over C and $B = \mathrm{End}_A(M)$ for a faithfully projective A-module M then the following sequence of Abelian groups is exact*:

$$1 \to \mathrm{Inn}\, B \to \mathrm{Aut}\, B \to \mathrm{Pic}(C) \to Q_A(M) \to 1,$$

where Inn B *stands for the group of inner automorphisms.*

The condition $\mathrm{Pic}(C) = 1$ is important in deducing an analogon of the Noether–Skolem theorem that connects the Brauer group to Galois cohomology. However it will show that this condition is not really necessary at least if one allows the use of generalized crossed products (strongly graded rings), instead of the usual crossed products. Along the way we mention two criteria allowing to recognize Azumaya algebras.

Let M be an (A, C)-bimodule. A derivation $g : A \to M$ is a C-linear map satisfying $g(ab) = ag(b) + g(a)b$ for $a, b \in A$. We say that g is an inner derivation whenever there is an $m \in M$ such that $g(a) = am - ma$ for all a in A. By $Z^1_C(A, M)$ we denote the derivations of A into M equipped with the C-module structure induced by $\mathrm{Hom}_C(A, M)$. By $B^1_C(A, M)$ we denote the inner derivations in $Z^1_C(A, M)$ and $H^1_C(A, M) = Z^1_C(A, M)/B^1_C(A, M)$. The latter is the first Hochschild cohomology module of A with coefficients in M.

THEOREM 3.6. *A is separable over C if and only if $H^1_C(A, M) = 0$ for every (A, C)-bimodule M.*

The following criterion stems from G. Azumaya's original approach with restricted attention to C-algebras A that are free C-modules of finite rank.

THEOREM 3.7. *Let A be a faithful C-algebra that is a free C-module of finite rank. The C-algebra A is an Azumaya algebra if and only if the matrix $(u_i u_j)_{ij}$ is invertible in $M_n(A)$, where $\{u_i\}_i$ is a basis for A over C.*

We assume some knowledge about separable and Galois extensions in the class of commutative rings; the basic account given in the book of F. De Meyer and F. Ingraham [168] is certainly sufficient. An extension $C \subset D$ of commutative rings that is separable and such that $C = D^G = \{d \in D,\ d^g = d \text{ for all } g \in G\}$ for a finite group G of C-automorphisms of D with the property that for every maximal ideal W of D and $\sigma \in G$, σ does not reduce to the identity modulo w, is a Galois extension with group G.

If the finite group G acts on an Abelian group X, the operation of which we also denote multiplicatively, then we define:

$$Z^1(G, X) = \{f : G \to X,\ f(\sigma\tau) = f(\sigma) f(\tau)^\sigma \text{ for } \sigma, \tau \in G\},$$

$$\begin{aligned} B^1(G,X) &= \{f \in Z^1(G,X),\ f(\sigma) = \sigma(x)x^{-1} \text{ for some } x \in X,\ \text{all } \sigma \in G\},\\ Z^2(G,X) &= \{f : G \times F \to X,\ f(\sigma\tau,\rho)f(\sigma,\tau) = f(\sigma,\tau\rho)f^{\sigma}(\tau,\rho),\\ &\qquad \text{all } \sigma,\tau,\rho \in G\},\\ B^2(G,X) &= \{f \in Z^2(G,X),\ f(\sigma,\tau) = g(\sigma)g(\tau)^{\sigma}g(\sigma\tau)^{-1},\ \text{all } \sigma,\tau \in G,\\ &\qquad \text{for some map } g : G \to X\}. \end{aligned}$$

In a similar way one can define $Z^i(G,X)$, $B^i(G,X)$ for $i > 2$ and we put: $H^i(G,X) = Z^i(G,X)/B^i(G,X)$.

The Z^i are called the i-cocycles, B^i consists of i-coboundaries and H^i is called the i-th cohomology group (of G in X). For an i-cocycle f we will write $[f]$ for its class in $H^i(G,X)$.

THEOREM 3.8. *Let D be a (commutative) Galois extension of C with Galois group G, then there exists a long exact sequence of Abelian groups*:

$$\begin{aligned} 0 &\to H^1(G,U(D)) \to \mathrm{Pic}(C) \to \mathrm{Pic}(D)\\ &\to H^2(G,U(D)) \to \mathrm{Br}(D/C) \to H^1(G,\mathrm{Pic}(D)) \to H^3(G,U/D), \end{aligned}$$

where $\mathrm{Br}(D/C)$ *is the subgroup in* $\mathrm{Br}(C)$ *consisting of those classes of Azumaya algebra over C split by the extension D of C.*

COROLLARY 3.9 (Hilbert's theorem 90). *If* $\mathrm{Pic}(C) = 0$ *then* $H^1(G,U(D)) = 0$.

COROLLARY 3.10 (The crossed product theorem). *If* $\mathrm{Pic}(D) = 0$ *then* $\mathrm{Br}(D/C) = H^2(G,U(D))$.

In fact the condition $\mathrm{Pic}(D) = 0$ is not really essential in Corollary 3.10 that is if one allows the replacement of crossed products by so-called generalized crossed products or strongly graded rings in the sense of E. Dade [150] or C. Năstăsescu and F. Van Oystaeyen [399]. When a ring R has a decomposition $R = \bigoplus_{\sigma\in G} R_\sigma$, where each R_σ is an additive subgroup of R, such that $R_\sigma T_\tau \subset R_{\sigma\tau}$ for all $\sigma,\tau \in G$, then R is said to be G-graded. When $R_\sigma R_\tau = R_{\sigma\tau}$ for all $\sigma,\tau \in G$ then R is said to be strongly-graded. It is clear from $R_\sigma R_{\sigma^{-1}} = R_e$, where e is the neutral element of G, that R_σ is an invertible R_e-bimodule for every $\sigma \in G$, and then $R_\sigma R_\tau = R_{\sigma\tau}$ is given by a system $\{f(\sigma,\tau),\ \sigma,\tau \in G\}$ of R_e-bimodule isomorphisms: $f(\sigma,\tau) : R_\sigma \otimes R_\tau \to R_{\sigma\tau}$, $a_\sigma \otimes b_\tau \mapsto a_\sigma b_\tau$. That f defines a 2-cocycle follows from the associativity diagram:

$$\begin{tikzcd} R_\sigma \otimes R_\tau \otimes R_\gamma \arrow[r, "f(\sigma,\tau)\otimes 1_\gamma"'] \arrow[d, "1_\sigma\otimes f(\tau,\gamma)"'] & R_{\sigma\tau} \otimes R_\gamma \arrow[d, "f(\sigma,\tau,\gamma)"] \\ R_\sigma \otimes R_{\tau\gamma} \arrow[r, "f(\sigma,\tau\gamma)"'] & R_{\sigma\tau\gamma} \end{tikzcd}$$

On the other hand $\sigma \mapsto [R_\sigma]$ defines a group homomorphism $\psi : G \to \mathrm{Pic}(R_e)$ and the data ψ together with the system $\{f(\sigma,\tau),\ \sigma,\tau \in G\}$ determines the graded ring structure of R completely.

COROLLARY 3.11. *If A is an Azumaya algebra over C containing the Galois extension D of C with Galois group G as a maximal commutative subring then A is G-strongly graded over D.*

Of course the R_e-bimodule isomorphism $f(\sigma,\tau): R_\sigma \otimes R_\tau \to R_{\sigma\tau}$ corresponds to a unit in $Z(R_e)$, so in the situation of Corollary 3.11 the class $[A]$ is determined by an element $\psi \in H^1(G, \mathrm{Pic}(D))$ and an element of $H^2(G, U(D))$. There are further cohomological interpretations but we prefer to present these in the framework of $\mathbb{Z}$-graded Azumaya algebras. In this way we will of course recover the ungraded theory by considering the trivial $\mathbb{Z}$-gradation but the gain is that we obtain immediate connections with the theory of the Brauer group of a projective scheme.

4. The graded Brauer group

After A. Grothendieck's "Dix Exposés" appeared in 1968, cf. [256], the Brauer group became an object of geometric interest, however one must recognize that pure geometrical applications of the Brauer group like M. Artin and D. Mumford's use of it in discussing examples of non-rational unirational varieties, cf. [41], are the exception not the rule. The cohomological interpretation of the geometric Brauer group (of a scheme) should consist in comparing it to the torsion of the second étale cohomology group and again the state of the art is not the most desirable one. Of course, there is O. Gabber's thesis, contained in [237], and so for projective varieties covered by two affines we do have satisfactory answers to the basic questions, but in full generality there is left much to be desired. It was J. Murre who raised the question of finding out the nature of the algebraic equivalent of the Brauer group of a projective scheme, the Brauer group of an affine scheme being the reasonably well-known Brauer group of its global sections ring. Starting from a positively graded ring, the first idea could be that the Brauer group of its projective scheme could be close to a certain graded version of the Brauer group, so we set out to study a graded Brauer group of general graded rings finding a theory with some interest in its own right but finally being frustrated (as far as the original question is concerned) by the fact that in the positively graded case our object reduced to the Brauer group of the degree zero part and this could not correspond to the geometrical situation. Finally it turned out that a combination of graded and "relative" techniques was necessary to obtain the algebraic equivalent of the Brauer group of Proj. Using a suitable affine covering one may describe the geometric Brauer group by Brauer groups of certain rings. The latter approach may make the use of étale cohomologic unnecessary but meanwhile the theory has not been carried much further and perhaps the most important job remains to be done.

Let R be a commutative $\mathbb{Z}$-graded ring. $R = \bigoplus_{n\in\mathbb{Z}} R_n$. We write $\mathbf{P}(R)$ for the category of all finitely generated projective R-modules with R-module homomorphisms and with product in fact being given by the direct sum $\oplus$. Then $\mathbf{FP}(R)$ is the category of all faithfully projective R-modules but with product being given by the tensor product $\otimes_R -$. Let $\mathbf{Pic}(R)$ equal to the category of invertible R-modules, the product again being given by the tensor product. Similarly we denote by $\mathbf{P}^g(R)$, $\mathbf{F}P^g(R)$ and $\mathbf{Pic}^g(R)$ the corresponding categories within the category of graded R-modules, $\mathbf{R}^g$-mod say, with R-module morphisms, but we write $\mathbf{P}_g(R)$, $\mathbf{FP}_g(R)$, $\mathbf{Pic}_g(R)$ for the categories consisting of the same objects but taking only graded R-module morphisms for the morphisms, i.e. working in $\mathbf{R}$-gr. Put $\mathrm{Pic}^g(R) = K_0\mathbf{Pic}^g(R)$ and $\mathrm{Pic}_g(R) = K_0\mathbf{Pic}_g(R)$; then $\mathrm{Pic}^g(R)$ embeds in $\mathrm{Pic}(R)$ and

$$K_1\mathbf{Pic}^g(R) = K_1\mathbf{Pic}(R) = U(R),$$

the unit group of R. On the other hand

$$K_1\mathbf{Pic}_g(R) = U_0(R) = U(R)_0,$$

the unit group of degree zero. Unlike the ungraded case there may be gr-semilocal rings having a nontrivial $\mathrm{Pic}^g(R)$, however in enough decent cases this graded-anomaly does not appear, e.g., when R is gr-local, strongly graded (that is $R_n R_m = R_{n+m}$, $n, m \in \mathbb{Z}$), positively graded or such that $J(R) = J^g(R)$ where the latter is the graded version of the Jacobson radical, that is always a graded ideal but it may be considerably larger than $J(R)$ (also a graded ideal) in general. Let $g(R)$ be the group of graded ring isomorphism classes of invertible graded R-modules that are equal to R as an R-module and that have a ring structure in this way, i.e. $g(R)$ is the group of all possible gradations on R as an R-module where the group law is induced by the tensor product.

PROPOSITION 4.1. *There is an exact sequence of Abelian groups*:

$$1 \to U_0(R) \to U(R) \to g(R) \to \mathrm{Pic}_g(R) \to \mathrm{Pic}^g(R) \to 1.$$

We write $H^0(R)$ for the additive group of continuous functions $\mathrm{Spec}\, R \to \mathbb{Z}$. Since every idempotent element of R is homogeneous it is clear that $H^0(R)$ is also the group of continuous functions from $\mathrm{Spec}^g\, R \to \mathbb{Z}$.

THEOREM 4.2. *There is a natural embedding $\mu : H^0(R) \to g(R)$ and when R is reduced it is in fact an isomorphism.*

COROLLARY 4.3. *If R is also connected then all units are homogeneous.*

Passing via a graded cancellation property, i.e. if P, Q in $\mathbf{FP}^g(R)$ are such that $[P] = [Q]$ in $K_0\mathbf{FP}^g(R)$ then $P^n \simeq Q^n$ for some $n > 0$, we may obtain a graded version of the stability theorem (Theorem III.3.1.4 in [59]). We do not drift further into graded K-theory but let us just point out the existence of [588].

A graded R-algebra is called a gr-Azumaya algebra if it satisfies one of the equivalent conditions in the following.

PROPOSITION 4.4. *For a graded R-algebra A the following assertions are equivalent*:

(1) *A is an (R)-Azumaya algebra.*

(2) *A is a faithfully projective R-module such that $A^e = A \otimes A^0 \cong \mathrm{END}_R(A)$ where $\mathrm{END}_R(A)$ consists of combinations of endomorphisms of arbitrary degrees and the isomorphism is a graded one.*

(3) *The functors $N \to A \otimes N$ and $M \to M^A = \{m \in M;\ am = ma,\ a \in A\}$ yield an equivalence of the categories R-gr and A^e-gr.*

(4) *There exists a graded R-algebra B and a graded faithfully projective R-module P such that $A \otimes_R B \cong_g \mathrm{END}_R(P)$ as graded algebras.*

Let $\mathbf{Az}_g(R)$ be the category of graded R-Azumaya algebras with graded R-algebra homomorphisms. The functor $\mathrm{END}_R : \mathbf{FP}_g(R) \to \mathbf{Az}_g(R)$ induces a map

$$K_0\mathbf{FP}_g(R) \to K_0Az_g(R)$$

and the graded Brauer group $\mathrm{Br}_g(R)$ is just the cokernel of

$$K_0\mathrm{END}_R : K_0\mathbf{FP}_g(R) \to K_0\mathbf{Az}(R).$$

Hence the graded Brauer group may also be defined by introducing gr-equivalence $\sim_g$ as follows $A \sim_g B$ for graded R-Azumaya algebras if there exist graded faithfully projective R-modules P, Q such that the graded R-algebras $A \otimes_R \mathrm{END}_R(P)$ and $B \otimes_R \mathrm{END}_R(Q)$ are graded isomorphic. The gr-equivalence class of A is denoted by $[A]$ and $[A].[B] = [A \otimes_R B]$, $[A]^{-1} = [A^0]$. It is easily checked that $[A] = [R]$ exactly then when $A \cong_g \mathrm{END}_R(E)$ for some graded faithfully projective R-module E.

THEOREM 4.5 (Graded version of H. Bass's exact sequence). *The sequence of functors*

$$\mathbf{Pic}_g(R) \xrightarrow{I} \mathbf{FP}_g(R) \xrightarrow[\mathrm{END}]{} \mathbf{Az}_g(R)$$

of categories with product, induces the following exact sequence:

$$\begin{aligned} U_0(R) = K_1\mathbf{Pic}_g(R) &\xrightarrow[K_1I]{} K_1\mathbf{FP}_g(R) \xrightarrow[K_1\mathrm{END}]{} K_1\mathbf{Az}_g(R) \to \mathrm{Pic}_g(R) \to \\ &\xrightarrow[K_0I]{} K_0\mathbf{FP}_g(R) \xrightarrow[K_0\mathrm{END}]{} K_0\mathbf{Az}_g(R) \to \mathrm{Br}_g(R) \to 1. \end{aligned}$$

THEOREM 4.6 (Second graded version of H. Bass's exact sequence). *Consider the following commutative diagram of functors between categories with product*:

$$\begin{array}{ccccc} \mathbf{Pic}^g(R) & \xrightarrow[I^g]{} & \mathbf{FP}^g(R) & \xrightarrow{\mathrm{END}} & Az^g(R) \\ \downarrow{\scriptstyle U_0} & & \downarrow{\scriptstyle U_1} & & \downarrow{\scriptstyle U_2} \\ \mathbf{Pic}(R) & \xrightarrow[I]{} & \mathbf{FP}(R) & \xrightarrow[\mathrm{END}]{} & \mathbf{Az}(R) \end{array}$$

Then this induces the following commutative diagram of exact sequences:

$$\begin{tikzcd}
U(R) = K_1\mathbf{Pic}^g(R) \arrow[r] \arrow[d, equal] & K_1\mathbf{FP}^g(R) \arrow[r] \arrow[d, equal] & K_1\mathbf{Az}^g(R) \arrow[r] \arrow[d, equal] & {} \\
K_1\mathbf{Pic}(R) \arrow[r] & K_1\mathbf{FP}(R) \arrow[r] & K^1\mathbf{Az}(R) \arrow[r] & {}
\end{tikzcd}$$

$$\begin{tikzcd}
{} \arrow[r] & \mathrm{Pic}^g(R) \arrow[r, "K_0 I^g"'] \arrow[d] & K_0\mathbf{FP}^g(R) \arrow[r] \arrow[d] & K_0\mathbf{Az}^g(R) \arrow[r] \arrow[d] & \mathrm{Br}^g(R) \arrow[r] \arrow[d] & 1 \\
{} \arrow[r] & \mathrm{Pic}(R) \arrow[r, "K_0 I"'] & K_0\mathbf{FP}(R) \arrow[r] & K_0\mathbf{Az}(R) \arrow[r] & \mathrm{Br}(R) \arrow[r] & 1
\end{tikzcd}$$

where $\mathrm{Br}^g(R) = K_0\mathrm{END}$, $\mathrm{Ker}(K_1 I) = U(R)_{\mathrm{tors}}$, $\mathrm{Ker}(K_0 I^g) = \mathrm{Pic}^g(R)_{\mathrm{tors}}$ *and* $\mathrm{Ker}(K_0 I) = \mathrm{Pic}(R)_{\mathrm{tors}}$.

In [99] we also obtained a relation between $\mathrm{Br}_g(R)$ and $\mathrm{Br}(R)$ in terms of a graded version of the so-called Brauer-class group, denoted by $GR(R)$, fitting in the exact sequence:

$$\mathrm{Pic}_g(R) \to \mathrm{Pic}(R) \to GR(R) \to \mathrm{Br}_g(R) \to \mathrm{Br}(R).$$

Avoiding too much detail we just mention the two graded versions of the "Chase–Rosenberg Exact Sequence" in one theorem:

THEOREM 4.7. *Let S be a commutative graded R-algebra that is faithfully projective as an R-module. Then we have the following two exact sequences*:

$$\begin{aligned}
\text{(a) } 1 &\to H^0(S/R, U_0) \to U_0(R) \to 1 \\
&\to H^1(S/R, U_0) \to \mathrm{Pic}_g(R) \to H^0(S/R, \mathrm{Pic}_g) \\
&\to H^2(S/R, U_0) \to \mathrm{Br}_g(S/R) \to H^1(S/R, \mathrm{Pic}_g) \to H^3(S/R, U_0),
\end{aligned}$$

$$\begin{aligned}
\text{(b) } 1 &\to H^0_{\mathrm{gr}}(S/R, U) \to U(R) \to g(R) \\
&\to H^1_{\mathrm{gr}}(S/R, U) \to \mathrm{Pic}^g(R) \to H^0(S/R, \mathrm{Pic}^g) \\
&\to H^2_{\mathrm{gr}}(S/R, U) \to \mathrm{Br}_g(S/R) \to H^1(S/R, \mathrm{Pic}^g) \to H^3_{\mathrm{gr}}(S/R, 0),
\end{aligned}$$

where $H^n_{\mathrm{gr}}(S/R, U) = H^n_{A_2}(A_1)$ and A_1, resp. A_2, is the Amitsur complex of the functor U, resp. $g(-)$.

COROLLARY 4.8. (1) *The crossed product theorems. Let S be a commutative graded R-algebra that is faithfully projective as an R-module then*:

(a) *If* $\mathrm{Pic}_g(S) = \mathrm{Pic}_g(S \otimes S) = 1$, *then* $\mathrm{Br}_g(S/R) \cong H^2(S/R, U_0)$,

(b) *If* $\mathrm{Pic}^g(S) = \mathrm{Pic}^g(S \otimes S) = 1$, *then* $\mathrm{Br}_g(S/R) \cong H^2_{\mathrm{gr}}(S/R, U)$.

(2) *If R is reduced and S a graded Galois extension of R such that S has no non-trivial idempotent elements then, if* $\mathrm{Pic}^g(S) = 1$ *then* $\mathrm{Ker}(H^2(G, U(S)) \to H^2(G, \mathbb{Z})) = \mathrm{Br}_g(S/R)$.

For graded Brauer groups of particular rings we refer to [99]. Let us just mention the following characterization of perfect fields in terms of graded Brauer groups:

THEOREM 4.9. *Let k be a field. Then k is perfect if and only if*

$$\operatorname{Br} k[T, T^{-1}] = \bigcup_{\deg T} \operatorname{Br}_g k[T, T^{-1}],$$

where the latter means that for each choice of $\deg T \in \mathbb{N}$ *we have identified* $\operatorname{Br}_g k[T, T^{-1}]$ *as a subgroup of* $\operatorname{Br} k[T, T^{-1}]$.

It may also be mentioned that every graded Azumaya algebra over a gr-local ring is equivalent to a generalized crossed product (that is a strongly graded ring for the Galois group of some graded Galois extension S/R splitting A). For $n \in \mathbb{N}$ we may define a graded ring $R^{(n)}$ by putting $R^{(n)}_{ni} = R_i$ for $i \in \mathbb{N}$, and as an extension of the techniques used in the proof of Theorem 4.9 we mention (because it has some use in the geometrical situation).

THEOREM 4.10. *Let R be a quasistrongly graded* gr*-local Krull domain containing a field of characteristic zero, then* $\operatorname{Br}(R) = U_n \operatorname{Br}_g(R^{(n)})$; *in particular* every *Azumaya algebra over R may be split by a graded Galois extension.*

NOTE 4.11. For a graded Krull domain R, $\operatorname{Br}_g(R) \hookrightarrow \operatorname{Br}(R)$ always holds. The quasistrongly graded property just means that $R^{(d)}$ is strongly graded for some $d \in \mathbb{N}$, or $Rd_n R\text{-}dn = R_0$ for all n for some d.

As pointed out before, for a positive $R = R_0 \oplus R_1 \oplus \cdots$, $R_+ = R_1 \oplus R_2 \oplus \cdots$, we have that $\operatorname{Br}_g(R) = \operatorname{Br}_g(R/R_+) \cong \operatorname{Br}_g(R_0) = \operatorname{Br}(R_0)$.

For a treatment of étale cohomology of graded rings we refer to [99, Ch. V]. A very interesting class of graded commutative rings where the graded Brauer group could be of some use is the class of regular graded rings. Obviously at this point we enter again into a geometric framework so we embed the treatment of this class of graded domains in our next section.

5. Separable algebras over rings regular in codimension n

In this section we restrict attention to Noetherian integrally closed domains R; when considering properties of a graded nature we assume that R is $\mathbb{Z}$-graded. We say that R is regular in codimension n, $n \in \mathbb{N}$, if for all prime ideals P of R of height less than n, the ring R_P is a regular local ring. Put $X^{(n)}(R) = \{P \in \operatorname{Spec} R, ht(P) \leqslant n\}$; so we obtain a filtration: $X^{(1)}(R) \subset \cdots \subset X^{(n)}(R) \subset \cdots \subset X^{(d)}(R) = X = \operatorname{Spec} R$, where $d = \operatorname{Kdim} R$, the Krull dimension of R. Put $X^i(R) = X^{(i)}(R) - X^{(i-1)}(R)$. It is obvious that any Zariski open set of X that contains $X^i(R)$ must also contain $X^{(i)}(R)$.

For n such that R is regular in codimension n we define

$$\operatorname{Br}^{(n)}(R) = \bigcap \{\operatorname{Br}(R_p), P \in X^{(n)}(R)\}.$$

Any Noetherian integrally closed domain is necessarily regular in codimension 1, hence we may always define

$$\beta(R) = \mathrm{Br}^{(1)}(R) = \bigcap \{\mathrm{Br}(R_p),\ P \in X^{(1)}(R)\}.$$

Note that all intersections written above are taken in $\mathrm{Br}(K)$, where K is the field of fractions of R, up to the obvious identification of $\mathrm{Br}(R_p)$ and its image in $\mathrm{Br}(K)$. When R is regular of dimension d, $\mathrm{Br}^{(d)}(R) = \mathrm{Br}(R)$. For a regular domain the filtrations $X^{(n)}(R)$ correspond to a filtration:

$$\mathrm{Br}(R) = \mathrm{Br}^{(d)}(R) \subset \mathrm{Br}^{(d-1)}(R) \subset \cdots \subset \beta(R) \subset \mathrm{Br}(K).$$

When R is also regular in codimension n, with $n < d$, we obtain:

$$\mathrm{Br}^{(n)}(R) \subset \cdots \subset \beta(R) \subset \mathrm{Br}(K).$$

When R is not regular in codimension n then $\mathrm{Br}^{(n)}(R)$ can still be defined (using relative Brauer groups in the sense of [609]) and we obtain a general sequence of group morphisms:

$$\begin{aligned}\mathrm{Br}(R) = \mathrm{Br}^{(d)}(R) &\to \mathrm{Br}^{(d-1)}(R) \to \cdots \to \mathrm{Br}^{(n)}(R) \hookrightarrow \mathrm{Br}^{(n-1)} \\ &\hookrightarrow \cdots \hookrightarrow \beta(R) \hookrightarrow \mathrm{Br}(K).\end{aligned}$$

LEMMA 5.1. *Let R be a ring as before regular in codimension n and let A be an R-algebra representing an element of* $\mathrm{Br}^{(n)}(R)$. *Then A is a maximal R-order in its classical ring of quotients $Q(A)$ which is a representative of the image of* $[A]$ *in* $\mathrm{Br}(K)$.

The group $\beta(R)$ is called the reflexive Brauer group of R, cf. S. Yuan [673], and an R-algebra A representing $\alpha \in \beta(R)$ is called a *reflexive Azumaya algebra*. For a ring R that is regular in codimension n, every $\gamma \in \mathrm{Br}^{(m)}(R)$, with $m \leqslant n$, may be represented by a reflexive Azumaya algebra Γ having some extra good properties at primes of codimension between 1 and m. We recall the sequence defining the Brauer class group $\mathrm{BCl}(R)$.

LEMMA 5.2. *The following sequence of Abelian groups is exact*:

$$0 \to \mathrm{Pic}(R) \to \mathrm{Cl}(R) \to \mathrm{BCl}(R) \to \mathrm{Br}(R) \to \beta(R).$$

Now consider the case where R is $\mathbb{Z}$-graded and let $X_g^{(n)}(R)$ be the subset of $X^{(n)}(R)$ consisting of graded prime ideals of height n. The following easy lemma is essential because it shows that we may replace *ht* for a prime ideal that happens to be graded by the graded version gr-ht.

LEMMA 5.3. *Let R be any $\mathbb{Z}$-graded Noetherian commutative ring. If P is a graded prime ideal of R with* $\mathrm{ht}(P) = n$ *then there exists a chain of graded prime ideal of R*:

$$P_0 \underset{\neq}{\subset} P_1 \underset{\neq}{\subset} \cdots \underset{\neq}{\subset} P_n = P.$$

As before we let K be the field of fractions of R and we sometimes consider the gr-field of fractions K^g obtained by inversing only homogeneous elements of R. Put $\mathrm{Br}_g^{(n)}(R) = \bigcap\{\mathrm{Br}_g Q_P^g(R),\ P \in X_g^{(m)}(R)\}$.

LEMMA 5.4. *Let R' be regular in codimension n and pick $m \leqslant n$, then $\mathrm{Br}_g^{(m)}(R) \subset \mathrm{Br}^{(m)}(R)$ and we obtain a commutative diagram of inclusions*:

$$\begin{tikzcd}
\mathrm{Br}^{(n)}(R) \subset \cdots \subset & \beta(R)/\mathrm{Br}^{(1)}(R) \subset \mathrm{Br}(K) \\
\mathrm{Br}_g^{(n)}(R) \subset \cdots \subset \arrow[u] & \beta_g(R) = \mathrm{Br}_g^{(1)}(R) \subset \mathrm{Br}_g(K^g) \subset \mathrm{Br}(K^g) \arrow[u]
\end{tikzcd}$$

In view of the ad-hoc definition of $\mathrm{Br}_g^{(n)}(R)$ the following proposition is not a complete triviality even if it expresses something one is willing to accept without proof.

PROPOSITION 5.5. *Let R be regular in codimension n and take $m \leqslant n$. if A is a $\mathbb{Z}$-graded R-algebra representing an $\alpha \in \mathrm{Br}^{(m)}(R)$ then $\alpha \in \mathrm{Br}_g^{(m)}(R)$.*

It is possible to describe $\beta(R)$ in terms of étale cohomology over the set of regular points X_{reg} in X, i.e. over the normal affine variety with coordinate ring R. In case $X = X_{\mathrm{reg}}$ the situation is easy enough.

THEOREM 5.6 [99]. *If R is a regular affine domain over a field k then $\beta(R) = \mathrm{Br}(R)$.*

THEOREM 5.7 [359]. *If R is a normal affine, i.e. the coordinate ring of a normal affine variety X over an algebraically closed field*

(1) $\beta(R) \cong H^2_{\text{ét}}(X_{\mathrm{reg}}, U)$; *note that since all appearing schemes are regular the cohomology groups mentioned are torsion.*
(2) $\beta(R) = \mathrm{Br}(V)$ *for some open subset V in X_{reg}.*
(3) *Put $Z = X - X_{\mathrm{reg}}$ and let $H^i_Z(X, U)$ be the cohomology groups with support on Z, cf. Milne* [384], *then we obtain a long exact sequence in étale cohomology.*

$$\begin{tikzcd}
H^1_Z(X,U) \arrow[d] & & & & H^3_Z(X,U) \\
H^1(X,U) \arrow[r] & H^1(X_{\mathrm{reg}},U) \arrow[r] & H^2_Z(X,U) \arrow[r] & H^2(X,U) \arrow[r] & H^2(X_{\mathrm{reg}},U) \arrow[u] \\
\mathrm{Pic}(R) \arrow[r] \arrow[u, "\cong"'] & \mathrm{Cl}(R) \arrow[r] \arrow[u, "\cong"'] & \mathrm{BCl}(R) \arrow[r] \arrow[u] & \mathrm{Br}(R) \arrow[r] \arrow[u] & \beta(R) \arrow[u]
\end{tikzcd}$$

OBSERVATIONS 5.8. In view of the earlier remarks we know that for a $\mathbb{Z}$-graded ring R regular in codimension n, any $\alpha \in \mathrm{Br}_g^{(m)}(R)$ may be represented by a $\mathbb{Z}$-graded maximal R-order A (for any $m \leqslant n$). Moreover, if R is a $\mathbb{Z}$-graded Noetherian integrally closed domain then an R-algebra A represents an element $\alpha \in \beta_g(R)$ if and only if for all $P \in X_g^{(1)}(R)$ we have that the graded localization of A at P, i.e. $Q_P^g(A)$, is an Azumaya-algebra over $Q_P^g(R)$.

For low-dimensional situations the following result of Auslander and Goldman [50] is very helpful.

LEMMA 5.9. *If R is a Noetherian domain of global dimension at most 2, then every finitely generated reflexive R-module is projective.*

COROLLARY 5.10. *If R is a Noetherian integrally closed domain of dimension at most two then* $\mathrm{Br}(R) = \beta(R)$.

In the presence of a positive $\mathbb{Z}$-gradation we are able to treat the case that actually corresponds to the calculation of the Brauer group of a projective scheme of dimension at most two; we first provide a purely algebraic statement and elaborate on the geometry afterwards.

THEOREM 5.11. *Let R be positively graded, with $R_0 = k$ being a field, and assume that R is a Noetherian integrally closed domain of* gr-*global dimension at most* 3. *If the R-algebra A represents $\alpha \in \beta_g(R)$ then for all $P \in \mathrm{Proj}(R) = \{P \in \mathrm{Spec}(R), P$ is graded and $P \not\supset R_+\}$ where $R_+ = \bigoplus_{n>0} R_n$, we have that $Q_P^g(A)$ is an Azumaya algebra.*

In general the Brauer group of a locally ringed space X may be defined in terms of locally separable sheaves of algebras. In case $X = \operatorname{Spec} R$ we let $\widetilde{R}$ be the structure sheaf of R defined over the topological space X. The locally separable sheaves then turn out to have stalks at $P \in X$ that are Azumaya algebras over the stalks $\widetilde{R}_p = R_p$. If R has a positive gradation then the space $\mathrm{Proj}(R) = Y$ may be viewed as the set consisting of the graded prime ideals in the open set $X(R_+)$ of $X = \mathrm{Spec}(R)$, endowed with the topology induced by the Zariski topology of X. The graded structure sheaf on Y is obtained by taking the graded ring of fractions $Q_f^g(R)$ as the ring of sections over $Y(f) \subset Y$ for each homogeneous $0 \neq f \in R$. The (projective) structure sheaf of Y is obtained by taking parts of degree zero in the graded structure sheaf and we denote it by $\mathbf{O}_Y$, writing $\mathbf{O}_Y^g$ for the graded structure sheaf. A locally separable sheaf of graded $\mathbf{O}_Y^g$-algebras will have a graded Azumaya algebra for the stalk at each $P \in Y$. Rephrasing Theorem 5.11 in this more geometric terminology learns that, if $\dim(R) \leqslant 3$, the graded structure sheaf as a reflexive graded Azumaya algebra A corresponds to an element of $\mathrm{Br}(\mathrm{Proj}(R))$. Note that any $\mathbb{Z}$-graded R-algebra A such that for all $P \in \mathrm{Proj}(R)$ we have that the graded stalk $Q_P^g(A)$ is an Azumaya algebra, is necessary a reflexive Azumaya algebra because $X_g^{(1)}(R) \subset \mathrm{Proj}(R)$ unless $R = k[X]$ a case that can be excluded for obvious reasons. In order to check that $\mathrm{Br}(\mathrm{Proj}(R))$, defined in terms of locally separable sheaves, is in fact isomorphic to $\beta_g(R)$, defined in terms of reflexive graded Azumaya algebras, one has still to verify that the similarity notions introduced in the definitions of both generalized Brauer groups do correspond well, but that is only a technicality, cf. [99]. So, for a connected regular projective variety Y over a field k, of dimension at most two (e.g., curves or surfaces) we find that

$\mathrm{Br}(Y) = \beta_g(\Gamma_*(Y))$ where $\Gamma(Y)$ is the homogeneous coordinate ring. Of course, for an affine scheme $X = \mathrm{Spec}\, R$ we have $\mathrm{Br}(X) = \mathrm{Br}(R)$. As an example of some interest one may consider the ring $R = \mathbb{C}[X, Y, Z]/(X^2 - YZ)$; then $\mathrm{Cl}(R) = \mathbb{Z}/2\mathbb{Z}$ and R describes an affine cone generated by a ruling P. The algebra $\Lambda = \mathrm{END}_R(R \oplus P)$ is a reflexive Azumaya algebra but not Azumaya. Nevertheless in the case $\dim R \leqslant 2$ we know that a reflexive module over a regular local ring is free and hence a reflexive Azumaya algebra is "Azumaya on the open set of regular points". So in this case $\beta(R) = \mathrm{Br}(X_{\mathrm{reg}})$ follows and $\mathrm{Br}(R) \neq \beta(R)$.

As we will point out at the end of the section, one can still say something about the algebraic equivalent of $\mathrm{Br}(X)$ for higher dimensional varieties (schemes) X. First let us concentrate on the case where X is a curve! For any commutative domain we write $\overline{S}$ for its integral closure in the field of fractions.

In general, if S is any positively graded Noetherian domain (as usual we assume that $S = S_0[S_1]$) then $\mathrm{Proj}(\overline{S})$ is the normalization of $\mathrm{Proj}(S)$ and the inclusion $S \hookrightarrow \overline{S}$ induces a scheme morphism $\mathrm{Proj}(\overline{S}) \to \mathrm{Proj}(S)$.

Now let R be a positively graded k-algebra with $R_0 = k$, where k is a fixed field, and assume that R is generated as a k-algebra by finitely many elements of R_1. Let c be the conductor ideal of $\overline{R}$ in R, $c = \{x \in \overline{R},\ xr \in R \text{ for all } r \in \overline{R}\}$. Assume moreover that $Y = \mathrm{Proj}(R)$ is a curve, then $\overline{Y} = \mathrm{Proj}(\overline{R})$ is its normalization and we let V, resp. $\overline{V}$, be the closed subscheme of Y, resp. $\overline{Y}$, determined by the ideal c. It is a classical result that $\mathrm{Pic}(\overline{V}) = 0$, cf. R. Hartshorne [263].

THEOREM 5.12. *The following sequence of Abelian groups is exact*:

$$0 \to \mathrm{Br}(Y) \to \mathrm{Br}(\overline{Y}) \oplus \mathrm{Br}(V) \to \mathrm{Br}(\overline{V}).$$

COROLLARY 5.13.
(1) $\mathrm{Br}(Y) = \beta_g(\Gamma_*(Y))$.
(2) *In case Y is a regular curve we obtain* $\mathrm{Br}(Y) = \mathrm{Br}_g(R)$.

After the foregoing we have to focus on singular curves. For $p \in V$, resp. $q \in \overline{V}$, let $K_y(p)$, resp. $K_{\overline{y}}(q)$, be the residue field at p in Y, resp. q in $\overline{Y}$.

THEOREM 5.14 [609]. *Suppose that $Y = \mathrm{Proj}(R)$ is a connected projective curve and let c, V and $\overline{V}$ be as before, then there is an exact sequence of Abelian groups*:

$$0 \to \mathrm{Br}(Y) \to \beta_g(R) \oplus \Big(\bigoplus_{p \in V} \mathrm{Br}\big(K_y(p)\big)\Big) \to \bigoplus_{q \in \overline{V}} \mathrm{Br}\big(K_{\overline{y}}(q)\big).$$

In the case $\dim Y = d$ one might hope to establish that $\mathrm{Br}(Y) \cong \mathrm{Br}_g^{(m)}(\Gamma_*(R))$ whenever R is regular in codimension $m < d$; this seems to follow from Lemma 5.4 and Proposition 5.5 together with the obvious properties of locally separable sheaves, however nobody seems to have verified that the similarity relations used in defining $\mathrm{Br}(Y)$ fit the definitions of the right hand member. Similarly, a version of Lemma 5.2 may be obtained for $\mathrm{Br}^{(m)}(R)$

if one replaces $\mathrm{Cl}(R)$ by the generalized class group $\mathrm{Cl}(R)_{(n)}$ as introduced by Fossum and Claborn [139], $\beta(R)$ by $\mathrm{Br}^{(m)}(R)$ and $\mathrm{BCl}(R)$ by the corresponding $\mathrm{BCl}_{(n)}(R)$ in fact defined by that sequence. More serious problems may arise in Theorem 5.7 if one replaces $\beta(R)$ by $\mathrm{Br}^{(m)}(R)$ and X_{reg} by $X_{m\text{-reg}}$ consisting of points corresponding to a local ring only regular in codimension m; this has not been tried out. Probably the main reason why the $\mathrm{Br}^{(m)}(R)$ have not been studied so much may reside in the fact that $\mathrm{Br}(\mathrm{Proj}(R)) = \mathrm{Br}_g(R, \kappa_+)$ in the sense of [609].

If we consider only Noetherian domains R then a Gabriel filter $\mathcal{L}$ is given by a multiplicative set of ideals of R and the localization of any R-module M at $\mathcal{L}$ is then given by

$$\lim_{I \in \mathcal{L}} \mathrm{Hom}_R(I, \overline{M}) = Q_{\mathcal{L}}(M)$$

where $\overline{M} = M/t_{\mathcal{L}}M$, $t_{\mathcal{L}}M = \{m \in M, Im = 0 \text{ for some } I \in \mathcal{L}\}$. In general one would have to use more complicated arguments of abstract torsion theory, but for Noetherian domains the theory simplifies a lot. In this case every Gabriel filter $\mathcal{L}$ corresponds to a so-called generically closed subset $X(\mathcal{L}) \subset X = \mathrm{Spec}(R)$ by putting $P \in X(\mathcal{L})$ if and only if P does not contain an element of $\mathcal{L}$; in fact this is a bijective correspondence between generically closed subsets of X and Gabriel filters determined by a multiplicative set of ideals. A Zariski open set $X(I)$ corresponds to the Gabriel filter generated by the powers of I. The ring of sections over $X(\mathcal{L})$ is nothing but $Q_{\mathcal{L}}(R) = \bigcap_{P \in X(\mathcal{L})} R_P$.

The R-torsion R-modules, that is those $M \in R$-mod such that $t_{\mathcal{L}}M = M$, form a Serre subcategory $T_{\mathcal{L}}$ of R-mod and the quotient category $R\text{-mod}/T_{\mathcal{L}}$ is denoted by $(R, \mathcal{L})$-mod; it is the full subcategory of R-mod consisting of localized modules $Q_{\mathcal{L}}(M)$. The latter are $Q_{\mathcal{L}}(R)$-modules but $(R, \mathcal{L})\text{-mod} \neq Q_{\mathcal{L}}(R)\text{-mod}$ in general.

When R is $\mathbb{Z}$=-graded we may replace R-mod by the Grothendieck category of graded R-modules R-gr, consider multiplicative sets of graded ideals of R, $\mathcal{L}^g$ say, and redo the foregoing. In this way one defines $(\mathcal{R}, \mathcal{L}^g)$-gr consisting of graded $Q^g_{\mathcal{L}}(R)$-modules. If R is positively graded then the Gabriel filter generated by the powers of R_+ will be denoted by κ_+ and so we obtain the quotient category (R, κ_+)-gr. In the particular (geometric) case where $R = k[R_1]$, $k = R_0$, the finitely generated κ_+-torsion modules are exactly those of finite length. Now one may agree that $\mathrm{Proj}(R)$ is determined by the (quasi-) coherent sheaves of modules over it, so by Serre's global sections theorem it makes sense to identify the geometric object $\mathrm{Proj}(R)$ and the "categorical" object $(R, \kappa_+)^f$-gr, where f indicates that one starts everything by looking at finitely generated modules (but this is not really essential in the story). Applying this "philosophy" to the Brauer group one does expect something like $\mathrm{Br}(\mathrm{Proj}(R)) \cong \mathrm{Br}((R, \kappa_+)\text{-gr})$ and in fact the latter categorical object may be given a precise sense. In [609], $\mathrm{Br}_g(R, \kappa_+)$ and more generically, $\mathrm{Br}_g(R, \mathcal{L}^g)$ has been described in terms of graded relative Azumaya algebras. Roughly stated these algebras have nice properties at $P \in X(\mathcal{L}) \cap \mathrm{Proj}(R)$ but may be bad when localized outside $X(\mathcal{L})$.

PROPOSITION 5.15. *For* $Y = \mathrm{Proj}(R), \mathrm{Br}(Y) = \mathrm{Br}_g(R, \kappa_+)$.

Of course, now we have an algebraic description for $\mathrm{Br}(Y)$ in any dimension but as both objects are equally unknown in dimension strictly larger than two, that is all there is to that. Nevertheless, by passing to a finite affine covering of Y given by elements of degree one, say $x_1, \ldots, x_d$, one may describe $\mathrm{Br}_g(R, \kappa_+)$ by the usual Brauer groups $\mathrm{Br}((R_{x_i})_0) = \mathrm{Br}_g(R_{x_i})$ and perhaps there is more to be obtained from this approach. Let us point out a few special cases that may motivate this optimism. If $\mathcal{L}_n$ is the Gabriel filter corresponding to the the generically closed set $X^{(n)}(R)$ then $\mathrm{Br}(R, \mathcal{L}_n) = \mathrm{Br}^{(n)}(R)$ and $\mathrm{Cl}(R)_{(n)} = \mathrm{Pic}(R, \mathcal{L}_n)$ where $\mathrm{Pic}(R, \mathcal{L}_n)$ may be defined as $K_0\mathrm{Pic}((R, \mathcal{L}_n)\text{-mod})$. In particular $\beta(R) = \mathrm{Br}(R, \mathcal{L}_1)$ where $\mathcal{L}_1$ corresponds to the generically closed $X^{(1)}(R)$. In the presence of a $\mathbb{Z}$-gradation one arrives at: $\mathrm{Br}_g^{(n)}(R) = \mathrm{Br}_g(R, \mathcal{L}_n^g)$, where all notions have obtained their usual "graded" meaning. The facts claimed after Theorem 5.14 may be proved by using this interpretation of $\mathrm{Br}^{(n)}$ and $\mathrm{Cl}_{(n)}$ as "relative" Brauer, resp. class, group with respect to $\mathcal{L}_n$.

When trying to make a classification of Azumaya algebras over a given commutative ring, the step to the consideration of the Brauer group and its possible cohomological descriptions is a natural and a fruitful one. Depending on the area of origin of the commutative ring, e.g., a number-ring, a coordinate ring, a ring of continuous functions, ..., the theory of its Brauer group will definitely take on the flavour of number theory, algebraic geometry, ... We did not aim to provide a complete survey of Brauer groups, but in view of the scope of this chapter we only dealt with those aspects of the Brauer group that are close to the algebraic theory concerning the (generalized) separable algebras representing the elements. Needless to say that, even though Azumaya algebras are in many ways near to commutative rings, their explicit structure is on most occasions still mysterious. In fact it is safe to say that even matrix rings over a commutative ring may hold a few surprizes now and then as will be seen in the final section.

6. Separable algebras from the Schur subgroup

The *Schur subgroup* $S(R)$ of $\mathrm{Br}(R)$ is defined by taking the classes represented by Azumaya algebras that may be obtained as epimorphic images of group rings of finite groups, i.e. there exists a finite group G and an R-algebra epimorphism $\pi : RG \twoheadrightarrow A$.

The *projective Schur subgroup* $PS(R)$ of $\mathrm{Br}(R)$ is defined by replacing RG in the above by $R^\alpha G$, the twisted group ring with respect to a 2-cocycle α representing $[\alpha] \in H^2(G, R^*)$, where R^* denotes the multiplicative group of units in the R. The Azumaya algebras representing elements of the Schur subgroup, resp. the projective Schur subgroup, are called Schur algebras, resp. projective Schur algebras (Caution: the notion of a Schur rings as it appears in several papers on representation theory is completely different).

An Azumaya algebra over R is a *generalized Clifford system* over some finite group G if there exist R-submodules M_g of A such that $A = \sum_{g\in G} M_g$ and $M_h = M_g M_h \cong M_g \otimes_R M_h$. It is not hard to see that every projective Schur algebra over R is a generalized Clifford system but for the converse to hold we need that $\mathrm{Cl}(R) = 1$.

When $R = k$ is a field every $M_n(k)$ is a Schur algebra but in general we can only establish:

PROJECTION 6.1. If $2 \in R^*$ then $M_n(R)$ is a Schur algebra for every $n \in \mathbb{N}$.

THEOREM 6.2 (P. Nelis [405]). *When $n \neq 1$, $M_n(\mathbb{Z})$ is a Schur algebra if and only if* 8 *divides n.*

The proof of this theorem is not trivial at all; an essential ingredient is that a finite subgroup of $GL_n(\mathbb{Z})$ that spans $M_n(\mathbb{Z})$ as a $\mathbb{Z}$-module is necessarily a subgroup of the automorphism group of an even unimodular positive definite quadratic form. Thompson in [571] proved that such forms exist only when $8 \mid n$ or $n = 1$ and in order to prove that the condition is also sufficient P. Nelis in [405] used some details on the modular representations of S_8. Of course, because of the lack of non-trivial 2-cocycles, projective Schur algebras over $\mathbb{Z}$ are just Schur algebras but this will no longer be true over more general Dedekind domains.

By restricting attention to particular groups in the definition of $PS(R)$ one may define subgroups: $PS^{ab}(R)$, $PS^{\mathrm{nil}}(R)$, $PS^{P}(R)$, ... where the finite groups appearing in the definition are supposed to be Abelian, nilpotent, p-groups, ... We say that R satisfies the Merkurjev–Suslin condition whenever the m-torsion part of the Brauer group $\mathrm{Br}(r)_m$ equals $PS^{ab}(R)$, where m is the order of a root of unity contained in R.

THEOREM 6.3. *If R satisfies the Merkurjev–Suslin condition and if every cyclic extension of R is contained in a Galois Kummer extension of R then* $\mathrm{Br}_{\mathrm{rat}}(R) \subset PS(R)$, *where* $\mathrm{Br}_{\mathrm{rat}}$ *stands for the rational Brauer subgroup* $\mathrm{Br}(R)$.

If K is a complete non-Archimedean field with finite residue class field then the Hasse invariant yields an isomorphism $\mathrm{Inv}\colon \mathrm{Br}(K) \to \mathbb{Q}/\mathbb{Z}$. This allows to give a description of the Brauer group of an arbitrary number field by using a local-global procedure. Let $\mathcal{S}$ be the set of all places of a number field K, $q \in \mathcal{S}$, $I_q = \mathbb{Q}/\mathbb{Z}$ if K_q is non-Archimedean and $I_q = 2^{-1}\mathbb{Z}/\mathbb{Z}$ if $K_q \simeq \mathbb{R}$, $I_q = 0$ when $K_q \simeq \mathbb{C}$. For every $q \in \mathcal{S}$ we have an isomorphism: $\mathrm{Inv}_q \mathrm{Br}(K_q) \to I_q$ and we identify the groups I_q as subgroups of $\mathbb{Q}/\mathbb{Z}$.

THEOREM 6.4. *The following sequence is exact:*

$$0 \to \mathrm{Br}(K) \xrightarrow{f} \bigotimes_{q \in \mathcal{S}} I_q \xrightarrow{g} \mathbb{Q}/\mathbb{Z} \to 0,$$

where $f = \bigoplus_q \mathrm{Inv}_q$ and g is the "sum"-map.

This famous theorem may be restated as follows: the Brauer group of a number field is given by selecting Hasse invariants $a_q \in \mathbb{Q}/\mathbb{Z}$ such that their sum in $\mathbb{Q}/\mathbb{Z}$ is zero, cf. Pierce [436, Ch. 17.18].

Now, taking $R = \mathbb{Q}$, then the Merkurjev–Suslin theorem states that the Merkurjev–Suslin condition holds in this case, cf. van der Kallen [319, p. 148]. Moreover, by the Kronecker–Weber theorem every Abelian extension of $\mathbb{Q}$ is subcyclotomic and it is not hard to derive from this that $\mathrm{Br}(\mathbb{Q}) = PS(\mathbb{Q})$. In a similar way one establishes the following.

PROPOSITION 6.5. *For every numberfield K we have* $\mathrm{Br}(K) = PS(K)$.

BOLD CONJECTURE 6.6. For any field of characteristic zero: $\mathrm{Br}(K) = PS(K)$.

Note that the conjecture holds if K is such that we may apply the Merkurjev–Suslin theorem; could one obtain an independent non K-theoretic proof for that theorem by first establishing $\mathrm{Br}(K) = PS(K)$ in a completely algebraic (representation theoretic) way?

The effect of the M-S-theorem is that algebras representing elements of the Brauer group are Brauer-equivalent to tensor products of cyclic crossed products. In this way generalized Clifford algebras enter the picture. Recall that an R-algebra A is said to be a *generalized Clifford Algebra* of rank m whenever there exist R-generators $u_1, \ldots, u_m$ in A such that $u_i^n \in R^*$ and u_i and $v_j = wu_j u_i$, where w is a fixed n-th root of unity.

Now a 2-cocycle c representing $[c] \in H^2(G, R^*)$ where G is a finite Abelian group, defines a symplectic pairing $f_c : G \times G \to R^*$, by $f_c(g,h) = c(g,h)c(h,g)^{-1}$, and a corresponding radical $\mathrm{rad}(G) = \{\sigma \in G, f_c(\sigma, \tau) = 1 \text{ for all } \tau \in G\}$.

THEOREM 6.7 (Žmud [677]). (1) *A non-degenerate symplectic pairing on an Abelian group is isometric to an orthogonal sum of hyperbolic planes.*

(2) *A symplectic Abelian group with a cyclic radical is uniquely determined up to isometry by its invariants.*

A cocycle c on $G = C_n^m = C_n \times \cdots \times C_n$ is said to determine a generalized Clifford representation if the radical $\mathrm{rad}(G)$ is minimal in the sense that it is trivial for even m and isomorphic to C_n when m is odd. A result by S. Caenepeel and the author relates generalized Clifford representations and algebras.

THEOREM 6.8. *The generalized Clifford algebras are exactly the twisted group rings $(RC_n^m)^c$ where (C_n^m, c) determines a generalized Clifford representation. Moreover, a generalized Clifford algebra of even rank is isomorphic to a tensor product of generalized quaternion algebras.*

Now look again at $R = k$, a field. To $\alpha, \beta \in k^*$ we can associate the generalized quaternion algebra $(\alpha, \beta)_n$ with generators x, y and relations $x^n = \alpha$, $y^n = \beta$, $ny = wyx$ for some n-th root of unity w. By Matsumito's theorem (cf. [], Theorem 11.1, p. 93) the group $K_2(k)$ is generated by the so-called Steinberg symbols $\{\alpha, \beta\}$ and we may define a group morphism $\psi_n : K_2(k) \to \mathrm{Br}(k)$, $\{\alpha, \beta\} \mapsto [(\alpha, \beta)_n]$.

THEOREM 6.9. *Let k be a field, then we have*:

$$PS^{ab}(k) = \bigoplus_n \{\psi_n(K_2(k)),\ n \text{ such that } w_n \in k^*\}.$$

OBSERVATION 6.10. If k is a field such that the group of roots of unity of k has order n then $PS^{ab}(k) = \mathrm{Br}(k)_n$; if k contains arbitrary roots of unity then $PS^{ab}(k) = \mathrm{Br}(k)$. Of course for a ring R containing n-th roots of unity the inclusion $PS^{ab}(R) \subset \mathrm{Br}(R)_n$ is still valid. However, take for example R to be the ring of integers of a number field K having at least two real embeddings, then $PS^{ab}(R) \neq \mathrm{Br}(R)_n$.

If we write $\mathrm{Br}^{(P)}(R)$ for the p-subgroup of $\mathrm{Br}(R)$ then:

THEOREM 6.11. *With notation as before*:
(1) $PS^{ab}(R) \subset \mathrm{Br}(R)_m$,
(2) $PS^{\mathrm{nil}}(R) \subset \prod_{p/m} \mathrm{Br}^{(p)}(R)$,
(3) $PS^{(P)}(R) \subset \mathrm{Br}^{(p)}(R)$.

If L is a cyclotomic Galois extension of $\mathbb{Q}$ and d a 2-cocycle of $\mathrm{Gal}(L/K)$, taking values in the group of roots of unity of L^* then the crossed product algebra $(L, \mathrm{Gal}(L/K), d)$ is called a *cyclotomic algebra*. The following theorem determines $S(k)$ in case of a number field:

THEOREM 6.12 (Brauer–Witt). *The subgroup $S(k)$ is generated by the classes of cyclotomic algebras.*

In [406] the relation between the existence of Schur algebras over a number ring and the existence of a certain positive definite unimodular Hermitian form is investigated. However in the case of a number ring R there is no equivalent of the Brauer–Witt theorem reducing the determination of $S(R)$ to the study of a very particular class of algebras. Looking at the diagram of injective group morphisms:

$$\begin{array}{ccc} S(R) & \longrightarrow & \mathrm{Br}(R) \\ \downarrow & & \downarrow \\ S(K) & \longrightarrow & \mathrm{Br}(K) \end{array}$$

we may view $S(R)$ as a subgroup of $\mathrm{Br}(R) \cap S(K)$. It is known that $S(R) \neq \mathrm{Br}(R) \cap S(K)$ is possible but the following conjecture may be stated.

THE SCHUR GROUP CONJECTURE 6.13. If R is the ring of integers of a numberfield contained in a cyclotomic numberfield, or if R is a localization of a ring as mentioned before, then $S(R) = \mathrm{Br}(R) \cap S(K)$.

Since the Brauer group of rings R as in the conjecture is well-known the conjecture reduces the calculation of $S(R)$ to the calculation of $S(K)$ and the consideration of cyclotomic K-algebras. In more specific cases more precise description of Schur algebras becomes possible; using the classification of the finite subgroups of the quaternions over $\mathbb{Q}$ it is possible to list all number rings having a Schur algebra that is embeddable in a quaternion skew field. In quaternion skewfields one may even list all projective Schur algebras. For details we refer to [407], let us just mention one of the concrete results there.

THEOREM 6.14. *Let K be a numberfield having a real embedding. A non-trivial projective Schur algebra is realizable in a quaternion skewfield over K if and only if K satisfies one of the following properties.*

(a) *K has an imaginary unramified extension L of degree 2 such that its ring of integers S is a Kummer extension of R.*
(b) *There is a $\mu \in R$ such that $(2) = (\mu)^2$.*
(c) $Q(\sqrt{5}) \subset K$.

In each case the corresponding Azumaya algebras are given by:
(a) $A \cong \left(\frac{S,\beta}{R}\right)$*, the quaternion ring determined by some $\beta \in R^*$.*
(b) $A \cong R\{1, (1+i)\mu^{-1}, (1+j)\mu^{-1}, (1+i+j-k)\mu^{-1}\}$ *where* $(\mu)^2 = (2)$.
(c) $A \cong R\{1, \frac{1}{4}(-1+\sqrt{5}+2i+(1+\sqrt{5})j), \frac{1}{4}(-1-\sqrt{5}+(\sqrt{5}-1)j+2k), j\}$.

The Azumaya algebras in case (b) were also constructed by L. Childs, [134], using smash products for Hopf algebras of rank 2 over R. It is possible to give now an example of a generalized Clifford algebra that is not a projective Schur algebras over R.

EXAMPLE 6.15. Put $R = \mathbb{Z}[\sqrt{10}]$. We know that $\mathrm{Cl}(R) = \mathbb{Z}/2\mathbb{Z}$ and the fundamental unit of R is $3+\sqrt{10}$. Let P be the ideal $(2, \sqrt{10})$ in $\mathbb{Z}[\sqrt{10}]$. Then $P^2 = (2)$ but P is not principal since $X^2 - 10Y^2 = \pm 2$ cannot be solved in $\mathbb{Z}$ (reduce modulo 5). Define $A = R \oplus (i+i)P^{-1} \oplus (1+j)P^{-1} \oplus \frac{1}{2}(1+i+j+k)R$. Putting $L = \mathbb{Q}(\sqrt{10}, \sqrt{-1})$ we see that A is an Azumaya algebra (calculate the discriminant $\mathrm{disc}(A/R) = R$) contained in the quaternion algebra: $\left(\frac{L,\ -1}{\mathbb{Q}(\sqrt{10})}\right)$. One can embed the binary octahedral group E_{48} into A and verify that A is a generalized Clifford algebra but it cannot be a projective Schur algebra because $\mathbb{Q}(\sqrt{10})$ does not satisfy one of the conditions of Theorem 6.14.

References

This list of references is an extension of the check-list on Brauer Groups compiled by A. Verschoren (RUCA, Antwerp). It is more up to date as far as separable algebras are being concerned but we did not strive to keep up to the completeness standard where publications about the Brauer group in its own right are concerned.

[1] A.A. Albert, *Normal division algebras of degree four over an algebraic field*, Trans. Amer. Math. Soc. **34** (1932), 363–372.

[2] A.A. Albert, *Non-cyclic algebras of degree and exponent four*, Amer. Math. Soc. Transl., Vol. 35 (1933), 112–121.

[3] A.A. Albert, *Involutorial simple algebras and real Riemann matrices*, Ann. of Math. **36** (1935), 886–964.

[4] A.A. Albert, *Normal division algebras of degree p^e over F of characteristic p*, Trans. Amer. Math. Soc. **39** (1936), 183–188.

[5] A.A. Albert, *Simple algebras of degree p^e over a centrum of characteristic p*, Trans. Amer. Math. Soc. **40** (1936), 112–126.

[6] A.A. Albert, *A note on normal division algebras of prime degree*, Bull. Amer. Math. Soc. **44** (1938), 649–652.

[7] A.A. Albert, *Non-cyclic algebras with pure maximal subfields*, Bull. Amer. Math. Soc. **44** (1938), 576–579.

[8] A.A. Albert, *Structure of Algebras*, Amer. Math. Soc. Coll. Publications, Vol. 24 (1939).

[9] A.A. Albert, *On p-adic fields and rational division algebras*, Ann. of Math. **41** (1940), 674–693.

[10] A.A. Albert, *Division algebras over a function field*, Duke Math. J. **8** (1941), 750–762.

[11] A.A. Albert, *Two element generation of a separable algebra*, Bull. Amer. Math. Soc. **50** (1944), 786–788.

[12] A.A. Albert, *On associative division algebras of prime degree*, Proc. Amer. Math. Soc. **16** (1965), 799–802.

[13] A.A. Albert, *New results on associative division algebras*, J. Algebra **5** (1967), 110–132.
[14] A.A. Albert, *Tensor products of quaternion algebras*, Proc. Amer. Math. Soc. **35** (1972), 65–66.
[15] A.A. Albert and H. Hasse, *A determination of all normal division algebras over an algebraic number field*, Trans. Amer. Math. Soc. (1932), 722–726.
[16] S.A. Amitsur, *La représentation d'algèbres centrales simples*, C. R. Acad. Sci. Paris Sér. A **230** (1950), 902–904.
[17] S.A. Amitsur, *Construction d'algèbres centrales simples*, C. R. Acad. Sci. Paris. Sér. A **230** (1950), 1026–1028.
[18] S.A. Amitsur, *Differential polynomial and division algebras*, Ann. of Math. **59** (1954), 245–278.
[19] S.A. Amitsur, *On rings with identities*, J. London Math. Soc. **30** (1955), 370–455.
[20] S.A. Amitsur, *Generic splitting fields of central simple algebras*, Ann. of Math. **62** (1955), 8–43.
[21] S.A. Amitsur, *Finite subgroups of division rings*, Amer. Math. Soc. (1955), 361–386.
[22] S.A. Amitsur, *Some results on central simple algebras*, Ann. of Math. **63** (1956), 285–293.
[23] S.A. Amitsur, *Simple algebras and cohomology groups of arbitrary fields*, Trans. Amer. Math. Soc. **90** (1959), 73–112.
[24] S.A. Amitsur, *Homology groups and double complexes for arbitrary fields*, J. Math. Soc. Japan **14** (1962), 1–25.
[25] S.A. Amitsur, *Embeddings in matrix rings*, Pacific J. Math. **36** (1971), 21–39.
[26] S.A. Amitsur, *On central division algebras*, Israel J. Math. **12** (1972), 408–420.
[27] S.A. Amitsur, *Polynomial identities and Azumaya algebras*, J. Algebra **27** (1973), 117–125.
[28] S.A. Amitsur, *Generic crossed product for arbitrary groups and their generic splitting fields*, in preparation.
[29] S.A. Amitsur, *Extension of derivations and automorphisms*, to appear.
[30] S.A. Amitsur, L.H. Rowen and J.P. Tignol, *Division algebras of degree* 4 *and* 8 *with involution*, Israel J. Math. **33** (1979), 133–148.
[31] S.A. Amitsur and D. Saltman, *Generic Abelian crossed products and p-algebras*, J. Algebra **51** (1978), 76–87.
[32] D.D. Anderson and D.F. Anderson, *Divisorial ideals and invertible ideals in a graded integral domain*, J. Algebra **76** (1982), 549–569.
[33] J.K. Arason, *Cohomologische invarianten quadratischer Formen*, J. Algebra **36** (1975), 448–491.
[34] E. Artin and J. Tate, *Class Field Theory*, W.A. Benjamin (1967).
[35] E. Artin, C. Nesbitt and R. Thrall, *Rings with Minimum Condition*, Univ. of Michigan Press, Ann Arbor (1943).
[36] M. Artin, *Grothendieck Topologies*, Lecture Notes, Harvard University, Math. Dept., Cambridge, MA (1962).
[37] M. Artin, *Etale coverings of schemes over Hensel rings*, Amer. J. Math. **88** (1966), 915–934.
[38] M. Artin, *Azumaya algebras and finite dimensional representations of rings*, J. Algebra **11** (1969), 532–563.
[39] M. Artin, *On the joins of Hensel rings*, Adv. in Math. **7** (1971), 282–296.
[40] M. Artin and B. Mazur, *Etale Homotopy*, Lecture Notes in Math., Vol. 100, Springer, Berlin (1969).
[41] M. Artin and D. Mumford, *Some elementary examples of unirational varieties which are not rational*, Proc. London Math. Soc. **25** (1972), 75–95.
[42] K. Asano, *Arithmetische Idealtheorie in Nichtkommutativen Ringen*, Japan J. Math. **16** (1936), 1–36.
[43] K. Asano, *Arithmetik in schiefringen I*, Osaka Math. J. **1** (1949), 98–134.
[44] B. Auslander, *Finitely generated reflexive modules over integrally closed Noetherian domains*, PhD thesis, University of Michigan (1963).
[45] B. Auslander, *The Brauer group of a ringed space*, J. Algebra **4** (1966), 220–273.
[46] B. Auslander, *Central separable algebras which are locally endomorphism rings of free modules*, Proc. Amer. Math. Soc. **30** (1971), 395–404.
[47] B. Auslander and A. Brumer, *Brauer groups of discrete valuation rings*, Nederl. Akad. Wetensch. Proc. Ser. A **71** (1968), 286–296.
[48] M. Auslander and A. Brumer, *The Brauer group and Galois cohomology of commutative rings*, Preprint.
[49] M. Auslander and D.A. Buchsbaum, *On ramification theory in Noetherian rings*, Amer. J. Math. **81** (1959), 749–765.

[50] M. Auslander and O. Goldman, *The Brauer group of a commutative ring*, Trans. Amer. Math. Soc. **97** (1960), 367–409.
[51] M. Auslander and O. Goldman, *Maximal orders*, Trans. Amer. Math. Soc. **97** (1960), 1–24.
[52] J. Ax, *A field of cohomological dimension* 1 *which is not* C_1, Bull. Amer. Math. Soc. **71** (1965), 717.
[53] J. Ax, *Proof of some conjectures on cohomological dimension*, Proc. Amer. Math. Soc. **16** (1965), 1214–1221.
[54] G. Azumaya, *On maximally central algebras*, Nagoya Math. J. **2** (1951), 119–150.
[55] G. Azumaya, *The Brauer group of an adèle ring*, NSF Adv. Sci. Sem. Bourdoin College (1966).
[56] G. Azumaya, *Algebras with Hochschild dimension* $\leqslant 1$, Ring Theory, Proc. of a Conference on Ring Theory, R. Gordon, ed., Park City, UT, 1971, Academic Press, New York (1972), 9–27.
[57] G. Azumaya, *Separable algebras*, J. Algebra **63** (1980), 1–14.
[58] M. Barr and M.-A. Knus, *Extensions of derivations*, Proc. Amer. Math. Soc. **28** (1971), 313–314.
[59] H. Bass, *K-Theory and Stable Algebra*, Publ. Math. IHES 22 (1964).
[60] H. Bass, *Lectures on Topics in Algebraic K-Theory*, Tata Institute for Fundamental Research, Bombay (1967).
[61] H. Bass, *Algebraic K-Theory*, Benjamin, New York (1968).
[62] M. Beattie, *A direct sum decomposition for the Brauer group of H-module algebras*, J. Algebra **43** (1976), 686–693.
[63] M. Beattie, *The Brauer group of central separable G-Azumaya algebras*, J. Algebra **54** (1978), 516–525.
[64] M. Beattie, *Automorphisms of G-Azumaya algebras*, Canad. J. Math. **37** (1985), 1047–1058.
[65] M. Beattie, *The subgroup structure of the Brauer group of RG-dimodule algebras,* Ring Theory 1985, Lecture Notes in Math., Vol. 1197, Springer, Berlin (1987).
[66] M. Beattie, *Computing the Brauer group of graded Azumaya algebras from its subgroups*, J. Algebra **101** (1986), 339–349.
[67] M. Beattie and S. Caenepeel, *The Brauer long group of* $\mathbb{Z}/p^r\mathbb{Z}$*-dimodule algebras*, in preparation.
[68] L. Bégueri, *Schéma d'automorphisms. Application à l'étude d'extensions finies radicielles*, Bull. Sci. Math. **93** (1969), 89–111.
[69] M. Benard, *Quaternion constituents of group algebras*, Proc. Amer. Math. Soc. **30** (1971), 217–219.
[70] M. Benard, *The Schur subgroup I*, J. Algebra **22** (1972), 347–377.
[71] M. Benard and M. Schacher, *The Schur subgroup II*, J. Algebra **22** (1972), 378–385.
[72] H. Benz and H. Zassenhaus, *Über verschränkte Produktordnungen*, J. Number Theory **20** (1985), 282–298.
[73] A.J. Berkson and A. McConnell, *On inflation-restriction exact sequences in group and Amitsur cohomology*, Trans. Amer. Math. Soc. **141** (1969), 403–413.
[74] A. Blanchard, *Les Corps Non Commutatifs*, Presses Univ., France (1972).
[75] S. Bloch, K_2 *and algebraic cycles*, Ann. of Math. **99** (1974), 349–379.
[76] S. Bloch, *Torsion algebraic cycles,* K_2 *and Brauer groups of function fields*, Bull. Amer. Math. Soc. **80** (1974), 941–945.
[77] S. Bloch, *Torsion algebraic cycles,* K_2 *and Brauer groups of function fields*, Springer-Verlag, Berlin (1981), 75–102.
[78] S. Bloch and A. Ogus, *Gersten's conjecture and the cohomology of schemes*, Ann. Sci. École Norm. Sup. **7** (1974), 181–202.
[79] Z. Borevitch and I. Safranevic, *Number Theory*, Pure and Appl. Math., Vol. 20, Academic Press, New York (1967).
[80] N. Bourbaki, *Algèbre, Eléments de Mathématique*, 4, 6, 7, 11, 14, 23, 24, Hermann, Paris (1947–1959).
[81] N. Bourbaki, *Algèbre Commutative, Eléments de Mathématique*, 27, 28, 30, 31, Hermann, Paris (1961–1965).
[82] R. Brauer, *Über Systeme hyperkomplexer Zahlen*, Math. Z. **30** (1929), 79–107.
[83] R. Brauer, *Über der Index und der Exponenten von Divisionalgebren*, Tohoku Math. J. **37** (1933), 77–87.
[84] R. Brauer, *On sets of matrices with coefficients in a divisionring*, Trans. Amer. Math. Soc. **49** (1941), 502–548.
[85] R. Brauer, *On normal division algebras of index* 5, Proc. Nat. Acad. Sci. USA **24**, 243–246.
[86] R. Brauer, *Algebra der hyperkomplexen Zahlensystemen (Algebren)*, Math. J. Okayama Univ. **21** (1979), 53–89.

[87] R. Braun, *On M. Artin's theorem and Azumaya algebras*, University of Haifa, Report 2 (1980).
[88] J. Brezinski, *Algebraic geometry of ternary quadratic forms and orders in quaternion algebras*, Preprint (1980).
[89] J. Brezinski, *A characterization of Gorenstein orders in quaternion algebras*, Preprint (1980).
[90] A. Brumer, *Remarques sur les couples de formes quadratiques*, C. R. Acad. Sci. Paris, Sér. A **286** (1978), 679–681.
[91] A. Brumer and M. Rosen, *On the size of the Brauer group*, Proc. Amer. Math. Soc. **19** (1968), 707–711.
[92] I. Bucur and A. Deleanu, *Introduction to the Theory of Categories and Functors*, Wiley, London (1968).
[93] R.T. Bumby and D.E. Dobbs, *Amitsur cohomology of quadratic extensions: Formulas and number-theoretic examples*, Pacific J. Math. **60** (1975), 21–30.
[94] S. Caenepeel, *A cohomological interpretation of the graded Brauer group, I*, Comm. Algebra **11** (1983), 2129–2149.
[95] S. Caenepeel, *A cohomological interpretation of the graded Brauer group, II*, J. Pure Appl. Algebra **38** (1985), 19–38.
[96] S. Caenepeel, *Gr-complete and Gr-Henselian rings*, Methods in Ring Theory, Reidel, Dordrecht (1984).
[97] S. Caenepeel, *On Brauer groups of graded Krull domains and positively graded rings*, J. Algebra **99** (1986), 466–474.
[98] S. Caenepeel, *A graded version of Artin's refinement theorem*, Ring Theory, 1985, Lecture Notes in Math., Vol. 1197, Springer, Berlin (1986).
[99] S. Caenepeel and F. Van Oystaeyen, *Brauer Groups and Cohomology of Graded Rings*, Monographs, Vol. 121, Marcel Dekker, New York (1988).
[100] S. Caenepeel, *Cancellation theorems for projective graded modules*, Ring Theory, Granada 1986, Lecture Notes in Math., Vol. 1324, Springer, Berlin (1989).
[101] S. Caenepeel, M. Van den Bergh and F. Van Oystaeyen, *Generalized crossed products applied to maximal orders, Brauer groups and related exact sequences*, J. Pure Appl. Algebra **33** (1984), 213–149.
[102] S. Caenepeel and F. Van Oystaeyen, *Crossed products over graded local rings*, Brauer Groups in Ring Theory and Algebraic Geometry, Lecture Notes in Math., Vol. 917, Springer, Berlin (1982).
[103] S. Caenepeel and F. Van Oystaeyen, *A note on generalized Clifford algebras*, Comm. Algebra, to appear.
[104] S. Caenepeel and A. Verschoren, *A relative version of the Chase–Harrison–Rosenberg sequence*, J. Pure Appl. Algebra **41** (1986), 149–168.
[105] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton University Press, Princeton (1956).
[106] J.W.S. Cassels and A. Frölich, *Algebraic Number Theory*, Academic Press, London and New York (1967).
[107] P. Cartier, *Questions de rationalité des diviseurs en géométrie algébrique*, Bull. Soc. Math. France **86** (1958), 177–251.
[108] N.T. Čebotarev, *Introduction to the Theory of Algebras*, OGIZ, Moscow (1949) (in Russian).
[109] C.N. Chang, *The Brauer group of an Amitsur field I.*
[110] C.N. Chang, *The Brauer group of an Amitsur field II*, Proc. Amer. Math. Soc. **39**, 493–496.
[111] C.N. Chang, *The cohomology group of an Amitsur field*, Bull. Inst. Math. Acad. Sinica **5** (1977), 215–218.
[112] S.U. Chase, *On inseparable Galois theory*, Bull. Amer. Math. Soc. **77** (1971), 413–417.
[113] S.U. Chase, *On the automorphism scheme of a purely inseparable field extension*, Ring Theory, E. Gordon, ed., Academic Press, London (1972), 75–106.
[114] S.U. Chase, *Infinitesimal groupscheme actions on finite field extensions*, Amer. J. Math. **98** (1976), 441–480.
[115] S.U. Chase, *Some remarks on forms of algebra.*
[116] S.U. Chase, *On a variant of the Witt and Brauer group*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 148–187.
[117] S.U. Chase, P.K. Harrison and A. Rosenberg, *Galois theory and cohomology of commutative rings*, Mem. Amer. Math. Soc. **52** (1965), 1–19.
[118] S.U. Chase and A. Rosenberg, *Amitsur cohomology and the Brauer group*, Mem. Amer. Math. Soc. **52** (1965), 20–45.
[119] S.U. Chase and A. Rosenberg, *A theorem of Harrison, Kummer theory and Galois algebras*, Nagoya Math. J. (1966), 663–685.
[120] S.U. Chase and A. Rosenberg, *Centralizers and Forms of Algebras.*

[121] S.U. Chase and M.E. Sweedler, *Hopf Algebras and Galois Theory*, Lecture Notes in Math., Vol. 97, Springer, Berlin (1969).

[122] F. Chatelet, *Variations sur un thème de H. Poincaré*, Ann. École Norm. Sup. **61** (1944), 249–300.

[123] F. Chatelet, *Géométrie Diophantienne et théorie des algèbres*, Sém–Dubreil, Exp. 17 (1954–1955).

[124] L.N. Childs, *On projective modules and automorphisms of central separable algebras*, Canad. J. Math. **21** (1969), 44–53.

[125] L.N. Childs, *The exact sequence of low degree and normal algebras*, Bull. Amer. Math. Soc. **76** (1970), 1121–1124.

[126] L.N. Childs, *Abelian extensions of rings containing roots of unity*, Illinois J. Math. **15** (1971), 273–280.

[127] L.N. Childs, *On normal Azumaya algebras and the Teichmüller cocycle map*, J. Algebra **13** (1972), 1–17.

[128] L.N. Childs, *Brauer groups of affine rings*, Ring Theory, Proceedings of the Oklahoma Conference, Lecture Notes in Pure and Appl. Math., Marcel Dekker, New York (1974), 83–93.

[129] L.N. Childs, *Mayer–Vietories sequences and Brauer groups of non-normal domains*, Trans. Amer. Math. Soc. **196** (1974), 51–67.

[130] L.N. Childs, *The full Brauer group and Cech cohomology*, The Brauer Group of Commutative Rings, M. Orzech and C. Small, eds, Lecture Notes in Pure and Appl. Math., Marcel Dekker, New York (1975), 147–162.

[131] L.N. Childs, *The Brauer group of graded Azumaya algebras II, graded Galois extensions*, Trans. Amer. Math. Soc. **204** (1975), 137–160.

[132] L.N. Childs, *On Brauer groups of some normal local rings*, Brauer Groups, Evanston, 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 1–15.

[133] L.N. Childs, *The group of unramified Kummer extensions of prime degree*, Proc. London Math. Soc. **35** (1977), 407–422.

[134] L.N. Childs, *Azumaya algebras over number rings as crossed products*, Preprint, to appear.

[135] L.N. Childs and F. De Meyer, *On automorphisms of separable algebras*, Pacific J. Math. (1967), 25–34.

[136] L.N. Childs, G. Garfinkel and M. Orzech, *The Brauer group of graded Azumaya algebras*, Trans. Amer. Math. Soc. **175** (1973), 299–326.

[137] L.N. Childs, G. Garfinkel and M. Orzech, *On the Brauer group and factoriality of normal domains*, J. Pure Appl. Algebra **6** (1975), 111–123.

[138] M. Cipolla, *Remarks on the lifting of algebras over Henselian pairs*, Math. Z. **152** (1977), 253–257.

[139] L. Claborn and R. Fossum, *Generalization of the notion of class group*, Illinois J. Math. **12** (1968), 228–253.

[140] C. Cliff and A. Weiss, *Crossed product orders and wild ramification*, Lecture Notes in Math., Vol. 1142, 96–104.

[141] P.M. Cohn, *Morita equivalence and duality*, Queen Mary Coll. Math. Notes (1966).

[142] P.M. Cohn, *Algebra II*, Wiley, New York (1977).

[143] C.M. Cordes, *The Witt group and the equivalence of fields with respect to quadratic forms*, J. Algebra **26** (1973), 400–421.

[144] E.F. Cornelius, *A sufficient condition for separability*, J. Algebra **67** (1980), 476–478.

[145] R.C. Courter, *The dimension of maximal commutative subalgebras of K_n*, Duke Math. J. **32** (1965), 225–232.

[146] C.W. Curtis, *On commuting rings of endomorphisms*, Canad. J. Math. **8** (1956), 271–291.

[147] C.W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras*, Wiley Interscience, New York (1962).

[148] C.W. Curtis and I. Reiner, *Methods of Representation Theory with Applications to Finite Groups and Orders*, Wiley, New York (1981).

[149] E. Dade, *The maximal finite subgroups of* 4×4 *matrices*, Illinois J. Math. **9** (1965), 99–122.

[150] E. Dade, *Compounding Clifford's theory*, Ann. of Math. **91** (2) (1970), 236–270.

[151] J. Dauns, *Noncyclic division algebras*, Math. Z. **169** (1979), 195–204.

[152] A.P. Deegan, *A subgroup of the generalized Brauer group of Γ-Azumaya algebras*, J. London Math. Soc. **23** (1981), 223–240.

[153] J.P. Delale, *Algèbre d'Azumaya 1, Généralités*, Sém. d'Alg. Non Commutative (1974), Deuxième Partie, exp. 8 et 10, 26 pp, Publ. Math. Orsay, Université Paris XI, Orsay (1974), 107–7523.

[154] P. Deligne, *Cohomologie étale, SGA$\frac{1}{2}$*, with J.F. Boutot, L. Illusie, J.L. Verdier, Lecture Notes in Math., Vol. 569, Springer, Berlin (1977).

[155] F.R. De Meyer, *Some notes on the general Galois theory of rings*, Osaka J. Math. **2** (1965), 117–127.

[156] F.R. De Meyer, *The Brauer group of some separably closed rings*, Osaka J. Math. **3** (1966), 201–204.

[157] F.R. De Meyer, *The trace map and separable algebras*, Osaka J. Math. **3** (1966), 7–11.

[158] F.R. De Meyer, *Projective modules over central separable algebras*, Canad. J. Math. **21** (1969), 39–43.

[159] F.R. De Meyer, *On automorphisms of separable algebras II*, Pacific J. Math. **32** (1970), 621–631.

[160] F.R. De Meyer, *Separable polynomials over a commutative ring*, Rocky Mountain J. Math. **2** (1972), 299–310.

[161] F.R. De Meyer, *Separable polynomials over a commutative ring*, Ring Theory, Proc. of the Oklahoma Conference, B.R. McDonald, A.R. Magid and K.C. Smith, eds, Marcel Dekker, New York (1974), 75–82.

[162] F.R. De Meyer, *The Brauer group of a polynomial ring*, Pacific J. Math. **59** (1975), 391–398.

[163] F.R. De Meyer, *The Brauer group of a ring modulo an ideal*, Rocky Mountain J. Math. **6** (1976), 191–198.

[164] F.R. De Meyer, *The Brauer group of affine curves*, Brauer Groups, Evanston, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 16–24.

[165] F.R. De Meyer, *The closed socle of an Azumaya algebra*, Proc. Amer. Math. Soc. **78** (1980), 299–303.

[166] F.R. De Meyer and T.J. Ford, *The Brauer group of a surface*, Preprint (1981).

[167] F.R. De Meyer and T.J. Ford, *Computing the Brauer-Long group of $\mathbb{Z}/2$-dimodule algebras*, J. Pure Appl. Algebra, to appear.

[168] F.R. De Meyer and E. Ingraham, *Separable algebras over commutative rings*, Lecture Notes in Math., Vol. 181, Springer, Berlin (1971).

[169] F.R. De Meyer and G. Janusz, *Finite groups with an irreducible representation of a large degree*, Math. Z. **108** (1969), 145–153.

[170] F.R. De Meyer and M.A. Knus, *The Brauer group of real curve*, Proc. Amer. Math. Soc. **57** (1976), 227–232.

[171] F.R. De Meyer and R. Mollin, *The Schur group of a commutative ring*, J. Pure Appl. Algebra **35** (1985), 117–122.

[172] F.R. De Meyer and R. Mollin, *The Schur group*, Lecture Notes in Math., Vol. 1142, 205–209.

[173] M. Deuring, *Einbettung von Algebren, in Algebren mit kleineren Zentrum*, J. Reine Angew. Math. **175** (1936), 124–128.

[174] M. Deuring, *Algebren*, 2nd edn, Ergebnisse der mathematik und ihre Grenzgebiete, Springer, Berlin (1968).

[175] L.E. Dickson, *New division algebras*, Trans. Amer. Math. Soc. **28** (1926), 207–234.

[176] L.E. Dickson, *Construction of division algebras*, Trans. Amer. Math. Soc. **32** (1930), 319–334.

[177] J. Dieudonné, *Compléments à trois articles antérieurs*, Bull. Soc. Math. France **74** (1946), 59–68.

[178] J. Dieudonné, *La théorie de Galois des anneaux simples et semi-simples*, Comment. Math. Helv. **21** (1948), 154–184.

[179] J. Dieudonné, *Les extensions quadratiques de corps non commutatifs et leurs applications*, Acta Math. **87** (1952), 175–242.

[180] D.E. Dobbs, *On inflation and torsion in Amitsur cohomology*, Canad. J. Math. **24** (1972), 239–261.

[181] D.E. Dobbs, *Rings satisfying $x^{n(x)} = x$ have trivial Brauer group*, Bull. London Math. Soc. **5** (1973), 176–178.

[182] P. Donovan and M. Karoubi, *Graded Brauer groups and K-theory with local coefficients*, IHES Publ. **38** (1970), 5–25.

[183] L. Dornhoff, *Group Representation Theory*, Pure and Appl. Math., Vol. 7, Marcel Dekker Inc. (1971).

[184] A. Dress, *On the localization of simple algebra*, J. Algebra **11** (1969), 53–55.

[185] M. Eichler, *Bestimmung der Ideal Klassenzahl in gewissen Normalen einfachen Algebren*, J. Reine Angew. Math. **176** (1937), 192–202.

[186] M. Eichler, *Über die Idealklassenzahl total definiter Quaternion-Algebren*, Math. Z. **43** (1937), 102–109.

[187] M. Eichler, *Über die Idealklassenzahl hyperkomplexer Systeme*, Math. Z. **43** (1937), 481–494.

[188] S. Eilenberg and S. McLane, *Cohomology and Galois theory* 1, *normality of algebras and Teichmüllers cocycle*, Trans. Amer. Math. Soc. **64** (1948), 1–20.

[189] R. Elman and T.Y. Lam, *On the quaternion symbol homomorphism $g_F : k_2 F \to B(F)$*, Algebraic K-Theory II, Lecture Notes in Math., Vol. 342, Springer-Verlag (1973).

[190] R. Elman and T.Y. Lam, *Quadratic forms and the u-invariant I*, Math. Z. **131** (1973), 283–304.

[191] R. Elman and T.Y. Lam, *Quadratic forms under algebraic extensions*, Math. Ann. **219** (1976), 21–42.

[192] R. Elman, T.Y. Lam, J.-P. Tignol and A.R. Wadsworth, *Witt rings and Brauer groups under multiquadratic extensions* 1, Preprint.

[193] R. Elman, T.Y. Lam and A.-R. Wadsworth, *Quadratic forms under multiquadratic extensions*, Nederl. Akad. Wetensch. Proc. Ser. A **83** (1980), 131–145.

[194] S. Endo and Y. Watanabe, *On separable algebras over a commutative ring*, Osaka J. Math. **4** (1967), 233–242.

[195] S. Endo and Y. Watanabe, *The theory of algebras over a commutative ring*, Sûgaku **21** (1969), 24–41. (Japanese).

[196] G. Epiney, *Modules projectifs sur* $H[X, Y]$, travail de diplôme, E.P.F. Zürich.

[197] D.K. Faddeev, *Simple algebras over a field of algebraic functions of one variable*, Trudy Mat. Inst. Steklov. **38** (1951), 321–344; Amer. Math. Soc. Transl. Ser. II **3** (1956), 15–38.

[198] C. Faith, *Algebra I, Rings, Modules and Categories*, Grundlehren der math. Wissenschaft., Springer, Berlin (1973).

[199] C. Faith, *Algebra II, Ring Theory*, Grundlehren der math. Wissenschaft., B. 191, Springer, Berlin (1976).

[200] D.R. Farkas and R.I. Snider, *Commuting rings of simple* $A(k)$*-modules*.

[201] B. Fein, *A note on Brauer–Speiser theorem*, Proc. Amer. Math. Soc. **25** (1970), 620–621.

[202] B. Fein and M.M. Schacher, *Embedding finite groups in rational division algebras I*, J. Algebra **17** (1971), 412–428.

[203] B. Fein and M.M. Schacher, *Embedding finite groups in rational division algebras II*, J. Algebra **19** (1971), 131–139.

[204] B. Fein and M.M. Schacher, *Finite subgroups occuring in finite-dimensional division algebras*, J. Algebra (1974), 332–338.

[205] B. Fein and M.M. Schacher, *Brauer groups of fields algebraic over* Q, J. Algebra **43** (1976), 328–337.

[206] B. Fein and M.M. Schacher, *Galois groups and division algebras*, J. Algebra **38** (1976), 182–191.

[207] B. Fein and M.M. Schacher, *Ulm invariants of the Brauer group of a field*, Math. Z. **154** (1977), 41–50.

[208] B. Fein and M.M. Schacher, *Ulm invariants of the Brauer group of a field, II*, Math. Z. **163** (1978), 1–3.

[209] B. Fein and M.M. Schacher, *Brauer groups and character groups of function fields*, J. Algebra **61** (1979), 249–255.

[210] B. Fein and M.M. Schacher, *Divisible groups that are Brauer groups*, Comm. Algebra **7** (1979), 989–994.

[211] B. Fein and M.M. Schacher, *Strong crossed product division algebras*, Comm. Algebra **8** (1980), 451–466.

[212] B. Fein, M.M. Schacher and J. Sonn, *Brauer groups of rational function fields*, Bull. Amer. Math. Soc. **1**, 766–768.

[213] B. Fein, M.M. Schacher and A. Wadsworth, *Division algebras and the square root of* -1, J. Algebra **65** (1980), 340–346.

[214] W. Feit, *The computation of Schur indices*, Israel J. Math. **46** (4) (1983), 275–300.

[215] K.L. Fields, *On the Brauer–Speiser theorem*, Bull. Amer. Math. Soc. **77** (1971), 223.

[216] K.L. Fields, *On the Schur subgroup*, Bull. Amer. Math. Soc. **77** (1971), 477–478.

[217] K.L. Fields and I.N. Herstein, *On the Schur subgroup of the Brauer group*, J. Algebra **20** (1972), 70–71.

[218] J. Fischer-Palmqvist, *On the Brauer group of a closed category*, Proc. Amer. Math. Soc. **50** (1975), 61–67.

[219] J.M. Fontaine, *Sur la décomposition des algèbres de groupes*, Ann. Sci. École Norm. Sup. **4** (1971), 121–180.

[220] C. Ford, *Pure, normal maximal subfields for division algebras in the Schur subgroup*, Bull. Amer. Math. Soc. **78** (1972), 810–812.

[221] N.S. Ford, *Inertial subalgebras of central separable algebras*, Proc. Amer. Math. Soc. **60** (1976), 39–44.

[222] T.J. Ford, *Every finite Abelian group is the Brauer group of a ring*.

[223] T.J. Ford, *On the Brauer group of a Laurent polynomial ring*, to appear.

[224] E. Formanek, *Central polynomials for matrix rings*, J. Algebra **23** (1972), 129–132.

[225] E. Formanek, *The center of the ring of* 3×3 *generic matrices*, Linear and Multilinear Algebra **7** (1979), 203–212.

[226] E. Formanek, *The center of the ring of* 4×4 *generic matrices*, J. Algebra **62** (1980), 304–320.

[227] R.M. Fossum, *The Noetherian different of projective orders*, PhD thesis, Univ. of Michigan (1965).

[228] R.M. Fossum, *Maximal orders over Krull domains*, J. Algebra **10** (1968), 321–332.
[229] R.M. Fossum, *The Divisor Class Group of a Krull Domain*, Ergebnisse der Mathematik, B. 74, Springer, Berlin (1973).
[230] R.M. Fossum, P. Griffith and I. Reiten, *The homological algebra of trivial extensions of Abelian categories with application to ring theory*, Preprint.
[231] T.V. Fossum, *On symmetric orders and separable algebras.*
[232] A. Frölich, *The Picard group of non-commutative rings, in particular of orders*, Trans. Amer. Math. Soc. **180** (1973), 1–45.
[233] A. Frölich, I. Reiner and S. Ullom, *Class groups and Picard groups of orders*, Proc. London Math. Soc. **29** (1974), 405–434.
[234] A. Frölich and C.T.C. Wall, *Equivalent K-theory and algebraic number theory*, Lecture Notes in Math., Vol. 108, Springer, Berlin (1969), 12–27.
[235] A. Frölich and C.T.C. Wall, *Equivariant Brauer groups in algebraic number theory*, Bull. Soc. Math. France **25** (1971), 91–96.
[236] A. Frölich and C.T.C. Wall, *Generalizations of the Brauer group, I*, Preprint.
[237] O. Gabber, *Some theorems on Azumaya algebras*, PhD thesis, Harvard Univ. (1978); Groupe de Brauer, Lecture Notes in Math., Vol. 844, Springer, Berlin (1981), 129–209.
[238] J. Gamst and K. Hoechsmann, *Quaternions généralisés*, C. R. Acad. Sci. Paris **269** (1969), 560–562.
[239] G.S. Garfinkel, *Amitsur cohomology and an exact sequence involving pic and the Brauer group*, PhD thesis, Cornell University (1968).
[240] G.S. Garfinkel, *A Torsion version of the Chase–Rosenberg exact sequence*, Duke Math. J. **42** (1975), 195–210.
[241] G.S. Garfinkel, *A module approach to the Chase–Rosenberg–Zelinski sequences*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 50–62.
[242] G.S. Garfinkel and M. Orzech, *Galois extensions as modules over the group ring*, J. Math. **22** (1970), 242–248.
[243] M. Gerstenhaber, *On finite inseparable extensions of exponent one*, Bull. Amer. Math. Soc. **71** (1965), 878–881.
[244] M. Gerstenhaber and A. Zaromp, *On the Galois theory of purely inseparable field extensions*, Bull. Amer. Math. Soc. **76** (1970), 1011–1014.
[245] G.T. Georgantas, *Inseparable Galois cohomology*, J. Algebra **38** (1976), 368–379.
[246] G.T. Georgantas, *Derivations in central separable algebra*, Glasgow Math. J. **19** (1978), 75–77.
[247] W.D. Geyer, *Ein algebraischer Beweis des Satzes von Weichold über reelle algebraische Funktionenkörpen*, Algebraische Zahlentheorie. Bericht Tagung Math. Forschungsint. Oberwolfach, Bibliographisches Institut Mannheim (1967), 83–98.
[248] J. Giraud, *Cohomologie Non Abélienne*, Springer, Berlin (1971).
[249] W. Glassmire, *A note on noncrossed products*, J. Algebra **31** (1974), 206–207.
[250] A. Goldie, *Azumaya algebras and rings with polynomial identity*, Math. Proc. Cambridge Philos. Soc. (1976), 393–399.
[251] A. Goldie, *Addendum "Azumaya algebras and rings with polynomial identity"*, Math. Proc. Cambridge Philos. Soc. **80** (1976), 383–384.
[252] O. Goldman, *Quasi-equality in maximum orders*, J. Math. Soc. Japan **13** (1961), 371–376.
[253] M.J. Greenberg, *Lectures on Forms in Many Variables*, Benjamin, New York (1969).
[254] P.A. Griffith, *The Brauer group of $A[T]$*, Math. Z. **147** (1976), 79–86.
[255] A. Grothendieck, M. Artin and J.L. Verdier, *Théorie des Topos et Cohomologie Étale des Schémas (1963–64)*, Lecture Notes in Math., Vols 269, 270, 305, Springer, Berlin (1972–73).
[256] A. Grothendieck, *Le groupe de Brauer I. Algèbres d'Azumaya et interprétations diverses, II. Théorie cohomologique III. Examples et compléments*, Dix Exposés sur la Cohomologie des Schémas, North-Holland, Amsterdam (1968), 46–188.
[257] K. Grünberg, *Profinite groups*, Algebraic Number Theory, Academic Press, New York (1967), 116–127.
[258] D.E. Haille, *On central simple algebras of given exponent I*, J. Algebra **57** (1979), 449–456.
[259] D.E. Haille, *On central simple algebras of given exponent II*, J. Algebra **66** (1980), 205–219.
[260] D.E. Haille, *Abelian extensions of arbitrary fields*, Trans. Amer. Math. Soc. **52** (1963), 230–235.
[261] M. Harada, *Some criteria for hereditary of crossed product orders*, Osaka J. Math. **1** (1964), 69–80.

[262] D.K. Harrison, *Abelian Extensions of Commutative Rings*, Mem. Amer. Math. Soc. **52** (1977).

[263] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Math., Springer, Berlin (1977).

[264] A. Hattori, *Certain cohomology associated with Galois extensions of commutative rings*, Sci. Papers Coll. Gen. Educ. Univ. Tokyo **24** (1974), 79–91.

[265] A. Hattori, *On groups $H^n(S, G)$ and the Brauer group of commutative rings*, Sci. Papers Coll. Gen. Educ. Univ. Tokyo **28** (1978–79), 1–20.

[266] A. Hattori, *On groups $H^n(S/R)$ related to the Amitsur cohomology and the Brauer group of commutative rings*, Preprint.

[267] N. Heerema, *Higher derivations and automorphisms of complete local rings*, Bull. Amer. Math. Soc. **76** (1970), 1212–1225.

[268] N. Heerema, *A Galois theory for inseparable fields extensions*, Trans. Amer. Math. Soc. **154** (1971), 193–200.

[269] I.N. Herstein, *Noncommutative Rings*, Carus Math. Monographs, Vol. 15, Math. Ass. of America, Chicago (1968).

[270] I.N. Herstein, *On a theorem of A.A. Albert*, Scripta Math. **29** (1973), 391–394.

[271] A. Heuser, *Über den Funktionenkörper der Normfläche einer zentral einfachen Algeba*, J. Reine Angew. Math. **301** (1978), 105–113.

[272] D.G. Higman, *On orders in separable algebras*, Canad. J. Math. **7** (1955), 509–515.

[273] K. Hirata, *Some types of separable extensions of rings*, Nagoya Math. J. **33** (1968), 108–115.

[274] K. Hirata, *Separable extensions and centralizers in rings*, Nagoya Math. J. **35** (1969), 31–45.

[275] K. Hirata and K. Sugano, *On semisimple extensions and separable extensions over non commutative rings*, J. Math. Soc. Japan **18** (1966), 360–373.

[276] G. Hochschild, *Automorphisms of simple algebras*, Trans. Amer. Math. Soc. **69** (1950), 292–301.

[277] G. Hochschild, *Restricted Lie algebras and simple associative algebras of characteristic p*, Trans. Amer. Math. Soc. **80** (1955), 135–147.

[278] G. Hochschild, *Simple algebras with purely inseparable splitting fields of exponent one*, Trans. Amer. Math. Soc. **79** (1955), 477–489.

[279] K. Hoechsmann, *Algebras split by a given purely inseparable field*, Proc. Amer. Math. Soc. **14** (1963), 768–776.

[280] K. Hoechsmann, *Simple algebras and derivations*, Trans. Amer. Math. Soc. **108** (1963), 1–12.

[281] K. Hoechsmann, *Über nicht-kommutative abelsche Algebren*, J. Reine Angew. Math. **218** (1965), 1–5.

[282] A. Holleman, *Inseparable algebras*, J. Algebra **46** (1977), 415–229.

[283] R. Hoobler, *A generalization of the Brauer group and Amitsur cohomology*, PhD thesis, University of California, Berkeley (1966).

[284] R. Hoobler, *Non-Abelian sheaf cohomology by derived functors*, Category Theory, Homology and Their Applications III, Lecture Notes in Math., Vol. 99, Springer, Berlin (1969), 313–364.

[285] R. Hoobler, *Brauer groups of Abelian schemes*, Ann. Sci. École Norm. Sup. **5** (1972), 45–70.

[286] R. Hoobler, *Cohomology in the finite topology and Brauer groups*, Pacific J. Math. **42** (1972), 667–679.

[287] R. Hoobler, *Cohomology of purely inseparable Galois coverings*, J. Reine Angew. Math. **266** (1974), 183–199.

[288] R. Hoobler, *Purely inseparable Galois theory*, Ring Theory: Proceedings of the Oklahoma Conference, Lecture Notes in Pure and Appl. Math., Marcel Dekker, New York (1974), 207–240.

[289] R. Hoobler, *A cohomology interpretation of Brauer groups of rings*, Pacific J. Math. **86** (1980), 89–92.

[290] R. Hoobler, *When is* $\mathrm{Br}(X) = \mathrm{Br}'(X)$, Brauer Groups in Ring Theory and Algebraic Geometry, Lecture Notes in Math., Vol. 917, Springer, Berlin (1982).

[291] R. Hoobler, *Functors of graded rings*, Methods in Ring Theory, Reidel, Dordrecht (1984).

[292] J.M. Hood, *Central simple p-algebras with purely inseparable subfields*, J. Algebra **17** (1971), 299–301.

[293] J. Hübschmann, *Group extensions, crossed pairs and an eight term exact sequence*, J. Reine Angew. Math. **321** (1981), 150–172.

[294] W. Hürlimann, *Sur le groupe de Brauer d'un anneau de polygones en caracteristique p et la théorie des invariants*, Groupe de Brauer, Lecture Notes in Math., Vol. 844, Springer, Berlin (1981), 229–274.

[295] E.C. Ingraham, *Inertial subalgebras of algebras over commutative rings*, Trans. Amer. Math. Soc. **124** (1966), 77–93.

[296] E.C. Ingraham, *Inertial subalgebras of complete algebras*, J. Algebra **15** (1970), 1–11.

[297] E.C. Ingraham, *On the existence and conjugacy of inertial subalgebras*, J. Algebra **31** (1974), 547–556.

[298] B. Iversen, *Brauer group of a linear algebraic group*, J. Algebra **42** (1976), 295–301.

[299] N. Jacobson, *Non commutative and cyclic algebras*, Ann. of Math. (1933), 197–208.

[300] N. Jacobson, *Abstract derivations and Lie algebras*, Trans. Amer. Math. Soc. **42** (1937), 206–224.

[301] N. Jacobson, *p-algebras of exponent p*, Bull. Amer. Math. Soc. **43** (1967), 667–670.

[302] N. Jacobson, *The Theory of Rings*, Amer. Math. Soc., New York (1943).

[303] N. Jacobson, *Construction of central simple associative algebras*, Ann. of Math. **45** (1944), 658–666.

[304] N. Jacobson, *Galois theory of purely inseparable fields of exponent one*, Amer. J. Math. **66** (1944), 645–648.

[305] N. Jacobson, *Lectures in Abstract Algebra I*, Van Nostrand, Princeton (1951).

[306] N. Jacobson, *Structure of Rings*, Amer. Math. Soc., New York (1956).

[307] N. Jacobson, *Generation of separable and central simple algebra*, J. Math. Pure Appl. **36** (1957), 217–227.

[308] N. Jacobson, *Lectures in Abstract Algebra, III*, Van Norstrand, Princeton (1964).

[309] N. Jacobson, *Abraham Adrian Albert*, Bull. Amer. Math. Soc. **80** (1974), 1075–1100.

[310] N. Jacobson, *P.I. Algebras, an Introduction*, Lecture Notes in Math., Vol. 441, Springer, Berlin (1975).

[311] V.I. Jancevskii, *Fields with involution and symmetric elements*, Dokl. Akad. Nauk SSSR **18** (1974), 104–107 (in Russian).

[312] V.I. Jancevskii, *Division algebras over a Henselian discrete valued field k*, Dokl. Akad. Nauk SSSR **20** (1976), 971–974.

[313] G.J. Janusz, *Separable algebras over commutative rings*, Trans. Amer. Math. Soc. **122** (1966), 461–479.

[314] G.J. Janusz, *The Schur index and roots of unity*, Proc. Amer. Math. Soc. **35** (1972), 387–388.

[315] G.J. Janusz, *Generators for the Schur group of local and global number fields*, Pacific J. Math. **56** (1975), 525–546.

[316] G.J. Janusz, *The Schur group of cyclotomic fields*, J. Number Theory **7** (1975), 345–352.

[317] G.J. Janusz, *Simple components of* $Q[\mathrm{Sl}(2,q)]$, Comm. Algebra **1** (1974), 1–22.

[318] G.J. Janusz, *The Schur group of an algebraic number field*, Ann. of Math. **103** (1976), 253–281.

[319] W. van der Kallen, *The Mercurjev–Suslin theorem*, Lecture Notes in Math., Vol. 1142, 157–168.

[320] T. Kanzaki, *On commutator rings and Galois theory of separable algebra*, Osaka J. Math. **1** (1964), 103–115.

[321] T. Kanzaki, *On Galois algebra over a commutative ring*, Osaka J. Math. **2** (1965), 309–317.

[322] T. Kanzaki, *A note on Abelian Galois algebra over a commutative ring*, Osaka J. Math. **3** (1966), 1–6.

[323] T. Kanzaki, *On generalized crossed products and Brauer group*, Osaka J. Math. **5** (1968), 175–188.

[324] I. Kaplansky, *Frölich's local quadratic forms*, J. Reine Angew. Math. **74–77** (1969), 239–240.

[325] I. Kaplansky, *Commutative Rings*, Allyn and Bacon, Boston (1970).

[326] I. Kaplansky, *Commutative Rings*, University of Chicago Press, Chicago (1874).

[327] M. Kargapolov and I. Merzliakov, *Eléments de la théorie des groupes*, Mir, Moscow (1985).

[328] M. Karoubi, *Algèbres de Clifford et K-théorie*, Ann. Sci. École Norm. Sup. **1** (1968), 161–270.

[329] G. Karpilowski, *Projective Representations of Finite Groups*, Marcel Dekker Inc. (1985).

[330] V.J. Katz, *The Brauer group of a regular local ring*, PhD thesis, Brandeis Univ. (1968).

[331] I. Kersten, *Zyklische p-Algebren*, J. Reine Angew. Math. **297** (1978), 177–187.

[332] I. Kersten, *Konstruktion von zusammenhangende Galoisschen Algebren der characteristic* p^k *zu vorgegebener Gruppe der Ordnung* p^f, J. Reine Angew. Math. **306** (1979), 177–184.

[333] M.A. Knus, *Algebras graded by a group*, Category Theory, Homology Theory and their Applications II, Lecture Notes in Math., Vol. 92, Springer, Berlin (1969), 117–133.

[334] M.A. Knus, *Algèbres d'Azumaya et modules projectifs*, Comm. Math. Helv. **45** (1970), 372–383.

[335] M.A. Knus, *Sur le Théorème de Skolem–Noether et sur les Dérivations des algèbres d'Azumaya*, C. R. Acad. Sci. Paris, Sér. A **270** (1970), 637–639.

[336] M.A. Knus, *A Teichmüller cocycle for finite extensions*, Preprint.

[337] M.A. Knus and M. Ojanguren, *A note on the automorphisms of maximal orders*, J. Algebra **22** (1972), 573–577.

[338] M.A. Knus and M. Ojanguren, *Sur le polynome caractéristique et les automorphismes des algèbres d'Azumaya*, Ann. Scuola Norm. Sup. Pisa **26** (1972), 225–231.

[339] M.A. Knus and M. Ojanguren, *Sur quelques applications de la théorie de la descente à l'étude du group Brauer*, Comm. Math. Helv. **47** (1972), 532–542.

[340] M.A. Knus and M. Ojanguren, *Théorie de la descente et d'algèbres d'Azumaya*, Lecture Notes in Math., Vol. 389, Springer, Berlin (1974).
[341] M.A. Knus and M. Ojanguren, *A Mayer–Vietoris sequence for the Brauer group*, J. Pure Appl. Algebra **5** (1974), 345–360.
[342] M.A. Knus and M. Ojanguren, *Modules and quadratic forms over polynomial algebras*, Proc. Amer. Math. Soc. **66** (1977), 223–226.
[343] M.A. Knus and M. Ojanguren, *Cohomologie étale et groupe de Brauer*, Groupe de Brauer, Lecture Notes in Math., Vol. 844, Springer, Berlin (1981).
[344] M.A. Knus, M. Ojanguren and D.J. Saltman, *On Brauer groups in characteristic p*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 25–49.
[345] M.A. Knus, M. Ojanguren and R. Sridharan, *Quadratic forms and Azumaya algebras*, J. Reine Angew. Math. (1978), 231–248.
[346] M.A. Knus and R. Parimala, *Quadratic forms associated with projective modules over quaternion algebras*, J. Reine Angew. Math. **318** (1980), 20–31.
[347] H. Koch, *Galoissche Theorie der p-Erweiterungen*, Deutscher Verlag de Wissenschaften, Berlin (1970).
[348] A. Kovacs, *Homomorphisms of matrix rings into matrix rings*, Pacific J. Math. **49** (1973), 161–170.
[349] A. Kovacs, *Generic splitting fields*, Comm. Algebra **6** (1978), 1017–1035.
[350] H.F. Kreimer, *Abelian p-extensions and cohomology*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 104–111.
[351] H.F. Kreimer and P.M. Cook, *Galois theories and normal bases*, J. Algebra **43** (1976), 115–121.
[352] W. Kuijk, *Generic construction of non-cyclic division algebras*, J. Pure Appl. Algebra **2** (1972), 121–130.
[353] W. Kuijk, *The construction of a large class of division algebras*.
[354] P. La Follette, *Splitting of Azumaya algebras over number rings*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 100–103.
[355] T. Lam, *Algebraic Theory of Quadratic Forms*, Benjamin, Reading (1973).
[356] T. Lam, *Serre's Conjecture*, Lecture Notes in Math., Vol. 635, Springer, Berlin (1973).
[357] S. Lang, *Algebra*, Addison-Wesley, Reading, MA (1965).
[358] S. Lang, *Rapport sur la Cohomologie des Groupes*, Benjamin, New York (1966).
[359] L. Le Bruyn, *A cohomological interpretation of the reflexive Brauer group*, J. Algebra **105** (1987), 250–254.
[360] L. Le Bruyn and F. Van Oystaeyen, *Generalized Rees rings and relative maximal orders satisfying polynomial identities*, J. Algebra **83** (1983), 409–436.
[361] L. Le Bruyn, *Permutation modules and rationality questions* 1, UIA preprint (1988).
[362] L. Le Bruyn, M. Van den Bergh and F. Van Oystaeyen, *Graded Orders*, Birkhäuser, Boston (1983).
[363] F.W. Long, *The Brauer group of dimodule algebras*, J. Algebra **30** (1974), 559–601.
[364] F.W. Long, *A generalization of the Brauer group of graded algebras*, Proc. London Math. Soc. **29** (1974), 237–256.
[365] F. Lorenz and H. Opolka, *Einfache Algebren und Projektive Darstellungen über Zahlkörpern*, Math. Z. (1978), 175–186.
[366] A. Lue, *Non-Abelian cohomology of associative algebras*, Quart. J. Math. **19** (1968), 159–180.
[367] H. Maass, *Beweis des Normensatzes in einfachen Hyperkomplexen Systemen*, Abh. Math. Sem. Univ. Hamburg **12** (1937), 64–69.
[368] S. MacLane and O.F.G. Schilling, *A formula for the direct product of crossed product algebras*, Bull. Amer. Math. Soc. **48** (1942), 108–114.
[369] A.R. Magid, *Pierce's representation and separable algebras*, Illinois J. Math. **15** (1971), 114–121.
[370] A.R. Magid, *The separable closure of some commutative rings*, Trans. Amer. Math. Soc. **170** (1972), 109–124.
[371] A.R. Magid, *The Separable Galois Theory of Commutative Rings*, Lecture Notes in Pure and Appl. Math., Vol. 27, Marcel Dekker, New York (1974).
[372] A.R. Magid, *The Picard sequence of a fibration*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 71–85.
[373] A.R. Magid, *Brauer groups of linear algebraic groups with characters*, Proc. Amer. Math. Soc. **71** (1978), 164–168.

[374] K.I. Mandelberg, *Amitsur cohomology for certain extensions of rings of algebraic integers*, Pacific J. Math. **59** (1975), 161–197.

[375] Yu.I. Manin, *Cubic Forms*, 2nd edn, North-Holland (1986).

[376] Yu.I. Manin, *Le Groupe de Brauer–Grothendieck en géométrie diophantienne*, Actes du Congrès International des Mathématiciens, Vol. 1, Gauthier-Villars, Cubic Forms: Algebra, Geometry, Arithmetic, North-Holland, Amsterdam (1974).

[377] W. Massey, *Homology and Cohomology Theory*, Pure and Appl. Mathematics, Vol. 46, Marcel Dekker, New York.

[378] R. Matsuda, *On algebraic properties of infinite group rings*, Bull. Fac. Sci. Ibaraki Univ. Series A, Math. **7** (1975), 29–37.

[379] B. Mazur, *Notes on etale cohomology of number fields*, Ann. Sci. École Norm. Sup. **6** (1973), 521–556.

[380] A. Micali and O.F. Villamayor, *Sur les algèbres de Clifford*, Ann. Sci. École Norm. Sup. **1** (1968), 271–304.

[381] A. Micali and O.F. Villamayor, *Sur les algèbres de Clifford, II*, J. Reine Angew. Math. **242** (1970), 61–90.

[382] A. Micali and O.F. Villamayor, *Algèbres de Clifford et groupe de Brauer*, Ann. Sci. École Norm. Sup. **4** (1971), 285–310.

[383] J.S. Milne, *The Brauer group of a rational surface*, Invent. Math. **11** (1970), 304–307.

[384] J.S. Milne, *Etale Cohomology*, Princeton University Press, Princeton, NJ (1980).

[385] J. Milnor, *Introduction to Algebraic K-Theory*, Annals of Mathematics Studies, Princeton University Press, Princeton, NJ (1971).

[386] Y. Miyashita, *An exact sequence associated with a generalized crossed product*, Nagoya Math. J. **49** (1973), 21–51.

[387] R. Mollin, *Herstein's conjecture, automorphisms and the Schur group*, Comm. Algebra **6** (1978), 237–248.

[388] R. Mollin, *Uniform distribution classified*.

[389] R. Mollin, *Induced quaternion algebras in the Schur group*, Canad J. Math. **6** (1981), 1370–1379.

[390] S. Montgomery, *Centralizers of separable subalgebras*, Michigan Math. J. **22** (1975), 15–22.

[391] R.A. Morris, *On the Brauer group of $\mathbb{Z}$*, Pacific J. Math. **39** (1971), 619–630.

[392] M. Nagata, *Local Rings*, Interscience Tracts in Pure and Appl. Math., Vol. 13, Wiley, New York (1962).

[393] A. Nakajima, *On a group of cyclic extensions over commutative rings*, Math. J. Okayama Univ. **15** (1972), 163–172.

[394] A. Nakajima, *Some results on H-Azumaya algebras*, Math. J. Okayama Univ. **19** (1976–77), 101–110.

[395] T. Nakayama, *Division Algebren über diskret bewerteten perfekten Körpern*, J. Reine Angew. Math. **178** (1937), 11–13.

[396] C. Năstăsescu, E. Nauwelaerts and F. Van Oystaeyen, *Arithmetically graded rings revisited*, Comm. Algebra **14** (1986), 1991–2008.

[397] E. Năstăsescu and F. Van Oystaeyen, *Graded and Filtered Rings and Modules*, Lecture Notes in Math., Vol. 758, Springer, Berlin (1980).

[398] C. Năstăsescu and F. Van Oystaeyen, *Graded Ring Theory*, Library of Math., Vol. 28, North-Holland, Amsterdam (1982).

[399] C. Năstăsescu and F. Van Oystaeyen, *On strongly graded rings and crossed products*, Comm. Algebra **10** (1982), 2085–2106.

[400] C. Năstăsescu, F. Van Oystaeyen and M. Van den Bergh, *Separable functors, applied to graded rings*, J. Algebra **123** (1989), 397–413.

[401] E. Nauwelaerts and F. Van Oystaeyen, *The Brauer splitting theorem and projective representations of finite groups*, J. Algebra **112** (1988), 49–57.

[402] E. Nauwelaerts and F. Van Oystaeyen, *Finite generalized crossed products over tame and maximal orders*, J. Algebra **101** (1988), 61–68.

[403] P. Nelis and F. Van Oystaeyen, *Ray symmetric projective representations and subgroups of the Brauer-group*, Israel Math. Conf. Proc. **1** (1989), 262–279.

[404] P. Nelis and F. Van Oystaeyen, *The projective Schur subgroup of the Brauer group and root group of finite groups*, J. Algebra, to appear.

[405] P. Nelis, *Integral matrices spanned by finite groups*, UIA Preprint 12 (1990).

[406] P. Nelis, *The Schur group conjecture for the ring of integers of a number field*, Proc. Amer. Math. Soc., to appear.

[407] P. Nelis, *Schur and projective Schur groups of number rings*, Canad. J. Math., to appear.
[408] P. Nelis, *On hereditary and Gorensteinness of crossed product orders*, UIA Preprint 34 (1989).
[409] D. Northcott, *A note on polynomial rings*, J. London Math. Soc. **33** (1958), 36–39.
[410] D. Northcott, *An Introduction to Homological Algebra*, Cambridge University Press, Cambridge (1960).
[411] T. Nyman and G. Whaples, *Hasse's principle for simple algebras over function fields of curves, I, algebras of index* 2 *and* 3, *curves of genus* 0 *and* 1, J. Reine Angew. Math. **299/300** (1978), 396–405.
[412] T. Nyman and G. Whaples, *Correction to the paper "Hasse's Principle ..."*, J. Reine Angew. Math. **301** (1978), 211.
[413] M. Ojanguren, *A non-trivial locally trivial algebra*, J. Algebra **29** (1974), 510–512.
[414] M. Ojanguren, *A remark on generic splitting rings*, Arch. Math. **27** (1976), 369–371.
[415] M. Ojanguren, *Formes quadratiques sur les algèbres de polynomes*, C. R. Acad. Sci. Paris **287** (1978), 695–698.
[416] M. Ojanguren and R.R. Sridharan, *Cancellation of Azumaya algebras*, J. Algebra **18** (1971), 501–505.
[417] T. Okado and R. Saito, *A generalization of P. Roquette's theorems*, Hokkaido Math. J. **1** (1972), 197–205.
[418] S. Okubo and H.C. Myung, *Some new classes of division algebras*, J. Algebra **67** (1980), 479–490.
[419] O. O'Meara, *Introduction to Quadratic Forms*, Academic Press, New York (1963).
[420] O. O'Meara, *Introduction to Quadratic Forms*, Die Grundlehren der Mathematischen Wissenschaften, B. 117, Springer, Berlin (1970).
[421] H. Opolka, *Rationalitätsfragen bei Projektiven Darstellungen endlicher Gruppen*, Dissertation, Universität Münster (1987).
[422] M. Orzech, *A cohomological description of Abelian Galois extensions*, Trans. Amer. Math. Soc. **137** (1969), 481–499.
[423] M. Orzech, *On the Brauer group of modules having a grading and an action*, Canad. J. Math. **28** (1976), 533–552.
[424] M. Orzech, *Brauer groups of graded algebras*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976).
[425] M. Orzech and C. Small, *The Brauer Group of Commutative Rings*, Lecture Notes in Pure and Appl. Math., Vol. 11, Marcel Dekker, New York (1975).
[426] M. Orzech, *Correction to "On the Brauer group of modules having a grading and an action"*, Canad. J. Math. **32** (1980), 1523–1524.
[427] B. Pareigis, *Über normale, zentrale, separabele Algebres und Amitsur Kohomologie*, Math. Ann. **154** (1964), 330–340.
[428] B. Pareigis, *Non-additive ring and module theory IV: The Brauer group of a symmetric monoidal category*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 112–133.
[429] S. Parimala and R. Sridharan, *Projective modules over polynomial rings over division rings*, J. Math. Kyoto Univ. **15** (1975), 129–148.
[430] J.W. Pendergrass, *The* 2*-part of the Schur group*, J. Algebra **41** (1976), 422–438.
[431] J.W. Pendergrass, *Calculations of the Schur group*, Pacific J. Math. **69** (1977), 169–175.
[432] J.W. Pendergrass, *The Schur subgroup of the Brauer group*, Pacific J. Math. **69** (1977), 477–499.
[433] S. Perlis, *Cyclicity of division algebras of prime degree*, Proc. Amer. Math. Soc. **21** (1969), 409–411.
[434] D.J. Picco and M.I. Platzeck, *Graded algebras and Galois extensions*, Bol. Un. Math. Argentina **25** (1971), 401–415.
[435] G. Pickert, *Eine Normalform für Endliche rein-inseparabele Körpererweiterungen*, Math. Z. **53** (1950), 133–135.
[436] R.S. Pierce, *Associative Algebras*, Graduate Texts in Mathematics, Vol. 88, Springer, New York (1982).
[437] W. Plesken and M. Pohlst, *On maximal finite subgroups of* $\mathrm{Gl}_n(\mathbb{Z})$, 1. *The five and seven dimensional cases*, 2. *The six dimensional case*, Math. Comp. **31** (1977), 138, 537–573.
[438] E.C. Posner, *Prime rings satisfying a polynomial identity*, Proc. Amer. Math. Soc. **11** (1960), 180–183.
[439] C. Procesi, *On the identities of Azumaya algebras*, Ring Theory, R. Gordon, ed., Academic Press, New York (1972), 287–295.
[440] C. Procesi, *On a theorem of M. Artin*, J. Algebra **22** (1972), 309–315.
[441] C. Procesi, *Rings with Polynomial Identities*, Marcel Dekker, New York (1973).
[442] C. Procesi, *Finite dimensional representations of algebras*, Israel J. Math. **19** (1974), 109–182.

[443] H.G. Quebbeman, *Schiefkörper als Endomorphismenringe einfache Modulen einer Weyl-Algebra*, J. Algebra **59** (1979), 311–312.
[444] M.L. Racine, *On maximal subalgebras*, J. Algebra **30** (1974), 155–180.
[445] M.L. Racine, *A simple proof of a theorem of Albert*, Proc. Amer. Math. Soc. **43** (1974), 487–488.
[446] M.L. Racine, *Maximal Subalgebras of Central Separable Algebras*, Proc. Amer. Math. Soc. **68** (1978), 11–15.
[447] M. Ramras, *Maximal orders over regular local rings of dimension two*, Trans. Amer. Math. Soc. **142** (1969), 457–479.
[448] M. Ramras, *Maximal orders over regular local rings*, Trans. Amer. Math. Soc. **155** (1971), 345–352.
[449] M. Ramras, *Splitting fields and separability*, Proc. Amer. Math. Soc. **38** (1973), 489–492.
[450] M. Ramras, *The center of an order with finite global dimension*, Trans. Amer. Math. Soc. **210** (1975), 249–257.
[451] M.L. Ranga Rao, *Azumaya, semisimple and ideal algebras*, Bull. Amer. Math. Soc. **78** (1972), 588–592.
[452] M.L. Ranga Rao, *Semisimple algebras and a cancellation law*, Comment. Math. Helv. **50** (1975), 123–128.
[453] R. Rasala, *Inseparable splitting theory*, Trans. Amer. Math. Soc. **162** (1971), 411–448.
[454] M. Raynaud, *Anneaux Locaux Henséliens*, Lecture Notes in Math., Vol. 169, Springer, Berlin (1971).
[455] M.B. Rege, *A note on Azumaya algebras*, J. Indian Math. (N.S.) **37** (1973), 217–225.
[456] I. Reiner, *Maximal Orders*, LMS Monographs, Academic Press, New York (1975).
[457] G. Renault, *Algèbre Non Commutative*, Gauthier-Villars, Paris (1975).
[458] C. Riehm, *Linear and quadratic Schur groups*, Preprint (1988).
[459] C. Riehm, *The linear and quadratic Schur subgroups over the S-integers of a number field*, Proc. Amer. Math. Soc. **107** (1) (1989), 83–87.
[460] C. Riehm, *The Schur subgroup of the Brauer group of cyclotomic rings of integers*, Proc. Amer. Math. Soc. **103** (1) (1989), 83–87.
[461] D. Rim, *An exact sequence in Galois cohomology*, Proc. Amer. Math. Soc. **16** (1965), 837–840.
[462] L.J. Risman, *Stability: Index and order in the Brauer group*, Proc. Amer. Math. Soc. **50** (1975), 33–39.
[463] L.J. Risman, *Zero divisors in tensorproducts of division algebras*, Proc. Amer. Math. Soc. **51** (1975), 33–36.
[464] L.J. Risman, *Cyclic algebras, complete fields and crossed products*, Israel J. Math. **28** (1977), 113–128.
[465] L.J. Risman, *Group rings and series*, Ring Theory, Proceedings of the Antwerp Conference 1977, Marcel Dekker, New York (1977), 113–128.
[466] L.J. Risman, *Non-cyclic division algebras*, J. Pure Appl. Algebra **11** (1977/1978), 199–215.
[467] L.J. Risman, *Twisted rational functions and series*, J. Pure Appl. Algebra **12** (1978), 181–199.
[468] P. Roquette, *On the Galois cohomology of the projective linear group and its applications to the construction of generic splitting fields of algebras*, Math. Ann. **150** (1962), 411–439.
[469] P. Roquette, *Isomorphisms of generic splitting fields of simple algebras*, J. Reine Angew. Math. **214/215** (1964), 207–226.
[470] P. Roquette, *Splitting of algebras by function fields in one variable*, Nagoya Math. J. **27** (1966), 625–642.
[471] A. Rosenberg and D. Zelinski, *On Amitsur's complex*, Trans. Amer. Math. Soc. **97** (1960), 327–356.
[472] A. Rosenberg and D. Zelinski, *Automorphisms of separable algebras*, Pacific J. Math. **11** (1961), 1107–1117.
[473] A. Rosenberg and D. Zelinski, *Amitsur's complex for inseparable fields*, Osaka Math. **14** (1962), 219–240.
[474] S. Rosset, *Locally inner automorphisms of algebras*, J. Algebra **29** (1974), 88–97.
[475] S. Rosset, *Generic matrices, K_2 and unirational fields*, Bull. Amer. Math. Soc. **81** (1975), 707–708.
[476] S. Rosset, *A remark on the generation of the Brauer group.*
[477] S. Rosset, *Abelian splitting of division algebras of prime degrees*, Comment. Math. Helv. **52** (1977), 519–523.
[478] S. Rosset, *Group extensions and division algebras*, J. Algebra **53** (1978), 297–303.
[479] S. Rosset, *The corestriction preserves symbols.*
[480] L.H. Rowen, *Identities in algebras with involution*, Israel J. Math. (1975), 70–95.
[481] L.H. Rowen, *Central simple algebras with involution*, Bull. Amer. Math. Soc. **83** (1977), 1031–1032.
[482] L.H. Rowen, *Central simple algebras*, Israel J. Math. **29** (1978), 285–301.
[483] L.H. Rowen, *Division algebras, counterexample of degree* 8, Israel J. Math., to appear.

[484] L.H. Rowen, *Cyclic division algebras.*

[485] L.H. Rowen and D.J. Saltman, *Dihedral algebras are cyclic*, Proc. Amer. Math. Soc.

[486] A. Roy and B. Sridharan, *Derivations in Azumaya algebras*, J. Math. Kyoto Univ. **7** (1967), 161–167.

[487] A. Roy and B. Sridharan, *Higher derivations and central simple algebras*, Nagoya Math. J. **32** (1968), 21–30.

[488] S. Ryškov, *Maximal finite groups of integral $n \times n$ matrices and full groups of integral automorphism of positive definite quadratic forms (Bravais models)*, Proc. Steklov. Inst. Math. **128** (1972), 217–250.

[489] D.J. Saltman, *Azumaya algebras over rings of characteristics*, PhD thesis, Yale University (1967).

[490] D.J. Saltman, *Splitting of cyclic P-algebras*, Proc. Amer. Math. Soc. **62** (1977), 223–228.

[491] D.J. Saltman, *Azumaya algebras with involution*, J. Algebra **52** (1978), 526–539.

[492] D.J. Saltman, *Noncrossed product p-algebras and Galois p-extensions*, J. Algebra **52** (1978), 302–314.

[493] D.J. Saltman, *Noncrossed products of small exponent*, Proc. Amer. Math. Soc. **68** (1978), 165–168.

[494] D.J. Saltman, *Indecomposable division algebras*, Comm. Algebra **7** (1979), 791–818.

[495] D.J. Saltman, *On p-power central polynomial*, Proc. Amer. Math. Soc. **78** (1980), 11–13.

[496] D.J. Saltman, *Division algebras over discrete valued fields*, Comm. Algebra **8** (1980), 1749–1774.

[497] D.J. Saltman, *Norm polynomials and algebras*, J. Algebra **62** (1980), 333–348.

[498] P. Samuel, *Théorie Algébrique des Normes*, Hermann, Paris (1967).

[499] P. Samuel, *Algebraic Theory of Numbers*, Kershaw Publishing Company Ltd, London (1970).

[500] D. Sanders, *The dominion and separable subalgebra of finitely generated algebras*, Proc. Amer. Math. Soc. **48** (1975), 1–7.

[501] N. Sankaran, *Galois extensions of rings and generalized group extensions*, J. Math. Sci. **8** (1974), 10–16.

[502] M. Schacher, *Subfields of division rings, I*, J. Algebra **9** (1968), 451–477.

[503] M. Schacher, *Subfields of division rings, II*, J. Algebra **10** (1968), 240–245.

[504] M. Schacher, *More on the Schur subgroup*, Proc. Amer. Math. Soc. **31** (1972), 15–17.

[505] M. Schacher, *The crossed product problem*, Ring Theory II, Lecture Notes in Pure and Appl. Math., Vol. 26, Marcel Dekker, New York (1977), 237–245.

[506] M. Schacher, *Non-uniqueness in crossed products*, Representation Theory of Algebras, Lecture Notes in Pure and Appl. Math., Vol. 37, New York (1978), 447–449.

[507] M. Schacher, *Cyclotomic splitting fields.*

[508] M. Schacher and L.W. Small, *Noncrossed products in characteristic p*, J. Algebra **24** (1973), 100–103.

[509] W. Scharlau, *Über die Brauer-Gruppe einer algebraischen Funktionenkörper in einer Variablen*, J. Reine Angew. Math. **239/240** (1969), 1–6.

[510] W. Scharlau, *Über die Brauer Gruppe eines Hensel-Körpers*, Abh. Math. Sem. Univ. Hamburg **33** (1969), 243–249.

[511] W. Scharlau, *Zur Existenz von Involutionen auf einfachen Algebren*, Math. Z. **145** (1975), 29–32.

[512] W. Scharlau, *Involutions on simple algebras and orders*, Canad. Math. Soc. Proc. **4** (1984), 141–147.

[513] W. Schelter, *Azumaya algebras and Artin's theorem*, J. Algebra **46** (1977), 303–304.

[514] O.F.G. Schelling, *Arithmetic in a special class of algebras*, Ann. of Math. **38** (1937), 116–119.

[515] O.F.G. Schelling, *The theory of valuations*, Amer. Math. Surveys **4** (1950).

[516] I. Schur, *Untersuchungen über die Darstellung der endlichen Gruppen durch gebrochene lineare Substitutionen*, Gesammelte Abhandlungen, B. 1, Springer, Berlin (1973), 198–250.

[517] R. Schwäling, *Separable Algebra über Kommutativen Ringen*, J. Reine Angew. Math. **262/263** (1973), 307–322.

[518] G. Seligman, *Modular Lie Algebras*, Ergebnisse der Mathematik, B. 40, Springer, Berlin (1967).

[519] J.-P. Serre, *Applications algébriques de la Cohomologie des groupes, Théorie des algèbres simples*, Sém. H. Cartan, Paris, exposés 5–7 (1950/51).

[520] J.-P. Serre, *Groupes Algébriques et Corps de Classes*, Hermann, Paris (1959).

[521] J.-P. Serre, *Cohomologie Galoisienne*, Lecture Notes in Math., Vol. 5, Springer, Berlin (1964).

[522] J.-P. Serre, *Algèbre Locale: Multiplicités*, Lecture Notes in Math., Vol. 11, Springer, Berlin (1965).

[523] J.-P. Serre, *Algèbres de Lie Sémi-Simple Complexes*, W.A. Benjamin Inc. (1966).

[524] J.-P. Serre, *Local class field theory*, Algebraic Number Theory, J.W.S. Cassels and A. Frölich, eds, Academic Press, New York (1967).

[525] J.-P. Serre, *Corps Locaux*, Hermann, Paris (1968).

[526] J.-P. Serre, *Cours d'Arithmétique*, P.U.F, Le Mathématicien, Vol. 2 (1970).

[527] J.-P. Serre, *Représentations Linéaires des Groupes Finis*, Hermann, Paris, Collections Méthodes (1971).

[528] D. Shapiro, J.-P. Tignol and A.R. Wadsworth, *Witt rings and Brauer groups under multi-quadratic extensions, II*, Preprint.

[529] S. Shatz, *Cohomology of Artinian group schemes over local fields*, Ann. of Math. **79** (1964), 411–449.

[530] S. Shatz, *Galois theory*, Category Theory, Homology Theory and their Applications I, Lecture Notes in Math., Vol. 86, Springer, Berlin (1969).

[531] S. Shatz, *Principal homogeneous spaces for finite group schemes*, Proc. Amer. Math. Soc. **22** (1969), 678–680.

[532] S. Shatz, *Profinite Groups, Arithmetic and Geometry*, Princeton Studies 67 (1972).

[533] L. Silver, *Tame orders, Tame ramification and Galois cohomology*, Illinois J. Math. **12** (1968), 7–34.

[534] C. Small, *The Brauer–Wall group of a commutative ring*, Trans. Amer. Math. Soc. **156** (1971), 455–491.

[535] C. Small, *The group of quadratic extensions*, J. Pure Appl. Algebra **2** (1972), 83–105.

[536] R. Snider, *Is the Brauer group generated by cyclic algebras?*, Ring Theory, Waterloo, 1978, Lecture Notes in Math., Vol. 734, Springer, Berlin (1979), 279–301.

[537] J. Sonn, *Class groups and Brauer groups*, Israel J. Math. **34** (1979), 97–106.

[538] Steinhausen, *Gerade, positiv definite, unimoduläre Gitter.*

[539] H.J. Stolberg, *Azumaya Algebras and Derivations.*

[540] R. Strano, *Azumaya algebras overe Hensel rings*, Pacific J. Math. (1975), 295–303.

[541] K. Sugano, *Note on semisimple extensions and separable extensions*, Osaka J. Math. **4** (1967), 265–270.

[542] K. Sugano, *Separable extensions and Frobenius extensions*, Osaka J. Math. **7** (1970), 291–299.

[543] K. Sugano, *Note on separability of endomorphism rings*, J. Fac. Sci. Hokkaido Univ. **21** (1971), 126–208.

[544] K. Sugano, *On automorphisms in separable extensions of rings*, Proc. of the 13th Symposium on Ring Theory, Okayama (1980), 44–55.

[545] D. Suprenko and R. Tyshkevich, *Commutative Matrices*, Academic Press, New York (1968).

[546] R. Swan, *Projective modules over group rings and maximal orders*, Ann. of Math. **76** (1) (1962), 55–61.

[547] M. Sweedler, *Structure of inseparable extensions*, Ann. of Math. **87** (1968), 401–410.

[548] M. Sweedler, *Correction to: "Structure of inseparable extensions"*, Ann. of Math. **89** (1969), 206–207.

[549] M. Sweedler, *Multiplication alteration by* 2*-cocycles*, Illinois J. Math. **15** (1971), 302–323.

[550] M. Sweedler, *Groups of simple algebras*, Publ. Math. IHES **44** (1974), 79–190.

[551] M. Sweedler, *Amitsur cohomology and the Brauer group*, Intl. Congress of Math. (1974).

[552] M. Sweedler, *Purely inseparable algebras*, J. Algebra **35** (1975), 342–355.

[553] M. Sweedler, *Something like the Brauer group*, Proc. Intl. Congress of Math. Vancouver 1974, Vol. 1, Canad. Math. Congress, Montreal, Que. (1975), 337–341.

[554] M. Sweedler, *Complete separability*, J. Algebra **61** (1979), 199–248.

[555] G. Szeto, *On a class of projective modules over central separable algebras*, Canad. Math. Bull. **14** (1971), 415–417.

[556] G. Szeto, *On a class of projective modules over central separable algebras*, Canad. Math. Bull. **15** (1972), 411–415.

[557] G. Szeto, *The Azumaya algebra of some polynomial closed rings*, Indiana Univ. Math. J. **22** (1973), 731–737.

[558] G. Szeto, *On the Wedderburn theorem*, Canad. J. Math. **25** (1973), 525–530.

[559] G. Szeto, *A characterization of Azumaya algebras*, J. Pure Appl. Algebra **9** (1976/77), 65–71.

[560] G. Szeto, *The Pierce representation of Azumaya algebras*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976).

[561] G. Szeto, *A generalization of the Artin-procesi theorem*, J. London Math. Soc. **15** (1977), 406–408.

[562] M. Takeushi, *On Long's H-Azumaya-algebra*, Proc. 10th Symp. on Ring Theory, Dept. Math. Okayama Univ., Okayama (1978), 46–61.

[563] S. Tanaka, *Construction and classification of irreducible representations of special linear group of the second order over a finite field*, Osaka J. Math. **4** (1967), 65–84.

[564] A. Tannenbaum, *The Brauer group and unirationality: An example of Artin–Mumford*, Groupe de Brauer, Lecture Notes in Math., Vol. 844, Springer, Berlin (1981), 103–128.

[565] J. Tate, *Global class fields theory*, Algebraic Number Theory, J. Cassels and A. Fröhlich, eds, Thompson Book Co. (1976).

[566] J. Tate, *On the conjectures of Birch and Swinnerton–Dyer and a geometric analog*, Dix exposés sur la Cohomologie des schémas, Sém. Bourbaki 1965–66, exposé 306, North-Holland, Amsterdam (1968), 189–214.

[567] J. Tate, *Relations between K_2 and Galois cohomology*, Invent. Math. **36** (1976), 257–274.

[568] O. Teichmüller, *p-Algebren*, Deutsche Math. **1** (1936), 362–388.

[569] O. Teichmüller, *Verschränkte Produkte mit Normalringen*, Deutsche Math. **1** (1936), 92–102.

[570] O. Teichmüller, *Über die sogenannte nicht kommutative Galoissche Theorie und die Relationen*, Deutsche Math. **5** (1940), 138–149.

[571] J.G. Thompson, *Finite groups and even lattices*, J. Algebra **38** (1976), 523–524.

[572] J.P. Tignol, *Sur les corps à involution de degré* 8, C. R. Acad. Sci. Paris, Sér. A **284** (1977), 1349–1352.

[573] J.P. Tignol, *On p-torsion in étale cohomology*, Notices Amer. Math. Soc. **24** A-6, Abstract 77 T-A 24 (1977).

[574] J.P. Tignol, *Décomposition et descente de produits tensoriels d'algèbres de quaternions*, Rapport Sém. Math. Pure UCL 76 (1978).

[575] J.P. Tignol, *Sur les classes de similitude de corps à involution de degré* 8, C. R. Acad. Sci. Paris, Sér. A **286** (1978), 875–876.

[576] J.P. Tignol, *Reflexive modules*, J. Algebra **54** (1978), 444–466.

[577] J.P. Tignol, *Central simple algebras with involution*, Ring Theory, Proc. of the 1978 Antwerp Conference, Lecture Notes in Pure and Appl. Math., Marcel Dekker, New York (1979), 279–288.

[578] J.P. Tignol, *Corps à involution de rang fini sur leur centre et de caractéristique différente de* 2, PhD thesis, UCL Louvain-la-Neuve (1979).

[579] J.P. Tignol, *On p-torsion in etale cohomology and in the Brauer group*, Proc. Amer. Math. Soc. **78** (1980), 189–192.

[580] J.P. Tignol, *Corps à involution neutralisés par une extension Abelienne elémentaire*, Groupe de Brauer, Lecture Notes in Math., Vol. 844, Springer, Berlin (1981), 1–34.

[581] J.P. Tignol, *Produits Croisés Abéliens*.

[582] F. Tilborghs, *The Brauer group of R-algebras wich have compatible G-action and $\mathbb{Z} \times G$-grading*, Preprint.

[583] F. Tilborghs and F. Van Oystaeyen, *Brauer–Wall algebras graded by $\mathbb{Z}_2 \times \mathbb{Z}$*, Comm. Algebra, to appear.

[584] D. Trappers, *Quadratic forms over graded arithmetical rings*, PhD thesis, UIA (1990).

[585] R. Treger, *Reflexive modules*, PhD thesis, University of Chicago (1976).

[586] C. Tsen, *Division Algebren über Funktionenkörpern*, Nachz. Ges. Wiss. Göttingen II **48** (1933), 335–339.

[587] M. Van den Bergh, *A note on graded Brauer groups*, Bull. Soc. Math. Belg. **39** (1987), 177–179.

[588] M. Van den Bergh, *A note on graded K-theory*, Comm. Algebra **14** (1986), 1561–1564.

[589] J.-P. Van Deuren, J. Van Geel and F. Van Oystaeyen, *Genus and a Riemann–Roch theorem for non-commutative function fields in one variable*, Sém. Dubreil (1980).

[590] J. Van Deuren and F. Van Oystaeyen, *Arithmetically graded rings*, Ring Theory, Antwerp, 1980, Lecture Notes in Math., Vol. 825, Springer, Berlin (1981).

[591] J. Van Geel, *Primes in algebras and the arithmetic in central simple algebras*, Comm. Algebra **8** (1980), 1015–1951.

[592] J. Van Geel, *Primes and value functions*, PhD thesis, University of Antwerp, UIA (1980).

[593] J. Van Geel and F. Van Oystaeyen, *About graded fields*, Indag. Math. **43** (1981), 273–286.

[594] F. Van Oystaeyen, *Pseudoplaces of algebras and the symmetric part of the Brauer group*, PhD thesis, Univ. of Amsterdam (1972).

[595] F. Van Oystaeyen, *Generic division algebras*, Bull. Soc. Math. Belg. **25** (1973), 259–285.

[596] F. Van Oystaeyen, *On pseudo-places of algebras*, Bull. Soc. Math. Belg. **25** (1973), 139–159.

[597] F. Van Oystaeyen, *The p-component of the Brauer group of a field of char. p*, Nederl. Akad. Wet. Proc. Ser. A. **77** (1974), 67–76.

[598] F. Van Oystaeyen, *Prime Spectra in Non-Commutative Algebra*, Lecture Notes in Math., Vol. 444, Springer, Berlin (1975).

[599] F. Van Oystaeyen, *Graded Azumaya algebras and Brauer groups*, Ring Theory, Antwerp 1980, Lecture Notes in Math., Vol. 825, Springer, Berlin (1981), 158–171.

[600] F. Van Oystaeyen, *On Brauer groups of arithmetically graded rings*, Comm. Algebra **9** (1981), 1873–1892.

[601] F. Van Oystaeyen, *Crossed products over arithmetically graded rings*, J. Algebra **80** (1983), 537–551.

[602] F. Van Oystaeyen, *Some constructions of rings*, J. Pure Appl. Algebra **31** (1984), 241–252.

[603] F. Van Oystaeyen, *Generalized Rees rings and arithmetically graded rings*, J. Algebra **82** (1983), 185–193.

[604] F. Van Oystaeyen, *On Clifford systems and generalized crossed products*, J. Algebra **87** (1984), 396–415.

[605] F. Van Oystaeyen, *Azumaya strongly graded rings and Ray classes*, J. Algebra **103** (1986), 228–240.

[606] F. Van Oystaeyen and A. Verschoren, *The Brauer group of a projective variety.*

[607] F. Van Oystaeyen and A. Verschoren, *On the Brauer group of a projective variety*, Israel J. Math. **42** (1982), 37–59.

[608] F. Van Oystaeyen and A. Verschoren, *On the Brauer group of a projective curve*, Israel J. Math. **42** (1982), 327–333.

[609] F. Van Oystaeyen and A. Verschoren, *Relative Invariants of Rings, Part I and II*, Marcel Dekker, New York (1983) and (1984).

[610] Van Rooij, *Non-Archimedean Functional Analysis*, Marcel Dekker, New York (1978).

[611] B. van der Waerden, *Moderne Algebra*, Springer, Berlin (1960).

[612] A. Verschoren, *Separable couples, Azumaya couples and sheaves*, Bull. Soc. Math. Belg. **32** (1980), 29–47.

[613] A. Verschoren, *Mayer–Vietoris sequences for Brauer groups of projective varieties.*

[614] A. Verschoren, *Mayer–Vietoris sequences for Brauer groups of graded rings*, Comm. Algebra **10** (1982), 765–782.

[615] A. Verschoren, *The cohomologcal Brauer group of a projective variety.*

[616] M.F. Vigneras, *Arithmétique des Algèbres de Quaternions*, Lecture Notes in Math., Vol. 800, Springer, Berlin (1980).

[617] O.E. Villamayor and D. Zelinski, *Galois theory for rings with finitely many idempotents*, Nagoya Math. J. **27** (1966), 721–731.

[618] O.E. Villamayor and D. Zelinski, *Galois theory with infinitely many idempotents*, Nagoya Math. J. **35** (1969), 83–98.

[619] O.E. Villamayor and D. Zelinski, *Brauer groups and Amitsur cohomology for general commutative ring extensions*, J. Pure Appl. Algebra **10** (1977), 19–55.

[620] C.T.C. Wall, *Graded Brauer groups*, J. Reine Angew. Math. **213** (1964), 187–199.

[621] C.T.C. Wall, *Graded algebras, anti-evolutions, simple groups and symmetric spaces*, Bull. Amer. Math. Soc. **74** (1968), 198–202.

[622] S.S.S. Wang, *Splitting ring of a monic separable polynomial*, Pacific J. Math. **75** (1978), 293–296.

[623] L. Washington, *Introduction to Cyclotomic Fields*, Graduate Texts in Mathematics, Vol. 83, Springer, New York (1980).

[624] Y. Watanabe, *Simple algebras over a complete local ring*, Osaka J. Math. **3** (1966), 73–74.

[625] Y. Watanabe, *On relatively separable subalgebras*, Osaka J. Math. **14** (1977), 165–172.

[626] W.C. Waterhouse, *A non-isiterion for central simple algebras, III*, J. Math. **17** (1973), 73–74.

[627] C.A. Weibel, *Homotopy in Algebraic K-Theory*, University of Chicago (1977).

[628] C.A. Weibel, *K-theory of Azumaya algebras*, Proc. Amer. Math. Soc. **81** (1981), 1–7.

[629] M. Weisfeld, *Purely inseparable extensions and higher derivations,* Trans. Amer. Math. Soc. **116** (1965), 435–449.

[630] S. Williamson, *Crossed products and hereditary orders*, Nagoya Math. J. **23** (1963), 103–120.

[631] S. Williamson, *Crossed products and maximal orders*, Nagoya Math. J. **25** (1965), 165–174.

[632] R. Wisbauer, *Separable Modulen von Algebren über Ringen*, J. Reine. Angew. Math. **303–304** (1978), 221–230.

[633] R. Wisbauer, *Zentrale Bimodulen and separabele Algebren*, Arch. Math. **30** (1978), 129–137.

[634] E. Witt, *Zerlegung reellen algebraischer Funktionen in Quadrata, Schiefkörper über reellen Funktionen Körper*, J. Reine Angew. Math. **171** (1934), 4–11.

[635] E. Witt, *Über ein Gegenbeispiel zum Normensatz*, Math. Z. **39** (1935), 462–467.

[636] E. Witt, *Schiefkörper über diskret bewerteten Körper*, J. Reine Angew. Math. **176** (1936), 153–156.

[637] E. Witt, *Zyklische Körper und Algebren der Charakteristik p vom Grade p^n*, J. Reine Angew. Math. **176** (1936), 126–140.

[638] E. Witt, *Die algebraische Struktur des Gruppen Ringen einer endlichen Gruppe über einen Zahlkörper*, J. Reine Angew. Math. **190** (1952), 231–245.

[639] P. Wolf, *Algebraische Theorie der Galoisschen Algebren*, Math. Forschungsberichte 3, Deutsche Verlag der Wissenschaften, Berlin (1956).
[640] T. Würfel, *Un critère de liberté pour les pro-p-groupes et la structure du groupe de Brauer*, C. R. Acad. Sci. Paris, Sér. A. **288** (1979), 793–795.
[641] T. Würfel, *Ein Freiheitskriterium für pro-p-Gruppen mit Anwendung auf die Strukture der Brauer-Gruppe*, Math. Z. **172** (1980), 81–88.
[642] T. Würfel, *Über die Brauer-Gruppe der Maximalen p-Erweiterung eines Körpers ohne p-te Einheitswurzeln*.
[643] T. Würfel, *On the divisible part the Brauer group of a field*.
[644] T. Yamada, *On the group algebras of metacyclic groups over algebraic number fields, I*, J. Fac. Sci. Univ. Tokyo **15** (1968), 179–199.
[645] T. Yamada, *On the group algebras of metacyclic groups over algebraic number fields, II*, J. Fac. Sci. Univ. Tokyo, Sect. I **16** (1969), 83–90.
[646] T. Yamada, *On the group algebras of metabelian groups over algebraic number fields I*, Osaka J. Math. **6** (1969), 211–228.
[647] T. Yamada, *On the Schur index of a monomial representations*, Proc. Japan. Acad. **45** (1969), 522–525.
[648] T. Yamada, *Characterization of the simple components of the group algebras over the p-adic number field*, J. Math. Soc. Japan **23** (1971), 295–310.
[649] T. Yamada, *Central simple algebras over totally real fields which appear in $Q[G]$*, J. Algebra **23** (1972), 382–403.
[650] T. Yamada, *Cyclotomic algebras over a 2-adic field*, Proc. Japan Acad. (1973), 438–442.
[651] T. Yamada, *Simple components of group algebras $Q[G]$ central over real quadratic fields*, J. Number Theory **5** (1973), 179–190.
[652] T. Yamada, *The Schur subgroup of the Brauer group I*, J. Algebra **27** (1973), 579–589.
[653] T. Yamada, *The Schur subgroup of the Brauer group II*, J. Algebra **27** (1973), 590–603.
[654] T. Yamada, *The Schur subgroup of the Brauer group*, Proc. Conf. on Orders Group Rings and Rel. Top., Ohio State 1972, Lecture Notes in Math., Vol. 353, Springer, Berlin (1973), 187–203.
[655] T. Yamada, *The Schur subgroup of the Brauer group*, Lecture Notes in Math., Vol. 397, Springer, Berlin (1974).
[656] T. Yamada, *On Schur subgroups*, Sugaku **26** (1974), 109–120. (Japanese).
[657] T. Yamada, *The Schur subgroup of 2-adic fields*, J. Math. Soc. Japan **26** (1974), 168–179.
[658] T. Yamada, *The Schur subgroup of Real Quadratic Field I*, Symp. Math., Vol. XV, Academic Press, London (1975), 413–425.
[659] T. Yamada, *The Schur subgroup of real quadratic field II*, J. Algebra **35** (1975), 190–204.
[660] T. Yamada, *The Schur subgroup of a p-adic field*.
[661] K. Yamazaki, *On projective representations and ring extensions of finite groups*, J. Fac. Sci. Univ. Tokyo, Sect. I **10** (1964), 147–195.
[662] K. Yokogawa, *Brauer group of algebraic function fields and their Adèle rings*, Osaka J. Math. **8** (1971), 271–280.
[663] K. Yokogawa, *On $S \otimes_R S$-module structure of S/R-Azumaya algebras*, Osaka J. Math. **12** (1975), 673–690.
[664] K. Yokogawa, *An application of the theory of descent to the $S \otimes_R S$-module structure of S/R-Azumaya algebras*, Osaka J. Math. **15** (1978), 21–31.
[665] K. Yokogawa, *$S - S$-bimodule structure of S/R-Azumaya algebra and 7-term exact sequence*, Proc. of the 10th Symposium on Ring Theory, Dept. Math. Okayama Univ., Okayama (1978), 38–45.
[666] S. Yuan, *Differentially simple rings of prime characteristic*, Duke Math. J. **31** (1964), 623–630.
[667] S. Yuan, *The Brauer groups of local fields*, Ann. of Math. **82** (1965), 434–444.
[668] S. Yuan, *On the p-algebras and the Amitsur groups for inseparable field extensions*, J. Algebra **5** (1967), 280–304.
[669] S. Yuan, *Inseparable Galois-theory of exponent one*, Trans. Amer. Math. Soc. **149** (1970), 163–170.
[670] S. Yuan, *Central separable algebras with purely inseparable splitting fields of exponent one*, Trans. Amer. Math. Soc. **153** (1971), 427–450.
[671] S. Yuan, *p-algebras arbitrary exponents*, Proc. Nat. Acad. Sci. USA **70** (1973), 533–534.
[672] S. Yuan, *Brauer-groups for inseparable fields*, Ann. of Math. **96** (1974), 430–447.

[673] S. Yuan, *Reflexive modules and algebra class groups over Noetherian integrally closed domains*, J. Algebra **32** (1974), 405–417.

[674] D. Zelinski, *Berkson's theorem*, Israel J. Math. **2** (1964), 205–209.

[675] D. Zelinski, *Long exact sequence and the Brauer group*, Brauer Groups, Evanston 1975, Lecture Notes in Math., Vol. 549, Springer, Berlin (1976), 63–70.

[676] D. Zelinski, *Brauer groups*, Ring Theory II, Lecture Notes in Pure and Appl. Math., Vol. 26, Marcel Dekker, New York (1977), 69–101.

[677] E.M. Žmud, *Symplectic geometries of finite Abelian groups*, Mat. Sbornik **68** (1), 128. English transl.: Math. USSR Sb., Amer. Math. Soc. (1971).

This Page Intentionally Left Blank

Section 3D
Deformation Theory of Rings and Algebras

This Page Intentionally Left Blank

Varieties of Lie Algebra Laws

Yu. Khakimdjanov

Faculté des Sciences et Techniques, Université de Haute Alsace, 4, rue des Frères Lumière, 68093 Mulhouse cedex, France

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

1. Introduction

The aim of this chapter is to put together all the results which were obtained recently in the study of Lie algebras from a geometrical point of view. Any Lie algebra law is considered as a point of an affine algebraic variety defined by the polynomial equations coming from the Jacobi identity for a given basis. This approach gives an explanation of the difficulties in classification problems concerning the classes of nilpotent and solvable Lie algebras and the relative facility of the classification of semisimple Lie algebras. Isomorphic Lie algebras correspond to the laws belonging to the same orbit relative to the action of the general linear group and classification problems (up to isomorphism) can be reduced to the classification of these orbits. An affine algebraic variety is a union of a finite number of irreducible components and the Zariski open orbits give interesting classes of Lie algebras to be classified. The Lie algebras of this class are called *rigid*. Any semisimple Lie algebra is rigid. There is also a sufficiently large class of solvable rigid Lie algebras. Now we do not have a complete description of rigid Lie algebras, but we have some results about their properties; and a description of some classes of such rigid ones. We also have low-dimensional results. The rigidity property can be considered as well in the varieties of nilpotent and solvable Lie algebra laws. Generally, rigidity in the variety of solvable Lie algebra laws is the same as rigidity in the variety of all Lie algebra laws. In the case of the variety of nilpotent Lie algebra laws the situation is very different. We do not know about the existence or not of rigid laws (excepting the low-dimensional case). In this paper we give a comprehensive account of the variety of nilpotent Lie algebra laws as recently elaborated. This study shows the reason of the difficulties in the classification problems and precises the place and role of different classes of such Lie algebras. We focus particularly on two classes of nilpotent Lie algebras, which are more or less models of nilpotent Lie algebras: those are filiform algebras (the least commutative ones) and the characteristically nilpotent Lie algebras (all derivations are nilpotent). The study of varieties of Lie algebra laws is essentially based on the cohomological study of Lie algebras and on deformation theory. We hope that this important and beautiful direction of research will continue to develop.

2. Varieties of laws of Lie algebras

2.1. *Generalities*

Let V be a n-dimensional vector space over a field K and let $B^2(V)$ be the space of all bilinear applications $V \times V \to V$. A Lie algebra law on V is an element μ of $B^2(V)$ verifying the following relations

$$\mu(x, y) + \mu(y, x) = 0,$$
$$\mu\big(\mu(x, y), z\big) + \mu\big(\mu(y, z), x\big) + \mu\big(\mu(z, x), y\big) = 0$$

for all elements x, y, z in V. In a given basis of V a Lie algebra law μ is defined by a set of structure constants $\{c_{ij}^k\}$ in the space K^{n^3} verifying the following polynomial conditions

$$\begin{aligned} &c_{ij}^k + c_{ji}^k = 0, \\ &\sum_{k=1}^{n} \left(c_{ij}^k c_{km}^l + c_{jm}^k c_{ki}^l + c_{mi}^k c_{kj}^l\right) = 0. \end{aligned} \tag{1}$$

One identifies the law μ with the set of its structure constants $\{c_{ij}^k\}$. The subset $\mathcal{L}^n$ of the Lie algebra laws in the space K^{n^3} is defined by the system of polynomial equations (1); so $\mathcal{L}^n$ is a affine algebraic variety.

The nilpotent Lie algebra laws of nilpotency class $\leqslant p$ are singled out in $\mathcal{L}^n$ by means of a relation that is defined by polynomial equations for a fixed basis, and they define a closed subvariety $\mathcal{N}_p^n$ in $\mathcal{L}^n$. The nilpotency class of an n-dimensional Lie algebra does not exceed $n-1$, and so $\mathcal{N}_{n-1}^n$ consists of all nilpotent Lie algebra laws of dimension n. We denote this subvariety by $\mathcal{N}^n$ and call it the variety of n-dimensional nilpotent Lie algebra laws. Analogously we can consider the subvariety $\mathcal{R}^n$ of n-dimensional solvable Lie algebra laws in the variety $\mathcal{L}^n$.

The linear group $G = \mathrm{GL}(V)$ naturally acts on the space $B^2(V)$:

$$(g \cdot \mu)(x, y) = g\big(\mu\big(g^{-1}(x), g^{-1}(y)\big)\big).$$

The varieties $\mathcal{L}^n$, $\mathcal{N}_p^n$ and $\mathcal{R}^n$ are invariant with respect to this action. We denote by $O(\mu)$ the orbit of this action containing μ. Evidently, two Lie algebra laws μ_1 and μ_2 are isomorphic if and only if their corresponding orbits coincide.

The variety $\mathcal{L}^n$ can be provided with the classical Euclidean topology. It is also naturally provided with the Zariski topology which is less fine than the preceding topology.

We consider also the affine scheme defined by the ideal of the ring $K[X_1, \ldots, X_{n^3}]$, generated by the polynomials of the relations (1). We will not distinguish in notation the scheme $\mathcal{L}^n$ and the variety $\mathcal{L}^n$. We also consider the schemes $\mathcal{N}_p^n$ and $\mathcal{R}^n$ defined by the corresponding polynomial relations.

EXAMPLE 1. The variety $\mathcal{L}^1$ contains only one point ($c_{1,1}^1 = 0$). The corresponding Lie algebra is Abelian.

EXAMPLE 2. The variety $\mathcal{L}^2$ is defined by the relations

$$\begin{aligned} &c_{1,1}^k = c_{2,2}^k = 0, \quad k = 1, 2; \\ &c_{1,2}^k + c_{2,1}^k = 0, \quad k = 1, 2; \end{aligned}$$

they define a plane in the space K^8. The point with $c_{1,2}^1 = c_{1,2}^2 = 0$ corresponds to the 2-dimensional Abelian Lie algebra. All other points of $\mathcal{L}^2$ form the orbit of the linear group $\mathrm{GL}(K^2)$ containing the law with $c_{1,2}^2 = 1$. This orbit is evidently a Zariski open set in $\mathcal{L}^2$.

2.2. *Tangent space*

In this section we suppose that $K = \mathbb{R}$ or $\mathbb{C}$. The first step in the local study of the variety $\mathcal{L}^n$ consists in determining the tangent space $T_{\mu_0}(\mathcal{L}^n)$ at the point μ_0 to the variety $\mathcal{L}^n$. This may be given in cohomological terms:

THEOREM 3. *The tangent space $T_{\mu_0}(\mathcal{L}^n)$ at the point μ_0 to the scheme $\mathcal{L}^n$ identifies with the space $Z^2(\mu_0, \mu_0)$ of 2-cocycles for the Chevalley cohomology of μ_0 with values in the adjoint module. The tangent space at μ_0 to the variety $\mathcal{L}^n$ is a subspace of $Z^2(\mu_0, \mu_0)$.*

We note that if the scheme $\mathcal{L}^n$ is reduced then $Z^2(\mu_0, \mu_0)$ coincides with tangent space at the point μ_0 to the variety $\mathcal{L}^n$.

The tangent space at the point μ_0 to the orbit $O(\mu_0)$ coincides with the space of 2-coboundaries $B^2(\mu_0, \mu_0)$.

Let $\mu_0 \in \mathcal{L}^n$. The orbit $O(\mu_0)$ can be identified with $\mathrm{GL}(V)/\mathrm{Aut}(\mu_0)$, where $\mathrm{Aut}(\mu_0)$ is the group of automorphisms of μ_0. Thus,

$$\dim O(\mu_0) = n^2 - \dim \mathrm{Aut}(\mu_0).$$

The Lie algebra of the Lie group $\mathrm{Aut}(\mu_0)$ is the algebra of derivations $\mathrm{Der}(\mu_0)$ and so we have

$$\dim O(\mu_0) = n^2 - \dim \mathrm{Der}(\mu_0).$$

2.3. *Formal deformations*

The second step in the local study of the variety $\mathcal{L}^n$ consists in studying formal deformations in the variety $\mathcal{L}^n$.

DEFINITION 4. A *formal deformation* of a law μ_0 of $\mathcal{L}^n$ is a formal sequence with a parameter t

$$\mu_t = \mu_0 + \sum_{i=1}^{\infty} t^i \varphi_i, \tag{2}$$

where the elements φ_i, $1 \leqslant i < \infty$, are the skew-symmetric bilinear maps $V \times V \to V$ such that μ_t verifies the formal Jacobi identity: $\mu_t \circ \mu_t = 0$.

In this definition for $\varphi, \psi \in C^2(V, V)$, we denote by $\varphi \circ \psi$ the trilinear application $V^3 \to V$ defined by

$$\varphi \circ \psi(X, Y, Z) = \varphi\big(\psi(X, Y), Z\big) + \varphi\big(\psi(Y, Z), X\big) + \varphi\big(\psi(Z, X), Y\big).$$

Let μ_t be a deformation of $\mu_0 \in \mathcal{L}^n$. We have

$$\delta\varphi_1 = \mu_0 \circ \varphi_1 + \varphi_1 \circ \mu_0 = 0, \tag{3}$$

where δ is the coboundary operator for the Chevalley cohomology with values in the adjoint module, that is $\varphi_1 \in Z^2(\mu_0, \mu_0)$.

We also have the relations:

$$\begin{aligned}
&\varphi_1 \circ \varphi_1 = -\tfrac{1}{2}\, d\varphi_2,\\
&\varphi_1 \circ \varphi_2 + \varphi_2 \circ \varphi_1 = -d\varphi_3,\\
&\dots\\
&\sum_{i+j=k} (\varphi_i \circ \varphi_j + \varphi_j \circ \varphi_i) = -d\varphi_k.\\
&\dots
\end{aligned} \tag{4}$$

The elements of the space $Z^2(\mu_0, \mu_0)$ are called *infinitesimal deformations* of μ_0.

Let φ be an infinitesimal deformation of μ_0; φ is called *integrable*, if there exists a deformation (2) with $\varphi_1 = \varphi$. Not all infinitesimal deformations are integrable. A necessary and sufficient condition for integrability of $\varphi_1 \in Z^2(\mu_0, \mu_0)$ is the existence of a solution to the system (4). For example, if $H^3(\mu_0, \mu_0) = 0$ any infinitesimal deformation φ_1 is integrable.

Two deformations

$$\mu_t^1 = \mu_0 + \sum_{i=1}^{\infty} t^i \varphi_i$$

and

$$\mu_t^2 = \mu_0 + \sum_{i=1}^{\infty} t^i \psi_i$$

of μ_0 are called *equivalent, if* there exists a linear automorphism Φ_t of the space V of the following form:

$$\Phi_t = \mathrm{Id} + \sum_{i=1}^{\infty} t^i f_i,$$

where $f_i \in \mathrm{End}(V)$, such that

$$\mu_t^2(x, y) = \Phi_t^{-1}\big(\mu_t^1\big(\Phi_t(x), \Phi_t(y)\big)\big).$$

A deformation μ_t of μ_0 is called *trivial* if it is equivalent to the deformation $\mu_t' = \mu_0$. It is easy to see that $\varphi_1 - \psi_1 \in B^2(\mu_0, \mu_0)$ for equivalent deformations μ_t^1 and μ_t^2. Thus, we can consider the set of infinitesimal deformations of μ_0 as parametrized by $H^2(\mu_0, \mu_0) = Z^2(\mu_0, \mu_0)/B^2(\mu_0, \mu_0)$.

2.4. *Deformations in the variety $\mathcal{L}^n$*

In this section we suppose that $K = \mathbb{R}$ or $\mathbb{C}$. The third step consists in determining the "genuine" local deformations in the variety $\mathcal{L}^n$. In the real case that is smooth curves parametrized by a variable t on $\mathcal{L}^n$:

$$c : [-\varepsilon, \varepsilon] \to \mathcal{L}^n$$

containing a given point of $\mathcal{L}^n$ for the value of t equal to 0. For this, a study of the convergence of formal deformations is necessary. Sometimes there are only a finite number of terms in the formula of deformations. For example, if an infinitesimal deformation φ is also a Lie algebra law, that is $\varphi \in \mathcal{L}^n$, the formula

$$\mu_t = \mu_0 + t\varphi$$

defines a deformation called *linear*. In this case the infinitesimal deformation φ is called *linearly integrable*.

We note that it is possible to consider another interval $[a, b]$ containing 0 for the application c.

In the case $K = \mathbb{C}$, we consider a smooth map $c : \Omega \to \mathcal{L}^n$, where Ω is an Euclidean open set in the complex plane $\mathbb{C}$, containing 0.

Analogously we can consider smooth deformations in the varieties $\mathcal{N}_p^n$ and $\mathcal{R}^n$.

Obviously, we can define different types of "smoothness" by considering c as continuous mappings, as analytical mappings, or otherwise.

EXAMPLE 5. Let $\mu_0 = 0$ be the Abelian law in $\mathcal{L}^n$ and let μ be a non Abelian law in $\mathcal{L}^n$. The one parameter family $\mu_t = \mu_0 + t\mu$ defines a linear deformation of μ_0. Evidently all Lie algebras corresponding to non zero values of the parameter t are isomorphic.

EXAMPLE 6. Let $\mu_0 \in \mathcal{L}^7$ defined by

$$\mu_0(X_1, X_i) = X_{i+1}, \quad i = 2, \ldots, 6;$$
$$\mu_0(X_3, X_4) = X_7, \qquad \mu_0(X_2, X_5) = -X_7,$$

where $(X_1, \ldots, X_7)$ is a basis in the space K^7 and let $\varphi \in Z^2(\mu_0, \mu_0)$ be defined by $\varphi(X_2, X_i) = X_{i+2}$, $i = 3, \ldots, 5$. The infinitesimal deformation φ of μ_0 is linearly integrable and the laws of the family $\mu_t = \mu_0 + t\varphi$ are two by two nonisomorphic.

2.5. *Local study of the variety $\mathcal{L}^n$*

In this section we suppose that $K = \mathbb{C}$. The local study of the variety $\mathcal{L}^n$ at a point μ_0 is based on the description of the infinitesimal deformation of μ_0 , that is on the description of $H^2(\mu_0, \mu_0)$. While, for the semisimple Lie algebras or their parabolic subalgebras this study is clear, for the nilpotent points of $\mathcal{L}^n$ this is more complicated. We do not know many results concerning the description of the space $H^2(\mu_0, \mu_0)$ for nilpotent laws μ_0. We know, for example, the local study of $\mathcal{L}^n$ in the neighborhood of the nilradical of a parabolic subalgebra of a semisimple Lie algebra. We give here the following results of this study.

THEOREM 7. *Let* $\mathfrak{n}$ *be the nilradical of a Borel subalgebra of a semisimple Lie algebra. Then the tangent space to the scheme* $\mathcal{L}^n$ *coincides with the tangent space to the corresponding reduced scheme at the point* $\mathfrak{n}$.

THEOREM 8. *Let* $\mathfrak{n}$ *be the nilradical of a parabolic subalgebra of a simple Lie algebra* $\mathfrak{g}$. *Let* $n = \dim \mathfrak{n}$, *and let* p *be the nilpotency class of* $\mathfrak{n}$, $p \leqslant n - 3$. *The Zariski tangent space to the scheme* $\mathcal{L}^n$ *at the point* $\mathfrak{n}$ *coincides with the tangent space to the subvariety of nilpotent Lie algebra laws of nilpotency class* $\leqslant p + 2$.

COROLLARY 9. *Suppose that the hypotheses of the previous theorem hold. Then the Zariski tangent space to the scheme* $\mathcal{L}^n$ *at the point* $\mathfrak{n}$ *coincides with the tangent space to the subscheme* $\mathcal{N}_n$ *of nilpotent Lie algebra laws at the same point.*

The demonstrations of these theorems are based on the determination of the space $H^2(\mathfrak{n}, \mathfrak{n})$ and on the study of integrability of the corresponding infinitesimal deformations.

2.6. *Contractions of Lie algebra laws*

In this section $K = \mathbb{C}$ or $\mathbb{R}$. Let $\varphi : \Omega \to \mathcal{L}^n$ a smooth curve, where Ω is an Euclidean open set in the complex plane $\mathbb{C}$ not containing 0, but $0 \in \overline{\Omega}$ (in the real case we can put $\Omega =]0, \varepsilon[$ or $]-\varepsilon, \varepsilon[\setminus \{0\}$), such that all Lie algebra laws in the image $\varphi(t)$ are isomorphic to a law μ. Suppose that $\lim_{t \to 0} \varphi(t)$ exists; we denote this limit by μ_0. The variety $\mathcal{L}^n$ is a closed set in K^{n^3} and the point μ_0 is an element of $\mathcal{L}^n$. The law μ_0 is called a *contraction* (or *degeneration*) of the law μ. The operation of contraction is in a certain manner opposite to the deformation of the law μ_0. Contractions can be used for the determination of the closure of an orbit in the variety $\mathcal{L}^n$. This notion has been defined for the first time in the theoretical physics literature. Contractions can be met also in the theory of symmetries of differential equations.

Evidently any Lie algebra can be contracted to the Abelian Lie algebra.

EXAMPLE 10. Let $\mathfrak{g}$ be a Lie algebra with a law μ and let $\mathfrak{h}$ be a subalgebra of $\mathfrak{g}$. Consider a complementary subspace $\mathfrak{p} \subset \mathfrak{g}$. Denote by $\pi_{\mathfrak{h}}$ and $\pi_{\mathfrak{p}}$ the projections of $\mathfrak{g}$ on $\mathfrak{h}$ and on $\mathfrak{p}$ parallel to $\mathfrak{p}$ and $\mathfrak{h}$ respectively. We put

$$\eta(t) = \pi_{\mathfrak{h}} + t \cdot \pi_{\mathfrak{p}}.$$

It is clear that $\eta(t) \in \mathrm{GL}(n)$ for $t \neq 0$. A direct verification shows that the limit of the law μ_t defined by

$$\mu_t(X, Y) = \eta(t)\big(\mu\big((\eta(t))^{-1}(X), (\eta(t))^{-1}(Y)\big)\big)$$

for $t \to 0$ exists and this limit gives the semidirect sum $\mu_0 = \mathfrak{h} \oplus \mathfrak{p}$. The subalgebra $\mathfrak{h}$ acts on $\mathfrak{p}$ as the adjoint action on the space $\mathfrak{g}/\mathfrak{h}$. Contractions of this type are called IW-contractions and have been considered for the first time by Inonu and Wigner.

EXAMPLE 11. A generalization of IW-contractions was considered later by Saletan. For this one considers a family $\eta(t) \in \mathrm{GL}(n)$ of the form $A + t \cdot B$, where $A, B \in \mathrm{gl}(n)$ and $A + B \in \mathrm{GL}(n)$.

REMARK 1. It would be very important to know a simple class of contractions, such that any point μ_0 of the closure $\overline{\mathcal{O}(\mu)}$ of the orbit containing the law μ can be obtained as a contraction of μ. Unfortunately we do not know such a class.

References for this section: [13,18,22,31,41,45,49,50].

3. The components of $\mathcal{L}^n$

3.1. *Rigid Lie algebras*

In this section we suppose that $K = \mathbb{R}$ or $\mathbb{C}$. The variety $\mathcal{L}^n$ is the union of a finite number of irreducible components. The determination of the number $N\ (\mathcal{L}^n)$ of the irreducible components and their description is a very difficult problem for an arbitrary n. We only have the results for $n \leqslant 7$ and some estimates for the number $N\ (\mathcal{L}^n)$. We also have a local study of $\mathcal{L}^n$ for some special points. For an evaluation of the number $N\ (\mathcal{L}^n)$ one studies a certain class of Lie algebra laws with open orbits. The closure of such orbits give the irreducible components of $\mathcal{L}^n$.

DEFINITION. A Lie algebra law $\mu \in \mathcal{L}^n$ is called *rigid* if its orbit $O(\mu)$ is a Zariski open set.

This definition is valid for an arbitrary field K. In the case $K = \mathbb{R}$ or $\mathbb{C}$ we can also say that the law $\mu \in \mathcal{L}^n$ is *rigid* if its orbit $O(\mu)$ is an open set relative to the Euclidean topology.

We also consider the notion of rigid Lie algebra in the varieties $\mathcal{R}^n$ and $\mathcal{N}^n$.

Evidently, if μ is rigid, the Zariski closure of its orbit is an irreducible component of the variety $\mathcal{L}^n$. Thus, an estimate of the number of rigid laws in $\mathcal{L}^n$ gives an estimate of the number of irreducible components of $\mathcal{L}^n$.

THEOREM 12. *Let $\mu \in \mathcal{L}^n$ with $H^2(\mu, \mu) = 0$. Then μ is rigid.*

This theorem gives a sufficient condition for rigidity. It is valid also in the case of Lie algebra laws over the field $\mathbb{R}$. The reverse of this theorem is false. There are rigid Lie algebra laws μ with $H^2(\mu, \mu) \neq 0$. Many papers are devoted to the description of families of rigid Lie algebras with $H^2(\mu, \mu) \neq 0$. The first example of such a Lie algebra was constructed by Richardson.

EXAMPLE 13. Let $\mathfrak{g}_n$ be the semi-direct sum $\mathrm{sl}_2(\mathbb{C}) \oplus_{\varphi_n} \mathbb{C}^n$, corresponding to the irreducible representation $\varphi_n : \mathrm{sl}_2(\mathbb{C}) \to \mathrm{gl}_n(\mathbb{C})$. Then $H^2(\mathfrak{g}_n, \mathfrak{g}_n) \neq 0$, but $\mathfrak{g}_n$ is rigid if $n \equiv 1 \pmod 4$ and $n \geqslant 13$.

As a corollary from this theorem 12 we have that any semisimple Lie algebra is rigid. We also know that $H^2(\mu, \mu) = 0$ for a parabolic subalgebra law μ of a semisimple Lie algebra ($K = \mathbb{C}$). In particular a Borel subalgebra of a semisimple complex Lie algebra is rigid.

Some constructions of solvable rigid Lie algebras are given in the chapter "Nilpotent and solvable Lie algebras" of this Handbook (Volume 2).

The following theorem is useful for the constructions of rigid Lie algebras in the varieties $\mathcal{L}^n$ and $\mathcal{R}^n$.

THEOREM 14. *A rigid (in the variety $\mathcal{L}^n$ or $\mathcal{R}^n$) Lie algebra $\mathfrak{g}$ is always algebraic and corresponds to one of the following cases*:

(1) *the radical of $\mathfrak{g}$ is not nilpotent and* $\dim \mathrm{Der}\, \mathfrak{g} = \dim \mathfrak{g}$.
(2) *the radical of $\mathfrak{g}$ is nilpotent and one of the following conditions is valid*:
   - $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$;
   - $\mathfrak{g} = \mathbb{C} \oplus \mathfrak{p}$ *(direct sum), where $\mathfrak{p}$ is rigid, $\mathfrak{p} = [\mathfrak{p}, \mathfrak{p}]$ and any derivation of $\mathfrak{p}$ is an interior derivation*;
   - *the Lie algebra $\mathfrak{g}$ and all of its 1-codimensional ideals do not admit exterior semisimple derivations.*

An estimate of the number of rigid Lie algebra laws in the variety $\mathcal{L}^n$ ($K = \mathbb{C}$) is given by the following theorem.

THEOREM 15. *The number N ($\mathcal{L}^n$) of irreducible components of the variety $\mathcal{L}^n$ is at least of the order* $\exp(kn(\ln n)^{-1})$, *where $k = (\ln 2)^2/4$, for sufficiently large n.*

For the determination of the irreducible components the following result is useful.

THEOREM 16. *Let C be an irreducible closed set of $\mathcal{L}^n$ and μ a point of C such that* $\dim C = \dim Z^2(\mu, \mu)$. *Then*

(1) *μ is a simple point of the scheme $\mathcal{L}^n$*;
(2) *C is the only irreducible component containing μ.*

3.2. *The irreducible components of $\mathcal{L}^n$ and $\mathcal{R}^n$ in the low-dimensional case*

We suppose $K = \mathbb{C}$. In dimension $\leqslant 7$ we have a complete classification of the irreducible components of $\mathcal{L}^n$. For this classification we introduce the following notations:

1. For $i \in \mathbb{N}$, V_i designates the simple sl(2)-module of dimension i.
2. For $i \in \mathbb{N}$, $\mathfrak{a}_i$ designates the Abelian Lie algebra of dimension i.
3. For a natural number $i \geqslant 3$ we denote by $\mathfrak{l}_i$ the nilpotent Lie algebra of dimension i defined in the basis $(X_1, \ldots, X_n)$ by the relations

$$[X_1, X_k] = X_{k+1}, \quad k = 2, \ldots, i-1.$$

4. We denote by $\mathfrak{r}_2$ the solvable Lie algebra of dimension 2 defined by $[X_1, X_2] = X_2$.
5. For the 5-dimensional nilpotent Lie algebras we use the following designations:

$$\begin{array}{lllll}
\mathfrak{n}_1^5: & [X_1, X_2] = X_5, & [X_3, X_4] = X_5, & & \\
\mathfrak{n}_2^5: & [X_1, X_2] = X_4, & [X_1, X_3] = X_5, & & \\
\mathfrak{n}_3^5: & [X_1, X_2] = X_3, & [X_1, X_3] = X_4, & [X_2, X_5] = X_4, & \\
\mathfrak{n}_4^5: & [X_1, X_2] = X_3, & [X_1, X_3] = X_4, & [X_2, X_3] = X_5, & \\
\mathfrak{n}_5^5: & [X_1, X_2] = X_3, & [X_1, X_3] = X_4, & [X_1, X_4] = X_5, & \\
\mathfrak{n}_6^5: & [X_1, X_2] = X_3, & [X_1, X_3] = X_4, & [X_1, X_4] = X_5, & [X_2, X_3] = X_5.
\end{array}$$

6. Let $\mathfrak{p} = \mathfrak{s} \oplus_* \mathfrak{u}$ be a semidirect sum, where $\mathfrak{s}$ is semisimple Lie algebra and $\mathfrak{u}$ is a nilpotent Lie algebra. We design by $\mathcal{L}^n(\mathfrak{p})$ the subset of $\mathcal{L}^n$ formed by all Lie algebras $\mathfrak{g}$ (Lie algebras laws) with the following property: the largest ideal of $\mathfrak{g}$ whose radical is nilpotent is equal to $\mathfrak{p}$. In the case $\mathfrak{p} = \mathfrak{n}$ (that is, $\mathfrak{p}$ is a nilpotent Lie algebra) we have $\mathcal{L}^n(\mathfrak{p}) \subset \mathcal{R}^n$ and we denote this subset of $\mathcal{R}^n$ by $\mathcal{R}^n$ $(\mathfrak{n})$.
7. Let $S \subset \mathcal{L}^n$. We denote by S_0 the subset of S formed by all decomposable laws. We recall that a Lie algebra $\mathfrak{g}$ is called decomposable if $\mathfrak{g} = \mathfrak{s} \oplus_* (\mathfrak{t} \oplus_* \mathfrak{n})$ where $\mathfrak{t} \oplus_* \mathfrak{n}$ is the radical of $\mathfrak{g}$, $\mathfrak{n}$ is the nilradical, $\mathfrak{s}$ is a Levi subalgebra and $\mathfrak{t}$ is an Abelian reductive in $\mathfrak{g}$ subalgebra with $[\mathfrak{s}, \mathfrak{t}] = 0$.

THEOREM 17. *The irreducible components of* $\mathcal{L}^n$, $n \leqslant 7$, *are given by the following list*:

$n = 1$. $\mathcal{L}^1$ *is irreducible and contains only one point* $\{0\}$.

$n = 2$. $\mathcal{L}^2$ *is irreducible*: $\mathcal{L}^2 = \overline{\mathcal{O}(\mathfrak{a}_2)}$ (dim $= 2$).

$n = 3$. $\mathcal{L}^3$ *is the union of* 2 *irreducible components of dimension* 6 (*with one open orbit*):

- $\overline{\mathcal{O}(\mathrm{sl}(2))}$ (*the Zariski closure of the orbit relative to group* GL(2) *corresponding to law of* sl(2)),
- $\mathcal{R}^3 = \overline{\mathcal{R}_0^3(\mathfrak{a}_2)}$ (dim $= 3$).

$n = 4$. $\mathcal{L}^4$ *is the union of* 4 *irreducible components* (*with* 2 *open orbits*):

- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus \mathbb{C})}$ (dim $= 12$),
- $\overline{\mathcal{O}(\mathfrak{r}_2 \oplus \mathfrak{r}_2)}$ (dim $= 12$),
- $\overline{\mathcal{R}_0^4(\mathfrak{a}_3)}$ (dim $= 12$),
- $\overline{\mathcal{R}_0^4(\mathfrak{l}_3)}$ (dim $= 12$).

$n = 5$. $\mathcal{L}^5$ *is the union of* 7 *irreducible components* (*with 3 open orbits*):

- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus \mathfrak{r}_2)}$ (dim $= 20$),
- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus_* V_2)}$ (dim $= 19$),
- $\overline{\mathcal{O}(T_2 \oplus_* \mathfrak{l}_3)}$ (dim $= 20$), *where*

$$T_2 = \langle t_1, t_2 \rangle, \quad [t_1, X_1] = X_1, \quad [t_1, X_3] = X_3, \quad [t_2, X_2] = X_2, [t_2, X_3] = X_3,$$

- $\overline{\mathcal{R}_0^5(\mathfrak{a}_3)}$ (dim $= 21$),
- $\overline{\mathcal{R}_0^5(\mathfrak{a}_4)}$ (dim $= 20$),
- $\overline{\mathcal{R}_0^5(\mathfrak{l}_3 \oplus \mathfrak{a}_1)}$ (dim $= 20$),
- $\overline{\mathcal{R}_0^5(\mathfrak{l}_4)}$ (dim $= 20$).

$n = 6$. *$\mathcal{L}^6$ is the union of* 17 *irreducible components* (*with* 6 *open orbits*):

- $\overline{\mathcal{O}(\mathrm{gl}(2) \oplus_* V_2)}$ (dim $= 30$),
- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus \mathrm{sl}(2))}$ (dim $= 30$),
- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus_\theta \mathfrak{l}_3)}$ (dim $= 30$), *where* $\theta : \mathrm{sl}(2) \to \mathrm{Der}\, \mathfrak{l}_3$ *with* $\ker\theta = 0$,
- $\overline{\mathcal{O}(\mathfrak{r}_2 \oplus \mathfrak{r}_2 \oplus \mathfrak{r}_2)}$ (dim $= 30$),
- $\overline{\mathcal{O}(T_2 \oplus_* \mathfrak{l}_4)}$ (dim $= 30$), *where* $T_2 = \langle t_1, t_2 \rangle$ *and*

$$[t_1, X_1] = X_1, \quad [t_1, X_3] = X_3, \quad [t_1, X_4] = 2X_4,$$
$$[t_2, X_2] = X_2, \quad [t_2, X_3] = X_3, \quad [t_2, X_4] = X_4,$$

- $\overline{\mathcal{O}(T_1 \oplus_* \mathfrak{n}_6^5)}$ (dim $= 30$), *where* $T_1 = \langle t \rangle$ *and*

$$[t, X_i] = iX_i, \quad 1 \leqslant i \leqslant 5,$$

- $\overline{\mathcal{L}_0^6(\mathrm{sl}(2) \oplus \mathfrak{a}_2)}$ (dim $= 30$),
- $\overline{\mathcal{R}_0^6(\mathfrak{a}_4)}$ (dim $= 32$),
- $\overline{\mathcal{R}_0^6(\mathfrak{l}_3 \oplus \mathfrak{a}_1)}$ (dim $= 31$),
- $\overline{\mathcal{R}_0^6(\mathfrak{a}_5)}$ (dim $= 30$),
- $\overline{\mathcal{R}_0^6(\mathfrak{l}_3 \oplus \mathfrak{a}_2)}$ (dim $= 30$),
- $\overline{\mathcal{R}_0^6(\mathfrak{l}_4 \oplus \mathfrak{a}_1)}$ (dim $= 30$),
- $\overline{\mathcal{R}_0^6(\mathfrak{n}_i^5)}$, $i = 1, 2, 3, 4, 5$ (dim $= 30$).

$n = 7$. *$\mathfrak{L}^7$ has* 49 *irreducible components* (*with* 14 *open orbits*):

- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus_* V_4)}$ (dim $= 41$),
- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus_* (V_2 \oplus V_2))}$ (dim $= 37$),
- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus \mathfrak{r}_2 \oplus \mathfrak{r}_2)}$ (dim $= 42$),
- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus \mathrm{sl}(2) \oplus \mathfrak{a}_1)}$ (dim $= 42$),
- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus_* (T_2 \oplus_* V_3)}$ (dim $= 42$), *where* $T_1 = \langle t \rangle$, t *is a semisimple derivation of* $V_3 = \mathfrak{a}_3$,
- $\overline{\mathcal{O}(\mathrm{sl}(2) \oplus_* (T_1 \oplus_* \mathfrak{l}_3)}$ (dim $= 42$), *where* $T_1 = \langle t \rangle$, t *is a semisimple derivation of* $\mathfrak{l}_3$, *with* $t|_{\mathrm{sl}(2)} = 0$,
- $\overline{\mathcal{O}(\mathfrak{g}_i)}$, $i = 1, 2, \ldots, 8$ (dim $= 42$), *where*

$\mathfrak{g}_1 = (T_2 \oplus_* \mathfrak{l}_3) \oplus \mathfrak{r}_2$ (*see* $n = 5$ *for* $T_2 \oplus_* \mathfrak{l}_3$),
$\mathfrak{g}_2 = (T_2 \oplus_* \mathfrak{n}_3^5)$, *where*

$$[t_1, X_1] = X_1, \quad [t_1, X_3] = 2X_3, \quad [t_1, X_4] = X_4, \quad [t_1, X_5] = 2X_5,$$
$$[t_2, X_2] = X_2, \quad [t_2, X_4] = X_4, \quad [t_2, X_5] = X_5,$$

$\mathfrak{g}_3 = (T_2 \oplus_* \mathfrak{n}_4^5)$, *where*

$$[t_1, X_1] = X_1, \quad [t_1, X_3] = X_3, \quad [t_1, X_4] = 2X_4, \quad [t_1, X_5] = X_5,$$
$$[t_2, X_2] = X_2, \quad [t_2, X_3] = X_3, \quad [t_2, X_4] = X_4, \quad [t_2, X_5] = 2X_5,$$

$\mathfrak{g}_4 = (T_2 \oplus_* \mathfrak{n}_5^5)$, *where*

$$[t_1, X_1] = X_1, \quad [t_1, X_3] = X_3, \quad [t_1, X_4] = 2X_4, \quad [t_1, X_5] = 3X_5,$$
$$[t_2, X_2] = X_2, \quad [t_2, X_3] = X_3, \quad [t_2, X_4] = X_4, \quad [t_2, X_5] = X_5,$$

$\mathfrak{g}_5 = (T_1 \oplus_* \mathfrak{n}) = \langle t_1, X_1, \ldots, X_6 \rangle$, *where*

$$[X_1, X_2] = X_4, \quad [X_1, X_4] = X_5, \quad [X_1, X_5] = [X_2, X_3] = [X_2, X_4] = X_6,$$
$$[t_1, X_1] = X_1, \quad [t_1, X_2] = 2X_2, \quad [t_1, X_3] = 3X_3,$$
$$[t_1, X_4] = 3X_4, \quad [t_1, X_5] = 4X_5, \quad [t_1, X_6] = 5X_6,$$

$\mathfrak{g}_6 = (T_1 \oplus_* \mathfrak{n}) = \langle t_1, X_1, \ldots, X_6 \rangle$, *where*

$$[X_1, X_i] = X_{i+1}, \quad i = 2, 3, 4, 5; \quad [X_2, X_3] = X_6,$$
$$[t_1, X_1] = X_1, \quad [t_1, X_i] = (i+1)X_i, \quad 2 \leqslant i \leqslant 6,$$

$\mathfrak{g}_7 = (T_1 \oplus_* \mathfrak{n}) = \langle t_1, X_1, \ldots, X_6 \rangle$, *where*

$$[X_1, X_i] = X_{i+1}, \quad i = 2, 3, 4, 5; \quad [X_2, X_3] = X_5, \quad [X_2, X_4] = X_6,$$
$$[t_1, X_i] = iX_i, \quad 1 \leqslant i \leqslant 6,$$

$\mathfrak{g}_8 = (T_1 \oplus_* \mathfrak{n}) = \langle t_1, X_1, \ldots, X_6 \rangle$, *where*

$$[X_1, X_2] = X_3, \quad [X_1, X_3] = X_4, \quad [X_1, X_4] = [X_2, X_3] = X_5,$$
$$[X_2, X_5] = -[X_3, X_4] = X_6,$$
$$[t_1, X_6] = 7X_6, \quad [t_1, X_i] = iX_i, \quad 2 \leqslant i \leqslant 5,$$

- $\overline{\mathcal{L}_0^7(\mathrm{sl}(2) \oplus \mathfrak{a}_3)}$ $(\dim = 42)$,
- $\overline{\mathcal{L}_0^7((\mathrm{sl}(2) \oplus_* V_2) \oplus \mathfrak{a}_1)}$ $(\dim = 42)$,
- $\overline{\mathcal{L}_0^7(\mathrm{sl}(2) \oplus_* \mathfrak{l}_3)}$ $(\dim = 42)$,
- $\overline{\mathcal{R}_0^7(\mathfrak{p}_i)}$, $i = 1, 2, \ldots, 26$; $(\dim = 42)$, *where* $\mathfrak{p}_1, \ldots, \mathfrak{p}_{26}$ *are all* 6*-dimensional nilpotent Lie algebras, excepting the nilradicals* $\mathfrak{n}$ *of* $\mathfrak{g}_5, \mathfrak{g}_6, \mathfrak{g}_7, \mathfrak{g}_8$,
- $\overline{\mathcal{R}_0^7(\mathfrak{a}_4)}$ $(\dim = 44)$,
- $\overline{\mathcal{R}_0^7(\mathfrak{a}_5)}$ $(\dim = 45)$,
- $\overline{\mathcal{R}_0^7(\mathfrak{l}_3 \oplus \mathfrak{a}_2)}$ $(\dim = 44)$,

- $\overline{\mathcal{R}_0^7(\mathfrak{l}_4 \oplus \mathfrak{a}_1)}$ (dim = 43),
- $\overline{\mathcal{R}_0^7(\mathfrak{n}_1^5)}$ (dim = 43),
- $\overline{\mathcal{R}_0^7(\mathfrak{n}_2^5)}$ (dim = 43).

The demonstration of this classification theorem is based on the following construction: Let $\mathfrak{n}$ be a nilpotent Lie algebra. Denote by $\mathcal{R}_n(\mathfrak{n})$ the subspace of the variety $\mathcal{R}_n$ formed by the solvable Lie algebras (laws) whose nilradical is isomorphic to $\mathfrak{n}$. The subset of decomposable laws, that is the laws of Lie algebras $\mathfrak{g}$ which can be presented as semidirect sum $\mathfrak{g} = \mathfrak{a} \oplus_* \mathfrak{n}$ in the set $\mathcal{R}_n(\mathfrak{n})$ we denote by $\mathcal{R}_n^0(\mathfrak{n})$, where $\mathfrak{a}$ is an Abelian subalgebra in $\mathfrak{g}$. For $n \leqslant 7$ the closure of $\mathcal{R}_n^0(\mathfrak{n})$ in $\mathcal{R}_n(\mathfrak{n})$ coincides with $\mathcal{R}_n(\mathfrak{n})$ and forms an irreducible component of the variety $\mathcal{L}^n$.

References for this section: [9–11,22,29,38,47,48,53].

4. Subvarieties of nilpotent and solvable Lie algebra laws

4.1. *The variety $\mathcal{R}^n$*

The existing results on the variety of solvable Lie algebra laws $\mathcal{R}^n$ are similar to those for the variety $\mathcal{L}^n$. Mainly these results concern the rigid solvable Lie algebras. In this way we will describe some irreducible components of $\mathcal{R}^n$. In the low-dimensional case ($n \leqslant 7$) all irreducible components of $\mathcal{R}^n$ are also the irreducible components of $\mathcal{L}^n$. See the classification in the previous section. From this classification we have the following table ($K = \mathbb{C}$):

n	1	2	3	4	5	6	7
Number of irreducible components	1	1	1	3	5	13	40
Number of open orbits	1	1	0	1	1	3	8

The following construction is useful for the determination of rigid solvable Lie algebras with a given nilradical.

A vector $X \in \mathfrak{t}$ is called *regular*, if

$$\dim\bigl(\mathrm{Ker}(\mathrm{ad}\, X)\bigr) = \min\bigl\{\dim\bigl(\mathrm{Ker}(\mathrm{ad}\, Y)\bigr)\colon\ Y \in \mathfrak{t}\bigr\}.$$

As the endomorphism $\mathrm{ad}\, X$ is semisimple and $\mathfrak{n}$ is $\mathrm{ad}\, X$-invariant, there is a basis

$$(X_1, \dots, X_{p+q}, \dots, X_n = X)$$

formed by the eigenvectors of $\mathrm{ad}\, X$ such that $(X_1, \dots, X_{p+q})$ is a basis of $\mathfrak{n}$, $(X_{p+q+1}, \dots, X_n)$ is a basis of $\mathfrak{t}$, $(X_{p+1}, \dots, X_n)$ is a basis of $\mathrm{Ker}(\mathrm{ad}\, X)$.

Consider the system of linear equations, denoted by $S(X)$, relative to the variables $x_1, x_2, \dots, x_{n-1}$ associated to the vectors $X_1, X_2, \dots, X_{n-1}$. The equations of this system are

$$x_i + x_j = x_k$$

for all i, j, k such that the projection of the vector $[X_i, X_j]$ on X_k is nonnull.

THEOREM 18. *If a decomposable Lie algebra* $\mathfrak{g} = \mathfrak{t} \oplus \mathfrak{n}$ *is rigid, then* $\operatorname{rank}(S(X)) = \dim \mathfrak{n} - 1$ *for any regular vector* X.

COROLLARY 19. *If a decomposable Lie algebra* $\mathfrak{g} = \mathfrak{t} \oplus \mathfrak{n}$ *is rigid, than* $\mathfrak{t}$ *is a maximal torus of derivations of* $\mathfrak{n}$.

For some classes of nilpotent Lie algebras $\mathfrak{n}$ we have more profound results. We have a necessary and sufficient condition for rigidity of a solvable Lie algebra with a filiform nilradical.

For other results concerning rigid solvable Lie algebras see the chapter "Nilpotent and solvable Lie algebras" of this Handbook (Volume 2).

4.2. *The tangent space to* $\mathcal{N}_q^n$

We suppose that $K = \mathbb{R}$ or $\mathbb{C}$. Let $\mathfrak{g}$ be a nilpotent Lie algebra. The lower central series

$$\mathfrak{g} = \mathcal{C}^0(\mathfrak{g}) \supset \mathcal{C}^1(\mathfrak{g}) \supset \cdots \supset \mathcal{C}^p(\mathfrak{g}) = 0$$

and the upper central series

$$0 = \mathcal{C}_0(\mathfrak{g}) \supset \mathcal{C}_1(\mathfrak{g}) \supset \cdots \supset \mathcal{C}_p(\mathfrak{g}) = \mathfrak{g}$$

have the same length p called the nilpotency class of $\mathfrak{g}$. Setting

$$\begin{array}{llll} S_q = \mathfrak{g} & \text{for } q \leqslant 1, & S_q = \mathcal{C}^{q-1}(\mathfrak{g}) & \text{for } q > 1, \\ T_q = \mathfrak{g} & \text{for } q \leqslant 1, & T_q = \mathcal{C}_{p-q+1}(\mathfrak{g}) & \text{for } q > 1, \end{array}$$

we obtain two descending filtrations of $\mathfrak{g}$. Considering the filtration $\{S_q\}$ in the Lie algebra $\mathfrak{g}$ and the filtration $\{T_q\}$ in the $\mathfrak{g}$-module $\mathfrak{g}$ (relative to the adjoint action), we can write

$$T_q = \mathcal{C}_{p-q+1}(\mathfrak{g}) = \{x \in \mathfrak{g}\colon \mathfrak{g}^{p-q+1} \cdot x = 0\}$$

for all $q \leqslant p$. This means that $[S_r, T_q] \subset T_{r+q}$ for all $r, q \in \mathbb{Z}$ and we obtain a filtration in the set of cochains compatible with the coboundary operator. Thus, for each natural j we obtain descending filtrations

$$\{F_k Z^j(\mathfrak{g}, \mathfrak{g})\}, \qquad \{F_k B^j(\mathfrak{g}, \mathfrak{g})\}, \qquad \{F_k H^j(\mathfrak{g}, \mathfrak{g})\}$$

on the spaces of cocycles, coboundaries, and cohomology. In particular, the space $F_r Z^j(\mathfrak{g}, \mathfrak{g})$ consists of the cocycles g for which

$$g(S_{i_1}, S_{i_2}, \ldots, S_{i_j}) \subset T_{i_1+\cdots+i_j+r}$$

for any choice $1 \leqslant i_1, i_2, \ldots, i_j \leqslant p$, and $F_r H^j(\mathfrak{g}, \mathfrak{g})$ consists of all $\overline{g}$ for which there exists a representative g satisfying this condition.

THEOREM 20. *Let $\mathfrak{g}$ be a nilpotent Lie algebra of nilpotency class p and let $n-1 \geqslant q \geqslant p$. The Zariski tangent space to the scheme $\mathcal{N}_q^n$ at the point $\mathfrak{g}$ coincides with the space of cocycles $\varphi \in Z^2(\mathfrak{g}, \mathfrak{g})$ such that their cohomology classes are in $F_{q-p}H^2(\mathfrak{g}, \mathfrak{g})$.*

In the particular case when $\mathfrak{g}$ is a filiform Lie algebra (that is $p = n-1$) we have the following theorem.

THEOREM 21. *Let $\mathfrak{g}$ be an n-dimensional filiform Lie algebra. The tangent space to the scheme $\mathcal{N}^n$ at the point $\mathfrak{g}$ coincides with the space of cocycles $\varphi \in Z^2(\mathfrak{g}, \mathfrak{g})$ such that their cohomology classes are in $F_0H^2(\mathfrak{g}, \mathfrak{g})$.*

4.3. *Local study of the variety $\mathcal{N}^n$*

We suppose that $K = \mathbb{C}$. The local study of the variety $\mathcal{N}^n$ at the point corresponding to a nilradical of a Borel subalgebra of a simple Lie algebra $\mathfrak{g}$ gives

THEOREM 22. *Let $\mathfrak{n}$ be the nilradical of the Borel subalgebra of a simple Lie algebra $\mathfrak{g}$ of rank $r \geqslant 1$. Let $n = \dim \mathfrak{n}$ and let p be the nilpotency class of $\mathfrak{n}$. Then the scheme $\mathcal{N}_p^n$ is smooth at the point $\mathfrak{n}$. If $\mathfrak{g}$ is of type A_2, A_3 or B_2, the orbit $G(\mathfrak{n})$ is a Zariski open set in $\mathcal{N}_p^n$. The dimension of these orbits is equal respectively to 3, 25, 9. If $\mathfrak{g}$ is of type A_4, C_3, B_3, D_4 or G_2, the dimension of $\mathcal{N}_p^n$ at the point $\mathfrak{n}$ is equal to $m + n^2 - n - 2r + 1$, where m is respectively one of the numbers 7, 4, 4, 6, 1. In the other cases the dimension of $\mathcal{N}_p^n$ at the point $\mathfrak{n}$ is equal to $n^2 - n - 2r + 1 + \frac{1}{2}(r^2 + r - 2)$.*

THEOREM 23. *Let $\mathfrak{n}$ be the nilradical of the Borel subalgebra of a simple Lie algebra $\mathfrak{g}$. Then the tangent space to the scheme $\mathcal{L}^n$ at the point $\mathfrak{n}$ coincides with the tangent space to the scheme $\mathcal{L}_{\mathrm{red}}^n$ at the same point.*

REMARK 2. Let C be the irreducible component of the variety $\mathcal{N}_p^n$ containing $\mathfrak{n}$, where $n = \dim \mathfrak{n} \geqslant 8$, p is the nilpotency class of $\mathfrak{n}$. Then C contains a Zariski open set whose elements are characteristically nilpotent Lie algebras.

The study of derivations of Lie algebras obtained by deformations of $\mathfrak{n}$, gives the following corollaries. For the definition and some properties of characteristically nilpotent Lie algebras see the chapter "Nilpotent and solvable Lie algebras" in this Handbook (Volume 2). Other largest families of this class of Lie algebras can be found in the subvariety $\mathcal{F}^n$ of filiform Lie algebra laws.

COROLLARY 24. *The variety $\mathcal{N}^n$ with $n \geqslant 8$ possesses a locally closed subset of dimension $\geqslant n^2 + n^2/16 + n/4$ consisting of characteristically nilpotent Lie algebras.*

COROLLARY 25. *The pairwise nonisomorphic characteristically nilpotent Lie algebras of the component C depend on at least $n^2/16 - 3n/4 + 1$ parameters.*

References for this section: [3,10,15,16,22,25,26,28–33,35–37,40,44,54,55].

5. Local study of $\mathcal{N}^n$ for the filiform points

We suppose that $K = \mathbb{C}$. The local study of the variety $\mathcal{N}^n$ for the filiform points is sufficiently well developed. This study allows to describe some irreducible components of $\mathcal{N}^n$.

5.1. *Some examples of filiform Lie algebras*

An n-dimensional Lie algebra is called *filiform* if its nilpotency class is maximal, that is equal to $n-1$. The set $\mathcal{F}^n$ of filiform Lie algebra laws is a Zariski open set in the variety $\mathcal{N}^n$ and the Zariski closure of any irreducible component of $\mathcal{F}^n$ gives an irreducible component of $\mathcal{N}^n$. In this subsection we suppose that $n \geqslant 3$.

The simplest $(n+1)$-dimensional filiform Lie algebra is defined on the basis $(X_0, X_1, \dots, X_n)$ by the brackets:

$$[X_0, X_i] = X_{i+1}, \quad i = 1, \dots, n-1.$$

We denote this Lie algebra by L_n; the corresponding law we denote by μ_0.

Some other remarkable filiform Lie algebras are the following ones:

The Lie algebra Q_n $(n = 2k+1)$:

$$\begin{aligned} &[X_0, X_i] = X_{i+1}, && i = 1, \dots, n-1, \\ &[X_i, X_{n-i}] = (-1)^i X_n, && i = 1, \dots, k. \end{aligned}$$

In the basis $(Z_0, Z_1, \dots, Z_n)$, where $Z_0 = X_0 + X_1$, $Z_i = X_i$, $i = 1, \dots, n$; this Lie algebra is defined by

$$\begin{aligned} &[Z_0, Z_i] = Z_{i+1}, && i = 1, \dots, n-2, \\ &[Z_i, Z_{n-i}] = (-1)^i Z_n, && i = 1, \dots, k. \end{aligned}$$

The Lie algebra R_n:

$$\begin{aligned} &[X_0, X_i] = X_{i+1}, && i = 1, \dots, n-1, \\ &[X_1, X_j] = X_{j+2}, && j = 2, \dots, n-2. \end{aligned}$$

The Lie algebra W_n:

$$\begin{aligned} &[X_0, X_i] = X_{i+1}, \quad i = 1, \dots, n-1, \\ &[X_i, X_j] = \frac{6(i-1)!(j-1)!(j-i)}{(i+j)!} X_{i+j+1}, \quad 1 \leqslant i, j \leqslant n-2,\ i+j+1 \leqslant n. \end{aligned}$$

This Lie algebra can be defined also relative to a basis $(Y_1, Y_2, \dots, Y_{n+1})$ by the brackets

$$[Y_i, Y_j] = (j-i) Y_{i+j}, \quad i+j \leqslant n+1.$$

The Lie algebra T_n $(n = 2k)$:

$$
\begin{aligned}
&[X_0, X_i] = X_{i+1}, && i = 1, \ldots, n-1,\\
&[X_{k-i-1}, X_{k+i}] = (-1)^i X_n, && i = 0, 1, \ldots, k-2.
\end{aligned}
$$

The Lie algebra T_n $(n = 2k + 1)$:

$$
\begin{aligned}
&[X_0, X_i] = X_{i+1}, && i = 1, \ldots, n-1,\\
&[X_{k-i-1}, X_{k+i+j}] = (-1)^i C_{i+j}^i X_{n+j-1}, && i = 0, 1, \ldots, k-2,\ j = 0, 1.
\end{aligned}
$$

The Lie algebra P_n $(n = 2k)$:

$$
\begin{aligned}
&[X_0, X_i] = X_{i+1}, \quad i = 1, \ldots, n-1; \qquad [X_{k-1}, X_k] = X_n;\\
&[X_{k-i-1}, X_{k+i}] = (-1)^i \left(1 - \frac{2}{(k-1)(k-2)} C_{i+1}^{i-1}\right) X_n, \quad i = 1, \ldots, k-2;\\
&[X_{k-i-2}, X_{k+i+j-1}] = (-1)^i \frac{2}{(k-1)(k-2)} C_{i+j}^i X_{n+j-2},\\
&\quad i = 0, 1, \ldots, k-3;\ j = 0, 1.
\end{aligned}
$$

Here the C_j^i denote binomial coefficients.

THEOREM 26. *Let $\mathfrak{g}$ be a m-dimensional filiform Lie algebra ($m \geqslant 3$). The graded Lie algebra gr $\mathfrak{g}$ associated to the natural filtration of $\mathfrak{g}$ (filtration by descending central series) is isomorphic to L_{m-1}, if m is odd, and isomorphic to L_{m-1} or Q_{m-1}, if m is even.*

THEOREM 27. *Let $\mathfrak{g}$ be a m-dimensional filiform Lie algebra ($m \geqslant 12$) defined in the basis $(X_1, X_2, \ldots, X_m)$ by the brackets $[X_i, X_j] = c_{i,j} X_{i+j}$, $1 \leqslant i, j \leqslant m$, $i + j \leqslant m$. If m is odd and $\geqslant 15$, then $\mathfrak{g}$ is isomorphic to one of the Lie algebras L_n, R_n, W_n, T_n, P_n $(n = m - 1)$. If $m = 13$, then $\mathfrak{g}$ is isomorphic to one of the Lie algebras L_{12}, R_{12}, W_{12}, T_{12}. If m is even, $\mathfrak{g}$ is isomorphic to one of the Lie algebras L_n, R_n, W_n, T_n $(n = m - 1)$.*

5.2. *Deformations of the filiform Lie algebra L_n*

The following theorem shows that for the study of the variety $\mathcal{F}^n$ is suffices to describe the deformations of the simplest filiform Lie algebra L_n.

THEOREM 28. *Any $(n + 1)$-dimensional filiform Lie algebra is isomorphic to a linear deformation of L_n.*

We give an explicit construction of this linear deformation.

Let Δ be the set of pairs of integers (q,r) such that $1 \leqslant q \leqslant [\frac{n-1}{2}]$, $2q+1 \leqslant r \leqslant n$, $r \geqslant 4$. For any element $(q,r) \in \Delta$, we can associate a 2-cocycle for the Chevalley cohomology of L_n with coefficients in the adjoint module, denoted $\Psi_{q,r}$, and defined by

$$\Psi_{q,r}(X_i, X_j) = (-1)^{q-i} C_{j-q-1}^{q-i} (\operatorname{ad} X_0)^{i+j-2q-1} X_r$$

if $1 \leqslant i \leqslant q \leqslant j \leqslant n$ and $\Psi_{q,r}(X_i, X_j) = 0$ otherwise. We remark that this formula for $\Psi_{q,r}$ is uniquely determined by the conditions:

$$\Psi_{q,r}(X_q, X_{q+1}) = X_r,$$
$$\Psi_{q,r}(X_i, X_j) \in Z^2(L_n, L_n).$$

THEOREM 29. *The cohomology classes of the cocycles $\Psi_{q,r}$, where $(q,r) \in \Delta$, form a basis of the space $F_0 H^2(L_n, L_n)$.*

REMARK 3. We identify the space $F_0 H^2(L_n, L_n)$ with the subspace of $F_0 Z^2(L_n, L_n)$ engendered by the cocycles $\Psi_{q,r}$, where $(q,r) \in \Delta$.

COROLLARY 30. *The dimension of the space $F_0 H^2(L_n, L_n)$ is equal to $\frac{1}{4}(n^2 - 2n - 3)$, if n is odd, and $\frac{1}{4}(n^2 - 2n - 4)$, if n is even.*

COROLLARY 31. *The dimension of the space $F_0 Z^2(L_n, L_n)$ is equal to $\frac{1}{4}(3n^2 - 4n + 1)$, if n is odd, and $\frac{1}{4}(3n^2 - 4n)$, if n is even.*

Let Δ' be the set of pairs of integers (q,r) such that $1 \leqslant q \leqslant [\frac{n-1}{2}]$, $2q+1 < r \leqslant n$ (if n is odd we suppose that Δ' contains also the pair $(\frac{n-1}{2}, n)$). Then, any $(n+1)$-dimensional filiform Lie algebra law $\mu \in \mathcal{F}^{n+1}$ is isomorphic to $\mu_0 + \Psi$ where μ_0 is the law of L_n and Ψ is the 2-cocycle defined by

$$\Psi = \sum_{(q,r) \in \Delta'} a_{q,r} \Psi_{k,r}$$

and verifying the relation $\Psi \circ \Psi = 0$ with (as usual)

$$\Psi \circ \Psi(x,y,z) = \Psi\big(\Psi(x,y), z\big) + \Psi\big(\Psi(y,z), x\big) + \Psi\big(\Psi(z,x), y\big).$$

Let $\mathfrak{g}$ be an $(n+1)$-dimensional filiform Lie algebra with law μ. A basis $(X_0, X_1, \ldots, X_n)$ of $\mathfrak{g}$ is called *adapted*, if

$$\mu(X_i, X_j) = \mu_0(X_i, X_j) + \Psi(X_i, X_j), \quad 0 \leqslant i, j \leqslant n,$$

where

$$\Psi = \sum_{(q,r) \in \Delta'} a_{q,r} \Psi_{k,r}.$$

REMARK 4. In the case $n \geqslant 6$ we have $\Delta' \subsetneq \Delta$. Thus, the scheme $\mathcal{N}^n$ is not reduced.

5.3. *On the irreducible components of $\mathcal{N}^{n+1}$ meeting $\mathcal{F}^{n+1}$*

In this subsection we suppose that $n \geqslant 11$. In the case $n \leqslant 10$ we have a complete classification of filiform Lie algebras up to isomorphism and a description of the irreducible components of the $\mathcal{F}^{n+1}$.

Let V be the linear space engendered by the cocycles $\Psi_{q,r}$, where $(q,r) \in \Delta'$. We denote by M the affine algebraic variety in V, defined by the relation $\Psi \circ \Psi = 0$. With respect to a given adapted basis this variety is given by the polynomial relations in the variables $a_{q,r}$. The following theorem shows that the study of the variety $\mathcal{F}^n$ can be reduced to the study of M.

THEOREM 32. *Let C be an irreducible component of M. Then $\overline{G(\mu_0 + C)}$ is an irreducible component of $\mathcal{N}^{n+1}$. Its dimension is equal to $n^2 + \dim C$, and the mapping which associates C to its image in $\mathcal{N}^{n+1}$ is bijective on the set of irreducible components of $\mathcal{N}^{n+1}$ meeting the open set $\mathcal{F}^{n+1}$.*

PROOF. Let C be an irreducible component of M and C_1 an irreducible component of $\mathcal{N}^{n+1}$ containing the irreducible set $G(\mu_0 + C)$. Clearly $\mu_0 \in C_1$. Consider the open set U of $\mathcal{N}^{n+1}$ containing μ_0 and the laws μ such that the vectors

$$\{X_0, X_1, Y_{i+1}(\mu) = (\mathrm{ad}_\mu X_0)^i (X_1),\ i = 1, \ldots, n-1\}$$

are linearly independent. We have

$$\mu(X_1, Y_2(\mu)) = \sum_{i \geqslant 3} \alpha_i(\mu) Y_i(\mu),$$

where the functions $\alpha_i(\mu)$ are rational and defined on U. We put

$$X_1(\mu) = \alpha_3(\mu) X_0 - X_1,$$
$$X_{i+1}(\mu) = (\mathrm{ad}_\mu(X_0))^i X_1(\mu), \quad i \geqslant 1.$$

The elements $X_0, X_1(\mu), X_2(\mu), \ldots, X_n(\mu)$ determine an adapted basis for the Lie algebra μ and the mapping

$$\mu \to (X_0, X_1(\mu), X_2(\mu), \ldots, X_n(\mu))$$

is a rational mapping. We have

$$\mu \circ \mu = \varphi(\mu)(\mu_0 + \beta(\mu)),$$

where $\varphi(\mu) \in G$ and $\beta(\mu) \in M$. The mapping $\mu \to \beta(\mu)$ is rational. Thus, $\beta(U \cap \mathcal{C}_1)$ is an irreducible set containing $\mathcal{C}$. This implies

$$\beta(U \cap \mathcal{C}_1) = \mathcal{C}$$

and

$$U \cap \mathcal{C}_1 \subset G(\mu_0 + \mathcal{C}).$$

Then

$$\mathcal{C}_1 = G(\mu_0 + \mathcal{C}).$$

Conversely, if $\mathcal{C}_1$ is an irreducible component of $\mathcal{N}^{n+1}$ meeting the open set of filiform laws, then $\mathcal{C}_1$ contains μ_0. The same arguments show that if $\mathcal{C}$ is an irreducible component containing $\beta(U \cap \mathcal{C}_1)$, then

$$G(\mu_0 + \mathcal{C}) = \mathcal{C}_1.$$

Finally, if $\mathcal{C}$ is an irreducible component of M, we have

$$\begin{aligned}\dim\big(G(\mu_0 + \mathcal{C})\big) &= \dim G(\mu_0) + \dim \mathcal{C} = (n+1)^2 - \dim \operatorname{Der} L_n + \dim \mathcal{C} \\ &= n^2 + \dim \mathcal{C}.\end{aligned}$$

This gives the theorem. □

5.4. *On the reducibility of the variety $\mathcal{N}^{n+1}$, $n \geqslant 11$*

In this subsection we also suppose $n \geqslant 11$.

Consider the following closed subsets of M:

(1) M_1 is the closed subset of M defined by the relations

$$a_{2,6} = a_{3,8} = \cdots = a_{\frac{n-2}{2},n} = 0$$

if n is even, and by the relations

$$a_{2,6} = a_{3,8} = \cdots = a_{\frac{n-3}{2},n} = 0$$

if n is odd.

(2) M_2 is the closed subset of M defined by the relations

$$a_{m,2m+2} = \left(4 + \frac{6}{m}\right) a_{m+1,2m+4}, \quad m = 1, 2, \ldots, \left[\frac{n-4}{2}\right],$$

if n is even, and by the relations

$$a_{m,2m+2} = \left(4 + \frac{6}{m}\right) a_{m+1,2m+4}, \quad m = 1, 2, \ldots, \left[\frac{n-5}{2}\right], \quad a_{\frac{n-3}{2},n} = 0$$

if n is odd.

(3) M_3 is the closed subset of M defined by the relations

$$a_{1,4} = a_{2,6} = \cdots = a_{\frac{n-4}{2},n-2} = 0$$

if n is even, and by the relations

$$a_{1,4} = a_{2,6} = \cdots = a_{\frac{n-3}{2},n-1} = 0$$

if n is odd.

(4) M_4 is the closed subset of M defined by the relations

$$a_{1,4} = a_{2,6} = \cdots = a_{k-3,n-3} = 0, \quad a_{k-1,n} = \frac{1}{2}(k-1)(k-2)a_{k-2,n-2}$$

if $n = 2k \geqslant 14$, and by the relations

$$a_{1,4} = a_{2,6} = \cdots = a_{\frac{n-5}{2},n-2} = 0, \quad a_{\frac{n-1}{2},n} = 0$$

if n is odd.

The existence of the remarkable filiform Lie algebras of Section 5.1 shows, that the closed subsets M_1, M_2, M_3, M_4 of M are nontrivial subsets, strictly contained in M.

THEOREM 33. $M = M_1 \cup M_2 \cup M_3 \cup M_4$.

PROOF. Let $\psi \in M$. The equalities

$$(\psi \circ \psi)(X_1, X_2, X_3) = 0, \qquad (\psi \circ \psi)(X_1, X_3, X_4) = 0,$$
$$(\psi \circ \psi)(X_2, X_3, X_4) = 0,$$

give the following polynomial relations

$$\begin{aligned}
&-3a_{2,6}^2 + a_{2,6}a_{3,8} + 2a_{1,4}a_{3,8} = 0, \\
&6a_{3,8}^2 - 4a_{2,6}a_{3,8} - a_{3,8}a_{4,10} + 2a_{1,4}a_{4,10} - a_{2,6}a_{4,10} = 0, \\
&-4a_{3,8}^2 + 3a_{3,8}a_{4,10} + 3a_{2,6}a_{4,10} = 0.
\end{aligned}$$

The solution of this system is the union of three straight-lines in the 4-dimensional space parametrized by $(a_{1,4}, a_{2,6}, a_{3,8}, a_{4,10})$ whose parametrizations are $(t, 0, 0, 0)$, $(0, 0, 0, t)$ and $(t, \frac{t}{10}, \frac{t}{70}, \frac{t}{420})$.

By successively using the relations

$$\begin{aligned}&(\psi \circ \psi)(X_1, X_4, X_5) = 0,\\ &(\psi \circ \psi)(X_1, X_5, X_6) = 0,\\ &\dots\\ &(\psi \circ \psi)(X_1, X_k, X_{k+1}) = 0,\end{aligned}$$

where $k = [\frac{n-3}{2}]$, and using an induction, we find the assertions:

(a) If $n = 2k$ is even, the vector $(a_{1,4}, a_{2,6}, \dots, a_{k-1,n})$ is equal to one of the following vectors:

$$\begin{aligned}&a_{1,4}(1, 0, \dots, 0),\\ &a_{k-1,n}(0, 0, \dots, 1),\\ &a_{1,4}(\alpha_1, \alpha_2, \dots, \alpha_{k-1}), \quad \alpha_1 = 1, \quad \alpha_{m+1} = \tfrac{m}{4m+6}\alpha_m, \quad 1 \leqslant m \leqslant k-2,\\ &a_{k-1,n}\big(0, 0, \dots, \tfrac{2}{(k-1)(k-2)}, 1\big) \quad (\text{if } n \geqslant 14).\end{aligned}$$

(b) If $n = 2k+1$ is odd, the vector $(a_{1,4}, a_{2,6}, \dots, a_{k-1,n-1}, a_{k,n})$ is equal to one of the following vectors:

$$\begin{aligned}&a_{1,4}(1, 0, \dots, 0),\\ &a_{k,n}(0, 0, \dots, 1),\\ &a_{k-1,n-1}(0, 0, \dots, 1, 0),\\ &a_{1,4}(\alpha_1, \alpha_2, \dots, \alpha_{k-1}, 0), \quad \alpha_1 = 1, \quad \alpha_{m+1} = \tfrac{m}{4m+6}\alpha_m, \quad 1 \leqslant m \leqslant k-2.\end{aligned}$$

This gives the theorem. □

COROLLARY 34. *The variety $\mathcal{F}^m$ (and then also the variety $\mathcal{N}^m$), with $m \geqslant 13$, contains at least four irreducible components. The variety $\mathcal{F}^{12}$ (and then also the variety $\mathcal{N}^{12}$) contains at least three irreducible components.*

5.5. *Description of an irreducible component of $\mathcal{N}^{n+1}$ containing the Lie algebra R_n*

In this subsection we also suppose $n \geqslant 11$. We denote by U_1 the Zariski open subset of M_1, defined by the inequation $a_{1,4} \neq 0$ (we use the previous notations). It is clear that $R_n \in U_1$.

A cocycle $\Psi = \sum_{(q,r)\in\Delta'} a_{k,r}\Psi_{k,r}$ belonging to $F_0H^2(L_n, L_n)$ is called *nongenerated in layer* q_0, if $a_{q_0,s} \neq 0$ for some integer s.

LEMMA 35. *Let*

$$\Psi = \sum_{(q,r)\in\Delta'} a_{q,r}\Psi_{q,r} \in U_1$$

be a nongenerated cocycle in layer 2 *and let* s_0 *be the minimal value of the index* s *such that* $a_{2,s_0} \neq 0$. *Then*

(a) $a_{3,l} = 0$, *if* $l < 2s_0 - 4$;
(b) $a_{3,2s_0-4} = \frac{(s_0-3)a_{2,s_0}^2}{2a_{1,4}}$;
(c) $a_{k,2k+1+r} = 0$, *if* $k > 3$, $1 \leqslant r \leqslant 2s_0 - 11$.

PROOF. First, we show the following assertion:

$a_{q,s} = 0$ *if* $q \geqslant 3$ *and if the degree of homogeneousness of the cocycle* $\Psi_{q,s}$ *is less than* $2s_0 - 11$ (*that is if* $s - 2q - 1 < 2s_0 - 11$).

Suppose the converse. Let Ψ be a cocycle such that $a_{q,s} \neq 0$, where $r = s - 2q - 1 < 2s_0 - 11$. We can suppose that r is the minimal value satisfying this hypothesis and that q is the maximal value such that $s - 2q - 1 = r$. By writing that the coefficient of X_{2q+1+r} in $(\psi \circ \psi)(X_1, X_{q-1}, X_q)$ is equal to 0, we obtain $a_{1,4}a_{q,2q+1+r} = 0$. As $a_{1,4} \neq 0$, then $a_{q,2q+1+r} = 0$. This is impossible.

From this assertion we have the case (a). To verify point (b), it is sufficient to consider the relation $(\psi \circ \psi)(X_1, X_2, X_3) = 0$ and to write that the coefficient of X_{2s_0-4} is zero. To end the proof, we can show that $a_{k,2k+1+r} = 0$, if $k > 3$ and $r = 2s_0 - 11$. Let $k > 3$ and $r = 2s_0 - 11$. We consider the relation $(\psi \circ \psi)(X_1, X_{k-1}, X_k) = 0$. The assertion allows us to see that $a_{k,2k+1+r} = 0$. □

LEMMA 36. *Let*

$$\Psi = \sum_{(q,r)\in\Delta'} a_{q,r}\Psi_{q,r} \in U_1$$

be a nongenerated cocycle in layers 2 *and* 3, *and let* s_0, t_0 *be the minimal values of the index* s *and* t *such that* $a_{2,s_0} \neq 0$, $a_{3,t_0} \neq 0$ (*from the previous lemma, we have* $t_0 = 2s_0 - 4$). *Then*

(a) $a_{4,l} = 0$, *if* $l < 3s_0 - 8$;
(b) $a_{4,3s_0-8} = \frac{(2s_0-7)(s_0-3)a_{2,s_0}^3}{4a_{1,4}^2}$, *if* $3s_0 - 8 \leqslant n$.

The proof of this lemma is analogous to the previous one.

LEMMA 37. *Let*

$$\Psi = \sum_{(q,r)\in\Delta'} a_{q,r}\Psi_{q,r} \in U_1$$

be a cocycle of U_1. *Then* $a_{2,s} = 0$, *if* $s \leqslant \frac{n+7}{3}$.

PROOF. If $n < 14$, the lemma is obvious. Let $n \geqslant 14$. We suppose now that the index s satisfies $s \leqslant \frac{n+7}{3}$ with $a_{2,s} \neq 0$ and we denote by s_0 the minimal value of such index. The relation $(\psi \circ \psi)(X_1, X_3, X_5) = 0$ gives

$$(-t_0 + 4)a_{2,s_0}a_{3,t_0} + 2a_{1,4}a_{4,3s_0-8} = 0.$$

From the previous lemmas we have $a_{2,s_0} = 0$. □

COROLLARY 38. *The condition $a_{4,l} \neq 0$ is possible only if $l = n$ and $n \equiv 1 (\operatorname{mod} 3)$.*

This corollary is a consequence of the observation that the inequalities $s_0 > \frac{n+7}{3}$ and $3s_0 - 8 \leqslant n$ can only hold simultaneously if $s_0 = \frac{n+8}{3}$.

LEMMA 39. *We consider arbitrary scalars $\alpha_{1,4}, \alpha_{1,5}, \ldots, \alpha_{1,n}, \alpha_{2,s_0}, \ldots, \alpha_{2,n}$ satisfying $s_0 = [\frac{n+10}{3}]$ and $\alpha_{1,4} \neq 0$. Then there is only one cocycle*

$$\Psi = \sum_{(q,r)\in\Delta'} a_{q,r}\Psi_{q,r} \in U_1$$

such that $a_{1,i} = \alpha_{1,i}$ and $a_{2,j} = \alpha_{2,j}$.

PROOF. We put

$$\Psi = \sum_{s=4}^{n} a_{1,s}\Psi_{1,s} + \sum_{s=s_0}^{n} a_{2,s}\Psi_{2,s} + \sum_{s=s_0-4}^{n} a_{3,s}\Psi_{3,s} + x_{4,n}\Psi_{4,n}.$$

We choose the values $x_{3,s}$ in order that the condition $(\psi \circ \psi)(X_1, X_2, X_3) = 0$ is verified. The coefficients of $X_{2s_0-4}, X_{2s_0-3}, \ldots, X_n$ in the left hand part of this condition are zero. We obtain the values of the variables $x_{3,s}$, where $s = 2s_0 - 4, \ldots, n$. They have the desired form. In particular, we have

$$x_{3,2s_0-4} = \frac{(s_0-3)\alpha_{2,s_0}^2}{2\alpha_{1,4}}.$$

If $n \equiv 1 (\operatorname{mod} 3)$, then $s_0 = \frac{n+8}{3}$ and we put

$$x_{4,3s_0-8} = \frac{(2s_0-7)(s_0-3)\alpha_{2,s_0}^3}{4\alpha_{1,4}^2}.$$

In other cases, we put $x_{4,n} = 0$. The previous lemmas show that the equalities

$$(\psi \circ \psi)(X_1, X_3, X_4) = 0, \qquad (\psi \circ \psi)(X_2, X_3, X_4) = 0,$$
$$(\psi \circ \psi)(X_1, X_3, X_5) = 0$$

also are verified. The equalities

$$(\psi \circ \psi)(X_1, X_2, X_4) = 0, \qquad (\psi \circ \psi)(X_1, X_2, X_5) = 0,$$
$$(\psi \circ \psi)(X_1, X_2, X_6) = 0$$

are the consequence of the previous equalities and of $(\psi \circ \psi)(X_1, X_2, X_3) = 0$. The other relations $(\psi \circ \psi)(X_i, X_j, X_k) = 0$ can be verified directly. □

COROLLARY 40. *The variety $\overline{U_1}$ is irreducible and its dimension is equal to* $2n - [\frac{n+16}{3}]$.

COROLLARY 41. *The dimension of tangent space to the variety $\overline{U_1}$ at the point* 0 *is equal to* $3(n - [\frac{n+7}{3}])$ *if* $n \equiv 0$ *or* $2 (\mathrm{mod}\, 3)$ *and is equal to* $3(n - [\frac{n+7}{3}]) + 1$ *if* $n \equiv 1 (\mathrm{mod}\, 3)$.

COROLLARY 42. *The dimension of the tangent space at an arbitrary point of U_1 to the variety $\overline{U_1}$ is equal to* $2n - [\frac{n+16}{3}]$.

Finally we have the following theorem:

THEOREM 43. *Let $n \geqslant 11$. The point R_n of the variety $\mathcal{N}^{n+1}$ is a simple point. The dimension of the irreducible component C of $\mathcal{N}^{n+1}$ containing the Lie algebra R_n is equal to $n^2 + 2n - [\frac{n+16}{3}]$. Moreover, all the points $\mu_0 + \Psi$ of C with $\Psi \in U_1$ are simple points.*

5.6. *Description of an irreducible component of $\mathcal{N}^{n+1}$ containing the Lie algebra W_n*

In this subsection we also suppose $n \geqslant 11$. We denote by U_2 the Zariski open set in M_2 defined by $a_{1,4} \neq 0$.

LEMMA 44. *Let be $2 \leqslant r \leqslant n-3$. Then there is a cocycle $\Psi = \Psi_1 + \Psi_r + \Psi_{r+1} + \cdots + \Psi_{n-3}$ in U_2 such that $\Psi_r \neq 0$ and $\Psi_i \in H_i^2(L_n, L_n)$ (that is $\Psi_i(X_j, X_m) = \lambda X_{i+j+m}$, where $\lambda \in \mathbb{C}$).*

THE IDEA OF THE PROOF. Consider the Lie algebra law $\mu = \mu_0 + \Psi_1$ corresponding to W_n. For each $2 \leqslant r \leqslant n-3$ we construct an endomorphism $g \in G = \mathrm{GL}(n)$ such that $g(\mu) = \mu_0 + \Psi$, where Ψ satisfies the lemma. □

LEMMA 45. *Let be $2 \leqslant r \leqslant n-10$ and $n \geqslant 12$. Consider the cocycle φ_1 given by*

$$\varphi_1 = \sum a_{i,2i+2}\Psi_{i,2i+2},$$

where $a_{1,4} = 1$, $a_{j,2j+2} = (4 + \frac{6}{j})a_{j+1,2j+4}$, $j = 1, \ldots, [\frac{n-4}{2}]$. Then there is an unique nonnull solution (up a constant factor) of the equation $\varphi_1 \circ x + x \circ \varphi_1 = 0$ belonging to $H_r^2(L_n, L_n)$.

THE IDEA OF THE PROOF. We consider the cocycle

$$x = \sum_{1 \leqslant i \leqslant m} x_i \Psi_{i,2i+1+r} \in H_r^2(L_n, L_n)$$

with $m = [(n-r-1)/2]$ and with unknown coefficients $x_1, \ldots, x_m$. The condition $\varphi_1 \circ x + x \circ \varphi_1 = 0$ applied to the vectors (X_1, X_2, X_3), (X_1, X_3, X_4), (X_2, X_3, X_4), gives a linear homogeneous system of three equations with four unknowns (x_1, x_2, x_3, x_4). The rank of this system is equal to 3 for all values $r \geqslant 2$. Then the solutions of this system generate a 1-dimensional vector subspace in the 4-dimensional vector space. The relation $\varphi_1 \circ x + x \circ \varphi_1 = 0$ applied to the vectors (X_1, X_q, X_{q+1}) allows us to compute $x_5, x_6, \ldots, x_m$ from the parameters x_1, x_2, x_3, x_4. $\square$

The cocycles x verifying the lemma for r, $2 \leqslant r \leqslant n - 10$, we denote by φ_r.

The following lemma can be easily verified.

LEMMA 46. *Let $n - 9 \leqslant r \leqslant n - 7$ $(n \geqslant 11)$, and let φ_1 be the cocycle defined in the previous lemma. The dimension of the space of solutions of the equation $\varphi_1 \circ x + x \circ \varphi_1 = 0$ belonging to $H_r^2(L_n, L_n)$ is equal to* 2.

Let $n - 9 \leqslant r \leqslant n - 7$. We denote by φ_r and φ_r' the cocycles of $H_r^2(L_n, L_n)$ verifying the lemma and such that

$$\begin{aligned} &\varphi_r(X_1, X_2) = X_{r+3}, \qquad &&\varphi_r(X_2, X_3) = 0, \\ &\varphi_r'(X_1, X_2) = 0, &&\varphi_r'(X_2, X_3) = X_{r+5}. \end{aligned}$$

We put

$$\begin{aligned} &\varphi_{n-6} = \psi_{1,n-3}, \qquad &&\varphi_{n-5} = \psi_{1,n-2}, \qquad &&\varphi_{n-4} = \psi_{1,n-1}, \\ &\varphi_{n-3} = \psi_{1,n}, &&\varphi_{n-6}' = \psi_{2,n-1}, &&\varphi_{n-5}' = \psi_{2,n}. \end{aligned}$$

We can show that the cocycles $\varphi_1, \varphi_2, \ldots, \varphi_{n-3}, \varphi_{n-5}', \ldots, \varphi_{n-9}'$ generate the tangent space to the variety $\overline{U_2}$ at the point L_n. As $\mu_0 \in \overline{G(\mu)}$ for any filiform Lie algebra law μ (μ_0 is the law of L_n), we have the following theorem.

THEOREM 47. *Let $n \geqslant 11$. There exists a unique irreducible component of the variety $\mathcal{N}^{n+1}$ containing W_n. This component is smooth* (*each point is a simple one*) *and its dimension is equal to $n^2 + n + 2$.*

COROLLARY 48. *Let $n \geqslant 11$ and let $\mathcal{C}$ be the irreducible component of the variety $\mathcal{N}^{n+1}$ containing W_n. Then* $\dim \mathcal{C} - \dim G(W_n) = 5$.

This corollary is very interesting: the difference $\dim \mathcal{C} - \dim G(W_n)$ does not depend on the dimension of the Lie algebra W_n.

References for this section: [3,22–26,34–37,55].

6. A bound on the number of irreducible components of $\mathcal{N}^n$

The methods used for estimates of the number of irreducible components of $\mathcal{L}^n$ or $\mathcal{R}^n$ is not valid in the nilpotent case. We do not know of the existence of rigid Lie algebra laws in the variety $\mathcal{N}^n$ for large values of n. To obtain a bound we use the fact of the existence of irreducible components of dimensions to order n^2 and to order n^3 and we construct a family of Lie algebras belonging to the different irreducible components.

Let $\mathfrak{g}$ be the n-dimensional Lie algebra whose law μ is defined in the basis

$$(X_0, X_1, \dots, X_{m-1}, Y_1, \dots, Y_k, Z_1, \dots, Z_k, T_1, \dots, T_k),$$

where $n = 3k + m$ is sufficiently large, by

$$\begin{aligned} &\mu(X_0, X_i) = X_{i+1}, \quad && 1 \leqslant i \leqslant m-1, \\ &\mu(Y_i, Z_i) = T_i, && 1 \leqslant i \leqslant k. \end{aligned}$$

This Lie algebra is a direct sum of the model filiform Lie algebra L_{m-1} and of k copies of the 3-dimensional Heisenberg algebra. The derivation algebra of $\mathfrak{g}_n$ can be easily determinated from derivation algebras of L_{m-1} and of the Heisenberg algebra H_1 and we have the following lemma.

LEMMA 49. $\dim \operatorname{Der} \mathfrak{g}_n = 2k^2 + 8k + 2m - 1$.

COROLLARY 50. *The dimension of the orbit containing $\mathfrak{g}_n$ is equal to $7k^2 + 6km + m^2 - 8k - 2m + 1$.*

Consider the following 2-cocycles in $Z^2(\mathfrak{g}_n, \mathfrak{g}_n)$ defined by (non defined values being null, in particular $X_i = 0$ if $i \geqslant m$):

$\Psi_{q,s}(X_i, X_j) = (-1)^{q-i} C_{j-q-1}^{q-i} X_{i+j+s-2q-1}$, where $1 \leqslant q \leqslant [\frac{m-2}{2}]$, $2q + \frac{m-4}{2} \leqslant s \leqslant m-1$, $s \geqslant 4$.

$\varphi_{r,s,t}(Y_r, Z_s) = T_t$, where $1 \leqslant r, s, t \leqslant k$ and $r \neq s \neq t$, $r \neq t$.

$\eta_{r,s,t}(Y_r, Y_s) = T_t$, where $1 \leqslant r, s, t \leqslant k$ and $r \neq s \neq t$, $r \neq t$.

$\xi_{r,s,t}(Z_r, Z_s) = T_t$, where $1 \leqslant r, s, t \leqslant k$ and $r \neq s \neq t$, $r \neq t$.

$\rho_{t,u}(X_0, Y_t) = T_u$, where $1 \leqslant t, u \leqslant k$ and $t \neq u$.

$\rho'_{t,u}(X_0, Z_t) = T_u$, where $1 \leqslant t, u \leqslant k$ and $t \neq u$.

$\theta_{t,u}(X_1, Y_t) = T_u$, where $1 \leqslant t, u \leqslant k$ and $t \neq u$.

$\theta'_{t,u}(X_1, Z_t) = T_u$, where $1 \leqslant t, u \leqslant k$ and $t \neq u$.

$\vartheta_{t,u}(Y_t, Z_u) = X_{m-1}$, where $1 \leqslant t, u \leqslant k$ and $t \neq u$.

$\vartheta'_{t,u}(Y_t, Y_u) = X_{m-1}$, where $1 \leqslant t, u \leqslant k$ and $t \neq u$.

$\vartheta''_{t,u}(Z_t, Z_u) = X_{m-1}$, where $1 \leqslant t, u \leqslant k$ and $t \neq u$.

LEMMA 51. *Let Ω be the vector subspace in $Z^2(\mathfrak{g}_n, \mathfrak{g}_n)$ generated by the previous cocycles. Then, for all $\psi \in \Omega$, $\psi \neq 0$, the bilinear map $\mu + \psi$ is a Lie algebra law belonging to $\mathcal{N}^n$ and not isomorphic to μ.*

Indeed, it is easy to see that $\mu + \psi \in \mathcal{N}^n$. If $\mu + \psi$ is in the orbit containing μ, then ψ is a coboundary, that is $\psi = \delta f$ for an endomorphism $f \in \mathrm{gl}(n)$. As any cocycle of Ω is not a coboundary, the law $\mu + \psi$ are not isomorphic to μ.

Let $\mathcal{C}_0$ be an irreducible component of $\mathcal{N}^n$ passing through $\mathfrak{g}_n$ and containing the family $\mu + \psi$, $\psi \in \Omega$.

PROPOSITION 52. $\dim \mathcal{C}_0 \geqslant M(k,m)$, *where* $M(k,m) = 2k^3 + 7k^2 + \frac{17}{16}m^2 + 6km - 10k - \frac{15}{8}m + \frac{13}{16}$.

Indeed, $\dim \mathcal{O}(\mathfrak{g}_n) + \dim \Omega \geqslant M(k,m)$. From the previous lemma, this number is a minimum for the number of nonorbital parameters of $\mathcal{C}_0$.

LEMMA 53. *Let $\mathfrak{g}$ be a Lie algebra presented as a direct sum of its ideals*: $\mathfrak{g} = \mathfrak{i}_1 \oplus \mathfrak{i}_2 \oplus \mathfrak{i}_3$. *Let $X_i \neq 0 \in \mathfrak{i}_i$ and suppose that $X_3 \notin Z(\mathfrak{i}_3)$, where $Z(\mathfrak{i}_i)$ is the center of $\mathfrak{i}_i$. Then a bilinear alternated mapping $\varphi : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying $\varphi(X_1, X_2) = X_3$ does not belong to the space* $Z^2(\mathfrak{g}, \mathfrak{g})$.

In effect, consider $Y_3 \in \mathfrak{i}_3$ such that $\mu(X_3, Y) \neq 0$, where μ is the law of $\mathfrak{g}$. Then $\delta\varphi(X_1, X_2, Y) \neq 0$.

PROPOSITION 54. *Let $\mathcal{C}$ be an irreducible component of $\mathcal{N}^n$ containing $\mathfrak{g}_n$. Then*

$$\dim \mathcal{C} \leqslant N(k,m), \quad \text{where } N(k,m) = 2k^3 + 36k^2 + 2m^2 + 14km - 31k - 2m.$$

Indeed, for the filiform Lie algebra L_s the dimension of the space $Z^2(L_s, L_s)$ is $\leqslant 2s^2 - 2s$. By looking at the structure of $\mathfrak{g}_n$, we see that there exists an ideal $\mathfrak{i}$ isomorphic to L_{m-1}. On the other hand, the dimension of the components passing through $\mathfrak{g}_n$ is smaller than $\dim Z^2(\mathfrak{g}_n, \mathfrak{g}_n)$.

Using these propositions, we can give bounds on the dimensions of the component $\mathcal{C}_0$. Suppose now that $k \geqslant m$. Then, for n sufficiently large, we have:

$$N(k-7, m+21) < M(k,m).$$

If k varies between $\frac{n}{3}$ and $\frac{n}{4}$, we obtain a minoration of the number of (nonfiliform) components of $\mathcal{N}^n$. This bound is of the order $\frac{n}{74}$. Finally we have the following theorem.

THEOREM 55. *For n sufficiently large, the number of irreducible components of the variety $\mathcal{N}^n$ is at least of order n.*

References for this section: [2,22].

7. Characteristically nilpotent Lie algebras in the variety $\mathcal{N}^n$

The definition, the first properties and some results about the characteristically nilpotent Lie algebras are given in the chapter "Nilpotent and solvable Lie algebras" of this Handbook (Volume 2). Here we give a geometrical point of view of these results, by considering them as points of the variety $\mathcal{N}^n$.

THEOREM 56. *Let $n \geqslant 8$. Any irreducible component of $\mathcal{F}^n$ contains a nonempty Zariski open set, whose elements are characteristically nilpotent Lie algebra laws.*

For some families of nonfiliform characteristically nilpotent Lie algebra laws see Section 4.3 "Local study of the variety $\mathcal{N}^n$".

References for this section: [12,19,22,23,28,35–37].

8. The irreducible components of $\mathcal{N}^n$ in the low-dimensional case

THEOREM 57. *The irreducible components of $\mathcal{N}^n$, $n \leqslant 7$, are given by the following list:*

$n = 1$. *$\mathcal{N}^1$ is irreducible and contains only one point* $\{0\}$.

$n = 2$. *$\mathcal{N}^2$ is irreducible and contains only one point* $\{0\}$.

$n = 3$. *$\mathcal{N}^3$ is irreducible*: $\mathcal{N}^3 = \overline{\mathcal{O}(L_2)}$; $\dim \mathcal{N}^3 = 4$.

$n = 4$. *$\mathcal{N}^4$ is irreducible*: $\mathcal{N}^4 = \overline{\mathcal{O}(L_3)}$; $\dim \mathcal{N}^4 = 9$.

$n = 5$. *$\mathcal{N}^5$ is irreducible*: $\mathcal{N}^5 = \overline{\mathcal{O}(\mu)}$, *where* $\mu = \mu_0 + \Psi_{1,4}$; $\dim \mathcal{N}^5 = 17$.

$n = 6$. *$\mathcal{N}^6$ is irreducible*: $\mathcal{N}^6 = \overline{\mathcal{O}(\mu)}$, *where* $\mu = \mu_0 + \Psi_{1,4} + \Psi_{2,5}$; $\dim \mathcal{N}^6 = 28$.

$n = 7$. *$\mathcal{N}^7$ is reducible*: *$\mathcal{N}^7$ is the union of two irreducible components*:

- $\overline{\mathcal{O}(\mu)}$, *where* $\mu = \mu_0 + \alpha\Psi_{1,4} + \Psi_{2,6}$; *The orbit $\mathcal{O}(\mu)$ is a unique irreducible component of $\mathcal{F}^7$. The dimension of this component is equal to* 40.
- $\overline{\mathcal{O}(\mu(\alpha))}$, *where $\mu(\alpha)$ is the family of Lie algebra laws given in the basis* $(X_1, X_2, \ldots, X_7)$ *by* $\mu(X_1, X_i) = X_{i-1}$, $i = 4, 5, 6, 7$; $\mu(X_2, X_6) = X_3$; $\mu(X_2, X_7) = X_3 + X_4$; $\mu(X_5, X_7) = \alpha X_3$; $\mu(X_6, X_7) = \alpha X_4 + X_2$; *This component is not filiform*; *its dimension is equal to* 40.

REMARK 5. In dimension 8 there is a study of irreducible components of $\mathcal{N}^8$ using the computer software system and language MATHEMATICA.

In the filiform case we can describe the irreducible components of $\mathcal{F}^n$ for $1 \leqslant n \leqslant 11$. For $n \leqslant 7$ the variety $\mathcal{F}^n$ is irreducible.

THEOREM 58. *The irreducible components of $\mathcal{F}^n$, $8 \leqslant n \leqslant 11$, are given by the following list*:

$n = 8$. *$\mathcal{F}^8$ is the union of two irreducible components of dimension* 55:

- $\overline{\mathcal{O}(\mu(\alpha))}$, *where* $\mu(\alpha) = \mu_0 + \Psi_{1,4} + \alpha\Psi_{1,5} - 2\Psi_{2,6} + \Psi_{3,7}$;
- $\overline{\mathcal{O}(\mu(\alpha))}$, *where* $\mu(\alpha) = \mu_0 + \alpha\Psi_{1,4} + \Psi_{2,6} + \Psi_{2,7}$;

$n=9$. $\mathcal{F}^9$ *is irreducible*: $\mathcal{F}^9=\overline{\mathcal{O}(\mu(\alpha,\beta))}$; *where* $\mu(\alpha,\beta)=\mu_0+\Psi_{1,4}+\alpha\Psi_{2,6}+\Psi_{2,7}+\beta\Psi_{2,8}+\frac{3\alpha^2}{\alpha+2}\Psi_{3,8}$; $\alpha\neq 2$; $\dim\mathcal{N}^9=72$.

$n=10$. $\mathcal{F}^{10}$ *is the union of three irreducible components of dimension* 91:

- $\overline{\mathcal{O}(\mu(\alpha,\beta))}$, *where* $\mu(\alpha,\beta)=\mu_0+\Psi_{1,4}+\alpha\Psi_{2,6}+\Psi_{2,7}+\beta\Psi_{2,8}+\frac{3\alpha^2}{\alpha+2}\Psi_{3,8}+\frac{14\alpha-2\alpha^2}{(\alpha+2)^2}\Psi_{3,9}$, $\alpha\neq-2$.
- $\overline{\mathcal{O}(\mu(\alpha,\beta,\gamma))}$, *where* $\mu(\alpha,\beta,\gamma)=\mu_0+\Psi_{1,4}-\Psi_{2,6}+\alpha\Psi_{2,7}+\beta\Psi_{2,8}+\gamma\Psi_{2,9}+3\Psi_{3,8}+(\beta-16\alpha)\Psi_{3,9}+\Psi_{4,9}$;
- $\overline{\mathcal{O}(\mu(\alpha,\beta,\gamma))}$, *where* $\mu(\alpha,\beta,\gamma)=\mu_0+\Psi_{1,4}+\Psi_{2,6}+\alpha\Psi_{2,7}+\beta\Psi_{2,8}+\gamma\Psi_{2,9}+\Psi_{3,8}+\frac{\beta+4\alpha}{3}\Psi_{3,9}+\Psi_{4,9}$.

$n=11$. $\mathcal{F}^{11}$ *is the union of two irreducible components of dimension* 112:

- $\overline{\mathcal{O}(\mu(\alpha,\beta,\gamma,\delta))}$, *where* $\mu(\alpha,\beta,\gamma,\delta)=\mu_0+\Psi_{1,4}+\alpha\Psi_{2,6}+\Psi_{2,7}+\beta\Psi_{2,8}+\gamma\Psi_{2,9}+a_{3,8}\Psi_{3,8}+a_{3,9}\Psi_{3,9}+a_{3,10}\Psi_{3,10}+a_{4,10}\Psi_{4,10}$, $\alpha\neq-2,\pm1$.
  In this formula we can express $a_{3,8}$, $a_{3,9}$, $a_{3,10}$ *and* $a_{4,10}$ *in terms of the other parameters*

$$
\begin{aligned}
a_{3,8} &= \frac{3\,a_{2,6}^2}{2\,a_{1,4}+a_{2,6}},\\
a_{3,9} &= \frac{a_{2,6}\,(-9\,a_{1,5}a_{2,6}+14\,a_{1,4}a_{2,7}-2\,a_{2,6}a_{2,7})}{(2\,a_{1,4}+a_{2,6})^2},\\
a_{3,10} &= \frac{Q}{2\,(2\,a_{1,4}+a_{2,6})^3\,({a_{1,4}}^2-{a_{2,6}}^2)},\\
a_{4,10} &= \frac{3\,(4\,a_{1,4}-7\,a_{2,6})\,{a_{2,6}}^3}{2\,(2\,a_{1,4}+a_{2,6})({a_{1,4}}^2-{a_{2,6}}^2)},
\end{aligned}
$$

  where

$$
\begin{aligned}
Q = {} & 54\,{a_{1,4}}^2{a_{1,5}}^2{a_{2,6}}^2-48\,{a_{1,4}}^3a_{1,6}{a_{2,6}}^2+24\,{a_{1,4}}^2a_{1,6}{a_{2,6}}^3\\
&-54\,{a_{1,5}}^2{a_{2,6}}^4-12\,a_{1,4}a_{1,6}{a_{2,6}}^4-18\,a_{1,6}{a_{2,6}}^5-84\,{a_{1,4}}^3a_{1,5}a_{2,6}a_{2,7}\\
&+66\,{a_{1,4}}^2a_{1,5}{a_{2,6}}^2a_{2,7}+84\,a_{1,4}a_{1,5}{a_{2,6}}^3a_{2,7}-66\,a_{1,5}{a_{2,6}}^4a_{2,7}\\
&+32\,{a_{1,4}}^4{a_{2,7}}^2-52\,{a_{1,4}}^3a_{2,6}{a_{2,7}}^2-12\,{a_{1,4}}^2{a_{2,6}}^2{a_{2,7}}^2\\
&+52\,a_{1,4}{a_{2,6}}^3{a_{2,7}}^2+64\,{a_{1,4}}^4a_{2,6}a_{2,8}-20\,{a_{2,6}}^4{a_{2,7}}^2-8\,{a_{1,4}}^3{a_{2,6}}^2a_{2,8}\\
&-60\,{a_{1,4}}^2{a_{2,6}}^3a_{2,8}-22\,a_{1,4}{a_{2,6}}^4a_{2,8}-{a_{2,6}}^5a_{2,8},
\end{aligned}
$$

- $\overline{\mathcal{O}(\mu(\alpha,\beta,\gamma,\delta))}$, *where*

$$
\begin{aligned}
\mu(\alpha,\beta,\gamma,\delta) = {} & \mu_0+\Psi_{1,5}+\alpha\Psi_{1,6}+\beta\Psi_{2,7}\\
&+(-4\beta^2+3\beta\gamma+3\gamma)\Psi_{2,8}+\gamma\Psi_{2,9}+\delta\Psi_{3,9}+a_{4,10}\Psi_{4,10}.
\end{aligned}
$$

References for this section: [1,3,4,20–22].

References

[1] J.M. Ancochéa-Bermudez and M. Goze, *On the varieties of nilpotent Lie algebras of dimension* 7 *and* 8, J. Pure Appl. Algebra **77** (1992), 131–140.
[2] J.M. Ancochéa-Bermudez, M. Goze and Y. Hakimjanov (Khakimdjanov), *Sur les composantes de la variété des algèbres de Lie nilpotentes*, C. R. Acad. Sci. Paris **314** (1) (1992), 657–660.
[3] J.M. Ancochéa-Bermudez, M. Goze and Y. Hakimjanov (Khakimdjanov), *Sur la réductibilité de la variété des algèbres de Lie nilpotentes complexes*, C. R. Acad. Sci. Paris, Sér. I **313** (1991) 59–62.
[4] J.M. Ancochéa-Bermudez, J.R. Gomez, M. Goze and G. Valeiras, *Sur les composantes irreductibles de la variété des lois d'algèbres de Lie nilpotentes,* J. Pure Appl. Algebra **106** (1996), 11–22.
[5] J.M. Ancochéa-Bermudez and M. Goze, *Le rang des systèmes linéaires d'une algèbre de Lie résoluble rigide*, Comm. Algebra **20** (3) (1992), 875–887.
[6] N. Bourbaki, *Groupes et Algèbres de Lie*, Ch. I, Hermann (1960).
[7] N. Bourbaki, *Groupes et Algèbres de Lie*, Ch. VII–VIII, Hermann (1975).
[8] F. Bratzlavsky, *Une algèbre de Lie caractéristiquement nilpotente de dimension* 6, C. R. Acad. Sci. Paris, Sér. A **276** (1973), 1035–1037.
[9] R. Carles and Y. Diakité, *Sur les variétés d'algèbres de Lie de dimension* $\leqslant 7$, J. Algebra **91** (1984), 65–82.
[10] R. Carles, *Sur la structure des algèbres de Lie rigides,* Ann. Inst. Fourier (Grenoble) **33** (3) (1984), 65–82.
[11] R. Carles, *Sur certaines classe d'algèbres de Lie rigides,* Math. Ann. **272** (1985), 477–488.
[12] R. Carles, *Sur les algèbres de Lie caractéristiquement nilpotentes*, Preprint No. 5, Univ. de Poitiers (1984).
[13] C. Conatser, *Contraction of low-dimensional Lie algebras,* J. Math. Phys. **13** (2), 197–203.
[14] J. Dixmier, *Sur les représentations unitaires de groupes de Lie nilpotentes III,* Canad. J. Math. **10** (1958), 321–348.
[15] J. Dixmier and W.G. Lister, *Derivations of nilpotent Lie algebras,* Proc. Amer. Math. Soc. **8** (1957), 155–158.
[16] J.L. Dyer, *A nilpotent Lie algebra with nilpotent automorphism group*, Bull. Amer. Math. Soc. **8** (1970), 155–158.
[17] G. Favre, *Une algèbre de Lie caractéristiquement nilpotente de dimension* 7, C. R. Acad. Sci. Paris, Sér. A **274** (1972), 1338–1339.
[18] M. Gerstenhaber, *On the deformation of rings and algebras,* Ann. Math. **74** (1) (1964), 59–103.
[19] J.R. Gómez, M. Goze and Yu. Khakimdjanov, *On the k-Abelian filiform Lie algebras,* Comm. Algebra **25** (2) (1997), 431–450.
[20] J.R. Gómez, A. Jiménez-Merchan and Yu. Khakimdjanov, *On the variety of nilpotent Lie algebra laws of dimension* 11, Rend. Sem. Fac. Sci. Univ. Cagliari **66** (2) (1996), 137–142.
[21] J.R. Gómez, A. Jiménez-Merchan and Yu. Khakimdjanov, *Low-dimensional filiform Lie algebras*, J. Pure Appl. Algebra **30** (1998), 133–158.
[22] M. Goze and Yu. Khakimdjanov, *Nilpotent Lie Algebras*, MIA 361, Kluwer Acad. Publ., Dordrecht (1996).
[23] M. Goze and Yu. Khakimdjanov, *Sur les algèbres de Lie nilpotentes admettant un tore de dérivations,* Manuscripta Math. **84** (1994), 115–224.
[24] M. Goze and Yu. Khakimdjanov, *Some nilpotent Lie algebras and its applications,* Algebra and Operator Theory, Proceedings of the Colloquium in Tashkent, Kluwer Acad. Publ. (1998), 49–64.
[25] M. Goze and J.M. Ancochéa-Bermudez, *Algèbres de Lie rigides*, Proc. Kon. Nederl. Acad. **88** (4) (1985), 397–415.
[26] M. Goze and J.M. Ancochéa-Bermudez, *Algèbres de Lie rigides*, Algebra and Operator Theory, Proceedings of the Colloquium in Tashkent, Kluwer Acad. Publ. (1998), 65–91.
[27] N. Jacobson, *Lie Algebras*, Intersc. Tracts, No. 10, Wiley, New York (1962).
[28] N. Jacobson, *A note on automorphisms and derivations of Lie algebras*, Proc. Amer. Math. Soc. **6** (1955), 281–283.
[29] Yu. Khakimdjanov, *On the cohomology of parabolic subalgebras of semisimple Lie algebras,* Uspekhi Mat. Nauk **32** (6) (1977), 237–238 (Russian).
[30] Yu. Khakimdjanov, *The variety of structures of nilpotent Lie algebras,* Uspekhi Mat. Nauk (6) (1990), 159–160 (Russian).
[31] Yu. Khakimdjanov (Hakimjanov), *Cohomologie et déformations de certaines algèbres de Lie nilpotentes*, Comm. Algebra **16** (10) (1989), 2149–2192.

[32] Yu. Khakimdjanov (Hakimjanov), *Sur les variétés d'algèbres de Lie,* Comm. Algebra **18** (4) (1990), 1147–1187.
[33] Yu. Khakimdjanov, *On the cohomology of some nilpotent Lie algebras*, Mat. Sb. **180** (5) (1989), 658–675 (Russian). English transl. in Math. USSR-Sb. **67** (1990).
[34] Yu. Khakimdjanov, M. Goze and Sh. Ayupov (eds), *Algebra and Operator Theory, Proceedings of the Colloquium in Tashkent*, Kluwer Acad. Publ. (1998).
[35] Yu. Khakimdjanov, *Variété des lois d'algèbres de Lie nilpotentes*, Geometriae Dedicata **40** (1991), 229–295.
[36] Yu. Khakimdjanov, *Characteristically nilpotent Lie algebras*, Mat. Sb. **181** (5) (1990), 642–655 (Russian). English transl. Math. USSR Sb. **70** (1) (1991).
[37] Yu. Khakimdjanov, *On the characteristically nilpotent Lie algebras*, Algebra i Logika **28** (6) (1989), 722–736.
[38] A.A. Kirillov and Yu. Neretin, *The Variety A_n of n-Dimensional Lie Algebra Structures*, Amer. Math. Soc. Transl. (2) (1987).
[39] G. Leger and E. Luks, *Cohomology of nilradicals of Borel subalgebras,* Trans. Amer. Math. Soc. **195** (1974), 305–316.
[40] G. Leger and S. Tôgô, *Characteristically nilpotent Lie algebras,* Duke Math. J. **26** (4) (1959), 623–628.
[41] M. Levy Nahas, *Deformations and contractions of a Lie algebra,* J. Math. Phys. **8** (6) (1967), 1211–1222.
[42] E. Luks, *What is the typical nilpotent Lie algebra*?, Computers in Nonassociative Rings and Algebras, Academic Press (1977).
[43] V.V. Morosov, *Classification of nilpotent Lie algebras,* Izv. Vyssh. Uchebn. Zaved. **4** (1958), 161–171 (Russian).
[44] O.A. Nielsen, *Unitary representations and coadjoint orbits of low-dimensional nilpotent Lie groups,* Quenn's Papers in Pure Appl. Math. **63** (1983).
[45] A. Nijenhuis and R.W. Richardson, *Cohomology and deformation of algebraic structures,* Bull. Amer. Math. Soc. **70** (1964), 406–411.
[46] A. Nijenhuis and R.W. Richardson, *Deformation of Lie algebra structures,* J. Math. Mech. **17** (1967), 89–105.
[47] G. Rauch, *Effacement et déformation,* Ann. Inst. Fourier (Grenoble) **22** (1) (1972).
[48] R.W. Richardson, *On the rigidity of semi-direct products of Lie algebras,* Pacific J. Math. **22** (2) (1967), 339–344.
[49] R.W. Richardson, *Deformation of Lie algebra structures*, J. Math. Mech. **17** (1) (1967), 89–105.
[50] E. Saletan, *Contraction of Lie groups,* J. Math. Phys. **2** (1) (1966), 1–22.
[51] C. Seeley, *Degenerations of* 6*-dimensional nilpotent Lie algebras on* $\mathbb{C}$, Comm. Algebra **18** (10) (1990), 3493–3505.
[52] J.P. Serre, *Lie Algebras and Lie Groups*, W.A. Benjamin, New York (1965).
[53] A.K. Tolpigo, *On the cohomology of parabolic Lie algebras,* Mat. Zametki **12** (3) (1972), 251–255.
[54] M. Vergne, *Variété des algèbres de Lie nilpotentes,* Thèse 3ème cycle, Paris (1966).
[55] M. Vergne, *Cohomologie des algèbres de Lie nilpotentes. Application à l'étude de la variété des algèbres de Lie nilpotentes,* Bull. Soc. Math. France **98** (1970), 81–116.

This Page Intentionally Left Blank

Section 4D
Varieties of Algebras, Groups, ...

This Page Intentionally Left Blank

Varieties of Algebras

V.A. Artamonov

Department of Mechanics and Mathematics, Moscow State University, Moscow, 119899 GSP Russia

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

A well known definition says that a group is a set X with a binary operation of multiplication xy, a nullary operation fixing a unity e and a unary operation x^{-1}. These three operations satisfy the identities $(xy)z = x(yz)$, $xe = ex = x$, $xx^{-1} = x^{-1}x = e$. In a similar way one can define a ring, a semigroup, an associative (Lie, Jordan, power-associative) algebra. There is a common element in the structure of all these definitions — classes of groups, semigroups, etc. are defined by a given type (a set of operational symbols) and by a given set of identities. This leads to the fruitful idea of the study of classes of general algebras which are defined by identities.

Many classes of algebras are defined by identities. For example, classes of Abelian groups, nilpotent and soluble groups (alternative, Lie, Jordan algebras) of bounded degree, of distributive, modular lattices, of periodic groups of bounded exponent and so on.

Varieties are defined in terms of identities. In order to give a precise definition of an identity we shall start with another approach to the theory of varieties based on G. Birkhoff's theorem. According to this theorem varieties are exactly the non-empty classes of algebras of a given type which are closed under subalgebras, homomorphic images and direct products.

The bibliography presented at the end of this paper certainly is not complete. Mostly it contains monographs and surveys. It should be mentioned that this text reflects the interests of the author and is influenced by his point of view on the whole subject.

It will be assumed that the reader is familiar with the notions of universal algebra such as algebras, subalgebras, congruences, homomorphisms, products and so on.

1. General theory

1.1. *(Pre)varieties, free algebras*

By a *class of algebras* we mean a class of algebras of a fixed type T. A class K is *abstract* if it is closed under isomorphisms of algebras. This means that if $A \in K$ and $B \simeq A$, then $B \in K$. For any class of algebras K denote:

SK is the class of all algebras isomorphic to subalgebras of algebras of K;

HK is the class of all homomorphic images of algebras of K;

PK is the class of algebras isomorphic to direct products of algebras of K.

It is quite clear that each class SK, HK, PK is abstract. If K is any class of algebras we shall call each member of K a K-algebra.

DEFINITION 1. A class K of algebras of a type T is a *pre-variety* if $\mathrm{SK} = \mathrm{PK} = K$ and K contains a one-element algebra E.

Examples of pre-varieties:

(i) the class of all commutative rings with unity;

(ii) the class of all (Abelian) groups;

(iii) the class of all torsion-free (Abelian) groups;

(iv) a class of all algebras M from a pre-variety V such that for a given set K of simple algebras no K-algebra is embeddable in a member of M.

Note that if V is the class of all (Abelian) groups, K is the class of all cyclic groups of prime order, then the class M, in (iv) is exactly the class of all torsion-free (Abelian) groups in (iii).

The notion of a pre-variety was introduced by A.I. Mal'cev [Mal1,Mal2] under the name of *replica complete* class. An explanation of this name is provided by

THEOREM 2. *For an abstract class K of a type T the following are equivalent:*

(i) *K is a pre-variety;*

(ii) *for any algebra A of a type T there exists an algebra $RA \in K$ and a surjective homomorphism $f : A \twoheadrightarrow RA$ such that for any homomorphism $g : A \to B$, $B \in K$, there exists a unique homomorphism $h : RA \to B$ such that $g = hf$.*

If M is an arbitrary class of algebras then $\mathrm{SPM} \cup E$ is the smallest pre-variety containing M, where E is a one-element algebra. It is called a *pre-variety generated by M*. It follows from Theorem 2 that for any algebra A and any pre-variety K the congruence $K(A) = \operatorname{Ker} f$ is the smallest congruence in A such that $A/K(A) \in K$. The congruence $K(A)$ is *completely invariant*. This means that if a is an endomorphism of A and a pair (x, y) belongs to $K(A)$, then $(a(x), a(y)) \in K(A)$. The congruence $K(A)$ is called the *verbal congruence* in A associated with K. Note that if K_1, K_2 are two pre-varieties, then $K_1 \subseteq K_2$ if and only if $K_1(A) \supseteq K_2(A)$ for any algebra A.

For example, let K be the pre-variety of all Abelian groups and let G be an arbitrary group. Then any congruence in G is determined by some normal subgroup in G. It is not difficult to show that the congruence $K(G)$ corresponds in this sense to the commutator subgroup $[G, G]$ in G.

Any pre-variety K is closed under subdirect products, Boolean products and inverse limits since these constructions can be obtained by taking direct products and subalgebras.

Assume that a pre-variety K is non-trivial. This means that K contains an algebra which has more than one element.

DEFINITION 3. Let X be a set. An algebra $F_K(X)$ in K is a *free K-algebra over* (*base*) X if

(i) $F_K(X)$ is generated by the set X;

(ii) if f is an arbitrary map from X to any K-algebra A then f can be extended to a homomorphism $F_K(X) \to A$.

It follows from (i) that the extension of f to a homomorphism $F_K(X) \to A$ is unique. The algebra $F_K(X)$ is determined by X (and K) uniquely up to isomorphism. The cardinality X is called the *rank* of $F_K(X)$.

EXAMPLES.

(1) Let K be the pre-variety of all semigroups. Then the free K-semigroup $F_K(X)$ on X is the semigroup of all words in the alphabeth X.

(2) Let K be the pre-variety of all (Abelian) groups. Then $F_K(X)$ is the free (Abelian) group over X.

(3) Let k be a commutative associative ring with unity and let K be the pre-variety of all associative k-algebras with unity. Let V be a free k-module with basis X. Then $F_K(X)$ is the tensor algebra $T(V)$.

(4) Let k be as in (3). Consider the pre-variety K of all commutative associative k-algebras with unity. Then $F_K(X)$ is the polynomial algebra $k[X]$ with set of variables X.

(5) Let k be as in (3). Consider the pre-variety K generated by a matrix algebra $\mathrm{Mat}(n, k)$. Let R be the polynomial algebra over k with a set of variables z_{ijx}, $i, j = 1, \dots, n$, $x \in X$. Then the subalgebra in $\mathrm{Mat}(n, R)$-generated by matrices

$$Z_x = (z_{ijx}) \in \mathrm{Mat}(n, R), \quad x \in X,$$

is the free K-algebra with a base $\{Z_x \mid x \in X\}$.

THEOREM 4. *Let K be a non-trivial pre-variety. Then for any set X there exists a free K-algebra $F_K(X)$ over X.*

Free algebras play a crucial role in the general theory of (pre)varieties. This is explained by

THEOREM 5. *Let K be a non-trivial pre-variety. If $A \in K$, then there exists a set X such that A is a homomorphic image of $F_K(X)$. In particular, A is isomorphic to a quotient-algebra of $F_K(X)$.*

Take a finite set $X = \{x_1, \dots, x_n\}$. According to Definition 3 each element of $F_K(X)$ can be considered as an n-ary term operation in any K-algebra in the following sense. Let $t \in F_K(X)$, $A \in K$ and $a_1, \dots, a_n \in A$. Consider the map $f : X \to A$ such that $f(x_i) = a_i$, $i = 1, \dots, n$. Then f can be extended to a homomorphism $f : F_K(X) \to A$. In this case put $t(a_1, \dots, a_n) = f(t)$.

A pair (t_1, t_2) of elements of $F_K(X)$ is *an identity* in a K-algebra A if $t_1(a_1, \dots, a_n) = t_2(a_1, \dots, a_n)$ for all elements $a_1, \dots, a_n$ in A.

If Y is a subset of the set X then the subalgebra in $F_K(X)$ generated by Y is isomorphic to $F_K(Y)$. Moreover, $F_K(X)$ is the set-theoretic union of subalgebras $F_K(Y)$, where Y runs over all finite subsets in X. Thus any identity $t_1 = t_2$ can be viewed as an element of an algebra $F_K(X)$ with a countable base X.

DEFINITION 6. A *variety* V is a class of algebras closed under taking subalgebras, homomorphic images and direct products, that is $V = \mathrm{SV} = \mathrm{HV} = \mathrm{PV}$.

If M is an arbitrary class of algebras, then $V = \mathrm{HSP}(M)$ is a smallest variety containing M. HSP(M) is called the *variety generated by the class M*. Any variety is a pre-variety. Hence if V is a non-trivial variety then for any set X there exists a free V-algebra over X.

THEOREM 7 (G. Birkhoff [Gr, p. 171]). *Let W be a variety of algebras. A subclass V in W is a (sub)variety if and only if there exists a set of identities $f_i = g_i$, $i \in I$, such that V consists of all W-algebras satisfying all these identities.*

In this theorem it is supposed that $f_i, g_i,\ i \in I$, are elements of a free W-algebra $F_W(X)$ with a countable base $X = \{x_1, x_2, \ldots\}$.

THEOREM 8. *Let a pre-variety W contain a non-trivial pre-variety V. Then $F_W(X)/V(F_W(X)) \simeq F_V(X)$ for any set X, where $V(F_W(X))$ is the verbal congruence in $F_W(X)$ associated with V.*

THEOREM 9. *Let W be a variety and V be a subvariety in W. If X is a countable set, then the verbal congruence $V(F_W(X))$ consists of all pairs (f, g) such that $f = g$ is an identity in V. In particular, if V_1, V_2 are subvarieties in W, then $V_1 \subseteq V_2$ if and only if $V_1(F_W(X)) \supseteq V_2(F_W(X))$.*

It follows from Theorems 8 and 9 that each variety V is uniquely determined by its free algebra $F_V(X)$ of countable rank. Note that if K is a non-trivial pre-variety, then HK is the variety generated by K. Moreover, for any set X, the free K-algebra on X is the free HK-algebra over X.

Let the class K consists of a single algebra A. There is a standard representation for the free algebras in HSP(K). If Y is a set, then the direct power A^Y can be viewed as the set of all maps $f: Y \to A$. Thus for a given set X and a fixed element $x \in X$ one can define the element

$$m_x \in A^{(A^X)}$$

by the map $m_x : A^X \to A$, given by $m_x(u) = u(x),\ u \in A^X$. Note that $A^{(A^X)}$ as a direct power is a K-algebra. If we interprete elements of this power as maps $A^X \to A$, then the basic operations act pointwise on these maps.

THEOREM 10 [C,Gr]. *Let F be the subalgebra in $A^{(A^X)}$ generated by all elements $m_x, x \in X$. If $|A| \geqslant 2$, then F is the free algebra in* HSP(K) *with basis $\{m_x \mid x \in X\}$.*

DEFINITION 11. A variety V is *locally finite* if every finitely generated V-algebra is finite.

A variety V is locally finite if and only if the free V-algebras of finite ranks are finite. It follows from Theorem 10 that if A is a finite algebra, then the variety generated by A is locally finite.

Examples of locally finite varieties:
(i) the variety of periodic nilpotent groups;
(ii) the variety of Boolean algebras;
(iii) the variety of distributive lattices.

Note that the varieties of all (Abelian) group, and of all modular lattices are not locally finite [S].

DEFINITION 12. An algebra F in a variety V is *relatively free* if there exists a generating set X in F with the following property. Any map $X \to F$ can be extended to an endomorphism of F.

THEOREM 13. *For an algebra F in V the following are equivalent*:

(i) *F is relatively free*;

(ii) *F is a free algebra in the subvariety* HSP(F);

(iii) *$F \simeq F_V(X)/c$, where $F_V(X)$ is the free V-algebra with a base X, c is a completely invariant cogruence in $F_V(X)$.*

There exists another approach to the notion of a variety of algebras. By a *clone* C we mean a family of sets $\{C_n,\ n \geqslant 0\}$ with a partial operation of superposition

$$C_n \times \underbrace{C_m \times \cdots \times C_m}_{n} \to C_m,$$

$$(f, g_1, \ldots, g_n) \to f(g_1, \ldots, g_n) \in C_m, \qquad f \in C_n,\ g_i \in C_m,$$

and with fixed elements $p_{in} \in C_n,\ n \geqslant 1,\ i = 1, \ldots, n$. The operation of superposition satisfies the *super-associative law*

$$f\big(g_1(h_1, \ldots, h_m), \ldots, g_n(h_1, \ldots, h_m)\big) = \big(f(g_1, \ldots, g_n)\big)(h_1, \ldots, h_m),$$

where $f \in C_n,\ g_i \in C_m,\ h_j \in C_k$. The elements p_{in}, which are called *projections*, satisfy identities

$$f = f(p_{1n}, \ldots, p_{nn}), \qquad p_{in}(g_1, \ldots, g_n) = g_i,$$

where $f \in C_n,\ g_j \in C_m$.

The main example of clones is the clone $O(X)$ of all finitary operations on a set X. By definition $O(X)$ is the family of sets $O_n(X)$, where $O_n(X)$ is a set of all maps (n-ary operations) $X^n \to X$. If $f \in O_n(X),\ g_1, \ldots, g_n \in O_m(X)$ and $x \in X^m$, then by definition

$$\big(f(g_1, \ldots, g_m)\big)(x) = f\big(g_1(x), \ldots, g_m(x)\big).$$

The element p_{in} is the i-th projection $X^n \to X$, that is

$$p_{in}(x_1, \ldots, x_n) = x_i,$$

where $x_1, \ldots, x_n \in X$.

Given a clone C and a set A consider a homomorphism of clones $f : C \to O(A)$. This is a collection of maps $f_n : C_n \to O_n(A),\ n \geqslant 0$, commuting with superpositions and preserving projections. Then each element of C_n acts as an n-ary operation in A. Thus A becomes an algebra of a type C. A class of all these C-algebras A form a variety V_C. The clone C can be recovered from V_C as the clone of all term operations in V_C. More precisely, if F is a free V_C-algebra of a countable rank, then C is isomorphic to the subclone in $O(F)$ generated by the basic operations. This subclone consists of all term operations of V_C. Moreover, a similar argument shows that every variety V can be presented as $V = V_C$ for some C. For details see [S, Vol. II, Ch. VI], [Pl,Sz].

This approach enables us to introduce the notions of rational equivalence and interpretations of varieties.

DEFINITION 14. Let V, U be varieties of algebras of types T and S. The varieties V and U are *rationally equivalent* (in the sense of A.I. Mal'cev) if their clones of term operations are isomorphic.

For example, the variety of Boolean algebras is rationally equivalent to the variety of Boolean rings.

THEOREM 15 [Mal2]. *For two varieties U, V the following are equivalent:*

(i) *V and U are rationally equivalent;*

(ii) *there exists an equivalence of categories $V \to U$ preserving the underlying sets of algebras.*

DEFINITION 16. Let V, W be varieties of algebras of types T and S with clones of term operations $C(V)$ and $C(W)$ respectively. Then V is *interpreted (presented) in* W (notation $V \leqslant W$) if there exists a homomorphism of clones $C(V) \to C(W)$.

THEOREM 17 [GT, p. 17]. *Let V, W be as in Definition* 16. *The following, are equivalent:*

(i) *V is interpreted in W;*

(ii) *there exists a functor $W \to V$ which commutes with the underlying set functor (preserving underlying sets of algebras).*

EXAMPLE. Let LIE be the variety of all Lie algebras, ASS be the variety of all associative algebras, JORD be the variety of all Jordan algebras. Then LIE and JORD are interpreted in ASS. This follows from the fact that any associative algebra is a Lie (Jordan) algebra with respect to the multiplication $[x, y] = xy - yx$ (respectively, $x \circ y = xy + yx$). For further details consult [A3, pp. 55–56], [GT,Pl], [S, Vol. II, Ch. VI].

Finally it is worth it to mention a categorical approach to the definition of varieties due to F. Lawvere [A1,Man,Pl]. Let T be a category with direct products whose objects are the non-negative integers and whose morphisms are generated by

$$\bigcup_{n\geqslant 0} \mathrm{Hom}(n, 1). \tag{$*$}$$

If n, m are objects of T, then it is assumed that the object $n + m$ is the direct product of the objects n and m. By a *T-algebra* we shall mean a product-preserving functor F from T to the category of sets. It is clear that a T-algebra F is an algebra in the usual sense on the underlying set $F(1)$ with $(*)$ as set of basic operations. In fact, if $f : n \to 1$ in T, then f induces a map

$$F(f) : F(n) = F(1)^n \to F(1).$$

Morphisms of T-algebras are exactly the morphisms of functors. The category of T-algebras is actually a variety of $(*)$-algebras in the previous sense, since identities correspond to the commutativity of suitable diagrams of morphisms in the category T. For further details see [Man,Pl].

1.2. *Mal'cev conditions. Congruences. Generation of varieties*

Let W be a variety of algebras of a type T and $F_W(X)$ be the free W-algebra with countable basis $X = \{x_1, x_2, \ldots\}$. By a *strong Mal'cev condition* we mean a formulae

$$(\exists p_1)\ldots(\exists p_n)\big((f_1 = g_1)\&\cdots\&(f_m = g_m)\big), \qquad (**)$$

where f_i, g_i are elements of $F_W(X)$ which contain only elements of X and functional symbols $p_1, \ldots p_n$. A class of subvarieties in W satisfying $(**)$ forms a *strong Mal'cev class* of subvarieties.

There is another equivalent definition of a strong Mal'cev class. Let K be a variety of a finite type S which is defined by a finite set of identities. Then a strong Mal'cev class in W consists of all subvarieties V in W such that K can be presented in V.

A *Mal'cev class* in W is a class of subvarieties V in W such that there exists a chain $K_1 \geqslant K_2 \geqslant \cdots$ with the following properties:

(i) each K_i is of a finite type and is defined by a finite set of identities;

(ii) $K_n \leqslant V$ for some $n \geqslant 1$.

Here "$\leqslant$" means an interpretation.

A weak Mal'cev class is any class which is an intersection of countably many Mal'cev classes. These definitions and the appelation "Mal'cev classes" come from (see [Mal1, Mal2,Sm]).

THEOREM 18 (A.I. Mal'cev). *Let V be a variety of algebras. The following are equivalent*:

(i) *in each V-algebra congruences permute*;

(ii) *there exists a ternary term $p(x, y, z)$ such that*

$$p(x, x, y) = p(y, x, x) = y$$

are identities in V.

Thus congruence-permutable varieties form a strong Mal'cev class.

DEFINITION 19. A variety V is *congruence-distributive* (*congruence-modular*) if the congruence lattice of every V-algebra is distributive (respectively, modular).

THEOREM 20 (B. Jonsson [Jo]). *A variety V is congruence-distributive if and only if for some non-negative integer n there are terms $d_0(x, y, z), \ldots, d_n(x, y, z)$ such that V satisfies the identities*

$$d_0(x, y, z) = x, \qquad d_i(x, y, x) = x,$$
$$d_{2i}(x, y, y) = d_{2i+1}(x, y, y),$$
$$d_{2i+1}(x, x, y) = d_{2i+2}(x, x, y),$$
$$d_n(x, y, z) = z.$$

EXAMPLES. The variety of all lattices is congruence distributive. In fact, take $n = 2$, $d_0(x, y, z) = x$, $d_1(x, y, z) = (x \wedge y) \vee (y \wedge z) \vee (z \wedge x)$, $d_2(x, y, z) = z$.

Another example is a variety of median algebras. This is the variety of algebras with one ternary operation $m(x, y, z)$ satisfying the identities $m(x, x, y) = m(x, y, x) = m(y, x, x) = x$. This variety is congruence-distributive. In fact, take $n = 2$, $d_0(x, y, z) = x$, $d_1(x, y, z) = m(x, y, z)$, $d_2(x, y, z) = z$.

THEOREM 21 (Gumm [FM,HM]). *A variety V is congruence-modular if and only if for some non-negative integer n there exist terms $d_0(x, y, z), \ldots, d_n(x, y, z), p(x, y, z)$ such that the following identities hold in V*

$$
\begin{aligned}
&d_0(x, y, z) = x, \qquad d_i(x, y, x) = x,\\
&d_{2i}(x, y, y) = d_{2i+1}(x, y, y),\\
&d_{2i+1}(x, x, y) = d_{2i+2}(x, x, y),\\
&d_n(x, y, y) = p(x, y, y), \qquad p(x, x, y) = y.
\end{aligned}
$$

Any congruence-permutable variety is congruence-modular. Thus varieties of groups, rings, modules, quasigroups are congruence-modular.

Another Mal'cev type characterization of congruence-modular varieties in terms of quaternary terms was obtained by A. Day [Gr, p. 355]. It follows from Theorems 20 and 21 that classes of congruence-distributive and congruence-modular varieties form strong Mal'cev classes. A general theory of Mal'cev classes is developed in [Sm]. The following theorem is one of the basic result in this theory. It is due to R. Freese, J.B. Nation, A. Pixley and R. Wille [HM, p. 143].

THEOREM 22. *Let I be an equation for lattices. The class of varieties whose congruences satisfy I is a weak Mal'cev class which is defined in terms of identities $f_i = g_i$ such that each of f_i and g_i has at most one occurences of an operation symbol.*

Further results on Mal'cev classes can be found in [A1, pp. 201–202], [A3, pp. 61–64], [Gr, §60], [HM, §9], [Sm]. A detailed survey on congruence-distributive varieties can be found in [Pi2], and one on congruence-modular varieties in [Pi1].

In connection with these consideration it is necessary to mention a result of S.V. Polin [Po].

THEOREM 23. *There exists a variety P of algebras with the following properties*:

(i) *the congruence lattices of P-algebras belong to a proper lattice variety*;

(ii) *there exists an algebra in P with a non-modular congruence lattice*;

(iii) *if $A \in P$, then there exists an epimorphism f from the congruence lattice* $\operatorname{Con} A$ *onto a distributive lattice D such that each lattice $f^{-1}(x)$, $x \in D$, is distributive.*

THEOREM 24 (A. Day and R. Freese [DF]). *Let K be a non-congruence-modular variety. Then the lattice variety generated by all congruence lattices* $\operatorname{Con} B$, $B \in K$, *contain all lattices* $\operatorname{Con} A$, $A \in P$.

Congruence-modular varieties are very close to varieties of groups and rings. For example, commutator theory can be extended to this class of varieties. Let A be an algebra in a congruence-modular variety V and $c, c' \in \operatorname{Con} A$. Denote by $[c, c']$ a set of all pairs $(x, y) \in A^2$ such that the quadriple $((x, x), (y, y))$ belongs to the congruence in $c \subseteq A^2$ generated by all $((a, a), (b, b)) \in c \times c$, where $(a, b) \in c'$.

THEOREM 25 [FM]. *$[c, c']$ is a congruence in A. The commutator of the congruences has the following properties*:

(i) $[c, c'] = [c', c] \subseteq c \wedge c'$;

(ii) $[c, \vee c'_i] = \vee[c, c'_i]$;

(iii) *if $f: B \to A$ is an epimorphism (surjective homomorphism) in V, then* $f^{-1}[c, c'] = [f^{-1}c, f^{-1}c']$.

Assume that there is a binary operation $\langle c, c'\rangle$ with the properties (i)–(iii) in $\operatorname{Con} A$ for each V-algebra A. Then:

THEOREM 26 [FM]. $\langle c, c'\rangle \subseteq [c, c']$.

Let A be an algebra in a congruence modular variety V. Denote by 1^A and 0^A the greatest and the smallest congruences in A. Then we can by induction introduce a lower central and derived series of congruences in A. Put $1^A_1 = 1^A$, $1^A_{k+1} = [1^A_k, 1^A]$, $1^A_{(1)} = 1^A$, $1^A_{(k+1)} = [1^A_{(k)}, 1^A_{(k)}]$.

DEFINITION 27. The chain of congruences $1^A_1 \geqslant 1^A_2 \geqslant \cdots$ is the *lower central series* in A. The chain of congruences $1^A_{(1)} \geqslant 1^A_{(2)} \geqslant \cdots$ is *the derived series* in A. An algebra A is nilpotent of class at most n if $1^A_{n+1} = 0^A$. An algebra A is soluble of class at most n if $1^A_{(n+1)} = 0^A$.

All nilpotent algebras (soluble algebras) of class at most n form a subvariety V_n (respectively $V_{(n)}$) in V. Algebras in $V_1 = V_{(1)}$ are called *Abelian*:

DEFINITION 28. An algebra B in V is *affine* if there exists a structure of an Abelian group on B with an addition $+$ such that each n-ary term t, $n \geqslant 0$, induces an affine map $t: B^n \to B$, that is $t(x + y) + t(0) = t(x) + t(y)$ for all $x, y \in B^n$.

Note that in this situation the map $x \to t(x) - t(0)$ is a homomorphism of Abelian groups $B^n \to B$.

THEOREM 29 [FM]. *In a congruence-modular variety an algebra is Abelian if and only if it is affine.*

EXAMPLES. In a group each congruence is determined by a normal subgroup. Hence, if the congruences c, c' correspond to normal subgroups N, N', then $[c, c']$ corresponds to the normal subgroup $[N, N']$. Thus Definition 27 can be considered as a generalization of the concepts of nilpotency and solvability in group theory.

In rings congruences correspond to ideals. Therefore, if the congruences c, c' in a ring R correspond to ideals I, J, then $[c, c']$ corresponds to the ideal $IJ + JI$. Hence, nilpotence and solvability in Definition 27 generalize corresponding notions in ring theory. For further details and results consult [A3,FM,HM,MMT,Pi1], [S, Vol. II, Ch. VI].

DEFINITION 30. An algebra A in an arbitrary variety V is *subdirectly irreducible* if $\operatorname{Con} A$ has a unique atom.

For example, if p is a prime then the residue rings $Z/p^k Z$ are subdirectly irreducible. They are not simple, if $k > 1$. Note that any simple algebra is subdirectly irreducible.

THEOREM 31 (G. Birkhoff [Gr]). *Let $A \in V$. Then A is a subdirect product of subdirectly irreducible algebras in V. In particular, every variety is generated by its subdirectly irreducible members, namely $V = \mathrm{SP}(V_{\mathrm{SI}})$, where V_{SI} is the class of subdirectly irreducible members.*

DEFINITION 32. A variety V is *residually small* if cardinalities of members in V_{SI} are bounded.

For example, the varieties of Abelian groups, semilattices, distributive lattices, and Boolean algebras are all residually small.

THEOREM 33 ([Pi1, Ch. 4], [BS]). *A congruence-distributive variety generated by a finite algebra is residually small.*

THEOREM 34 [Gr, pp. 366–367]. *A variety V is residually small if and only if every V-algebra is embeddable in an equationally compact algebra.*

Note that an algebra A is *equationally compact* if every infinite system of algebraic equations S over A in variables x_i, $i \in I$, is simultaneous solvable in A provided that every finite subset of S has a simultaneous solution. By an *algebraic equation* $p = q$ over A we mean an equation where p, q are elements of the free product (coproduct) of A and the free algebra F with basis x_i, $i \in I$.

There is a Mal'cev-type characterization of residually small varieties of semigroups, groups and rings (see, for example, [A3, p. 60]).

DEFINITION 35. A variety V has *definable principal congruences* (in abbreviated *DPC*) if there exists a formulae $F(x, y, z, t)$ in first order language such that for any algebra A in V and arbitrary elements $a, b, c, d \in A$, a pair (a, b) is in the principal congruence generated by a pair (c, d) if and only if $F(a, b, c, d)$ is satisfied in A.

THEOREM 36 [HM]. *A variety V has DPC if one of the following conditions is satisfied:*

(i) *V is directly presentable, that is $V = \mathrm{HSP}(A)$ for some finite algebra A and the set of directly indecomposable finite V-algebras is finite*;
(ii) *V is locally finite and has* CEP.

Note that a variety V has congruence extension property (in abbreviated CEP) if for any subalgebra B in a V-algebra A the restriction morphism $\mathrm{Con}\, A \to \mathrm{Con}\, B$ is surjective. Varieties with CEP form a strong Mal'cev class.

For locally finite varieties congruence distributivity is almost equivalent to the property DPC [Pi2, p. 62].

A directly representable variety V is congruence-permutable. Conversely, if a congruence-permutable variety V is generated by a finite algebra and V_{SI} contains only simple algebras, then V is directly representable [BS].

DEFINITION 37. A collection of varieties $K_1, \ldots, K_n$ is *independent* if there exists a term $p(x_1, \ldots, x_n)$ such that $p(x_1, \ldots, x_n) = x_i$ is an identity in K_i, $i = 1, \ldots, n$.

THEOREM 38. *Let $K_1 \vee \cdots \vee K_n$ be congruence-modular, and let $K_1, \ldots, K_n$ be independent. Then every algebra A in $K_1 \vee \cdots \vee K_n$ has a unique decomposition $A \simeq A_1 \times \cdots \times A_n$, $A_i \in K_i$. Conversely, if $n = 2$, $K_1 \vee K_2 = K_1 \times K_2$ is congruence-permutable, and $K_1 \wedge K_2 = E$, then K_1, K_2 are independent* ([A3, p. 340], [S, Vol. II, Ch. VI]).

1.3. *Lattices of varieties*

Let V be a variety and $L(V)$ be a set of all subvarieties in V. Then $L(V)$ is a partially ordered set with respect to inclusion. Moreover $L(V)$ is a coalgebraic lattice which is dually isomorphic to the lattice of verbal congruences in the free V-algebra $F_V(X)$, where X is a countable set $X = \{x_1, x_2, \ldots\}$. The lattice $L(V)$ contains atoms, which are called *minimal* or *equationally complete varieties.*

THEOREM 39 [Mal1]. *A variety of semigroups is minimal if and only if it is defined by one of the following sets of identities*:
(i) $xy = yx$, $x^2 = x$;
(ii) $xy = x$;
(iii) $xy = y$;
(iv) $xy = zt$;
(v) $xy = yx$, $x^p y = y$, *p a prime.*

THEOREM 40 [Mal1]. *A variety of groups is minimal if and only if it is defined by identities $xy = yx$, $x^p = 1$ for some prime p.*

THEOREM 41 [Mal1]. *A variety of power-associative rings is minimal if and only if it is defined by one of the following sets of identities* (p *is a prime*):
(i) $xy = px = 0$;

(ii) $(xy)z = x(yz)$, $x^p = x$.

Several authors have studied varieties of algebras V for which the lattice $L(V)$ has some nice properties. For example, (see [A2]) there exists in terms of identities, a characterization of varieties V of (power-)associative algebras in which $L(V)$ has a unique atom. In a serie of papers (see, for example [A2]) the author has determined, in terms of identities, varieties V of groups and (non-)associative k-algebras over a commutative associative Noetherian ground ring k such that $L(V)$ is a chain.

THEOREM 42. *Let k be a field,* $\operatorname{char} k = 0$, *and let V be a variety of (non-)associative k-algebras. The lattice $L(V)$ is a chain if and only if V satisfies one of the following sets of identities*:

(i) $xy = yx$, $(xy)z = x(yz)$;
(ii) $xy + yx = (xy)z + x(yz) = 0$;
(iii) $x^2 = (xy)z + (yz)x + (zx)y = (xy)(zt) = 0$;
(iv) $xy = yx$, $x^3 = (xy)(zt) = 0$;
(v) $x^2 = (xy)z + (yz)x + (zx)y = ((xy)y)y = 0$;
(vi) $xy = yx$, $x^3 = 0$, $((xy)z)t = ((xy)t)z$.

THEOREM 43. *Let V be a non-Abelian variety of groups such that $L(V)$ is a chain,* $\exp V \neq 4$, *and let the free V-group of a rank 4 be residually solvable. Then* $\exp V = p \geqslant 3$ *is prime, and V satisfies one of the identities*:

(i) $[[x, y], [z, t]] = 1$;
(ii) $[[[x, y], y], y] = 1$.

THEOREM 44. *Let V be a variety of groups,* $\exp V = 4$ *and assume that $L(V)$ is a chain. Then V satisfies one of the following properties*:

(i) $V \subseteq A_2^2$, *that is each V-group is an extension of a group of exponent* 2 *by a group of exponent* 2;
(ii) *the identity* (i) *from Theorem* 43 *holds in V*;
(iii) *the identity* $[[[y, x], z], t][[[z, x], y], t][[[t, x], y], z] = 1$ *holds in V.*

Conversely, let V a variety of groups defined by one of the sets identities:

(i) $xy = yx$, $x^{p^n} = 1$, p a prime;
(ii) $x^p = 1$, $p \geqslant 3$ a prime, and identity (ii) from Theorem 43;
(iii) $x^p = 1$, $p \geqslant 3$ a prime, and identity (i) from Theorem 43;
(iv) $V \subseteq A_2^2$;
(v) $x^4 = 1$ and identity (iii) from Theorem 44.

Then $L(V)$ is a chain. These varieties are locally nilpotent.

A detailed survey of these results can be found in [A1,A2,A3], see also [Gr, §63]. Note that V. Martirosyan obtained a characterization of varieties V of right alternative algebras over a field of characteristic zero whose lattice $L(V)$ is distributive (see [A3]). These varieties V are defined by identities of degree three. A survey of results on varieties of groups V with a distributive lattice $L(V)$ is presented in [S, Vol. I, p. 141]. Quite recently

M. Volkov has found all varieties of semigroups V with a modular lattice $L(V)$ [I]. A survey on this topic is published in [S, Vol. II, pp. 161–165, 283–292].

In 1967 A.I. Mal'cev [Mal2] introduced an operation of *products of classes* of algebras. Given two subclasses U, V in a clases K denote by $U_K V$ the class of all algebras A in K which have a congruence c such that $A/c \in V$ and each c-class in A that is a K-algebra belongs to U. If K is a pre-variety, then a class of all sub-pre-varieties in K is closed under products.

THEOREM 45 (A.I. Mal'cev [Mal2]). *Assume that K is a congruence-permutable variety of algebras with a unary term f such that K satisfies the identities $f(x) = f(y), t(f(x), \ldots, f(x)) = f(x)$ for any term $t(x_1, \ldots, x_n)$. Then the set of all subvarieties in K is closed under products.*

THEOREM 46 [BO, Ch. 3, §2]. *If K is the variety of all groups or of all Lie algebras over a field of characteristic zero. Then the set of all subvarieties in K with the respect to a Mal'cev product form a free semigroup with unity (a one-element variety) and with zero (a variety K).*

THEOREM 47. *If K is the variety of all associative algebras (semigroups), then the Mal'cev product operation is not associative.*

1.4. *Finiteness conditions. Invariants of varieties. Constructions*

A subvariety V in a variety W is *finitely based* if it is defined in W by a finite set of identities. A single algebra $A \in W$ is *finitely based* if the variety generated by A is finitely based. Any finite group (Oates, Powell, 1964), finite Lie algebra (Bahturin, Ol'shansky, 1974), finite alternative algebra (L'vov, 1977) is finitely based [BO, pp. 127–128, 199–206, 217–221]. There exists a 64-element non-associative algebra with identity $x(yz) = 0$ which has no finite base (S.V. Polin). Each semigroup of an order less than 6 is finitely based. There exists a non-finitely based six-element semigroup (Perkins) [Gr, §63]. If a finite algebra A generates a congruence-distributive variety, then A is finitely based (K.A. Baker [Gr, §62]).

Let $f(n)$ be the number of all non-isomorphic algebras of order n, and $l(n)$ the number of all finitely based non-isomorphic algebras of an order n. Then (V. Murskii [Gr, §63])

$$\lim_{n\to\infty} \bigl(l(n)/f(n)\bigr) = 1.$$

A finite algebra generating a residually small variety is finitely based R. McKenzie [A3, p. 57]. A locally finite variety V is finitely based if V_{SI} consists of a finite set of finite algebras and V has definable principal congruences (R. McKenzie). In this situation the set of all finitely generated subvarieties in V is a sublattice in $L(V)$ (W. Block, D. Pigozzi, [A3, p. 57]). If a variety V is inheritably finitely based (that is every subvariety in V is finitely based), then the lattice $L(V)$ has DCC.

THEOREM 48 (A. Kemer [K]). *Let k be a field,* char $k = 0$. *Then every variety of associative k-algebras is finitely based.*

THEOREM 49 (M. Vaughan-Lee, V. Drensky [BO]). *Let k be a field,* char $k = p > 0$. *There exists a variety of Lie k-algebras which is not finitely based.*

THEOREM 50 (A.Yu. Ol'shansky, S.I. Adyan, M. Vaughan-Lee [BO]). *There exists a variety of (soluble)groups which is not finitely based.*

A variety V of (non-associative) algebras is a *Specht* variety if V is hereditary finitely based, that is any subvariety in V is finitely based.

THEOREM 51 (A. Iltyakov [Il]). *Let V be a variety of Lie algebras generated by a finite-dimensional Lie algebra. Then V a Specht variety.*

DEFINITION 52. Let V be a subvariety in a variety W. The axiomatic rank $r_a(V, W)$ is the least non-negative integer r such that V is defined in W be identities depending on r variables. If there is no such r we put $r_a(V, W) = \infty$.

DEFINITION 53. The basic rank $r_b(V)$ of a variety V is the least non-negative integer r such that V is generated by a free algebra of a rank r. If there is no such r then $r_b(V) = \infty$.

If a variety V is finitely based then $r_a(V, W)$ is finite.

The basic rank of the varieties of all groups, semigroups, associative algebras, Lie algebras is equal to 2, since the free algebra of rank two in these varieties contains a free algebra of countable rank. In contrast, the basic rank, of the varieties of alternative and Mal'cev non-associative algebras is infinite (I. Shestakov, V. Philippov [S, Vol. I, Ch. III, §3]). A detailed review of results on axiomatic and basic rank of varieties of groups, simigroups, algebras is presented in [S, v. II, Ch. IV, §7], [BO, §3].

DEFINITION 54. *The spectrum* Spec V of a variety V is the set of orders of finite algebras in V.

It is clear that Spec V is a multiplicative submonoid of positive integers N^*. For any submonoid M in N^* there exists a variety V such that $M = \text{Spec}\, V$. For a finite algebra A the following are equivalent [S, Vol. II, p. 348]):

(i) Spec(HSP(A)) is a cyclic monoid generated by A;

(ii) A is a simple algebra, any proper subalgebra in A has order one, and HSP(A) is a congruence permutable variety.

For example, let $k = \mathbb{F}_q$ be a finite field and let V the variety of commutative, associative k-algebras generated by $A = \mathbb{F}_q$. Then Spec V is the cyclic monoid generated by q.

Another important invariant of a variety V is its *fine spectrum*. This is the function f_V defined on the "set of cardinals" such that $f_V(a)$ is the cardinality of a set of non-isomorphic algebras in V of a cardinality a. For example, let A be a finite simple algebra of order m as in (ii) above. Then the restriction of f_V to N^* coincides with one of the following functions [Pi1, p. 69],

$$f(n) = \begin{cases} 1 & \text{if } n \text{ is a power of } m, \\ 0 & \text{otherwise;} \end{cases}$$

$$f(n) = \begin{cases} 1 & \text{if } n = 1, \\ 2 & \text{if } n \text{ is a power of } m, \\ 0 & \text{otherwise.} \end{cases}$$

DEFINITION 55. A variety V has *categoricity at the cardinal* a, if $f_V(a) = 1$.

Varieties with categoricity at any infinite cardinal a were characterized by E. Palyutin and S. Givant [Gr, p. 393].

DEFINITION 56. *A free spectrum* of a locally finite variety V is the function $s(n)$, where $s(n)$ is the order of the free V-algebra of a rank n.

THEOREM 57 [HM, p. 164]. *Let A be a finite algebra and $V = \mathrm{HSP}(A)$. Then either $s(n) \leqslant Cn^k$ for some fixed integers c, k, or $s(n) \geqslant 2^{n-k}$ for some k.*

Note that by Theorem 10 we have in Theorem 57

$$s(n) \leqslant |A|^{(|A|^n)}.$$

If V is the variety of groups generated by the symmetric group S_3, then $s(n) = 2^n 3^{2^n(n-1)+1}$ [HM]. A locally finite variety of groups is nilpotent if and only if $ln(s(n))$ is a polynomial in n. There exists distint locally finite varieties of groups with the same free spectrum (A.Yu. Ol'shansky, Yu. Kleiman, see Section 3.1).

Let V be a variety of algebras and D be a diagram in the category V.

THEOREM 58 [Gr, §21]. *There exist in V a direct and inverse limits of any diagram D.*

In particular, if all morphisms in D are identities, then the inverse limit of D is the direct product of objects in D and the direct limit is a free product (coproduct) of objects in D. A theory of direct decompositions in congruence-modular varietes was developed by A.G. Kurosh and O. Ore [Ku1].

Let A be a finite simple algebra generating a congruence-permutable variety. Then any algebra in HSP(A) is a subdirect product of primal algebras (D. Clark, P.H. Krauss [KC], [A3, p. 58]). If V is a minimal locally finite permutable variety, then $V = \mathrm{SP}(A) \cup E$, where A is a finite simple algebra (D. Clark [A3, p. 58]), and conversely.

The Nielsen–Schreier theorem on subgroups of free groups and Kurosh's theorem of subgroups of free products of groups have inspired a series of papers on the structure of subalgebras of free algebras and free products of algebras in varieties. Detailed surveys on this topic can be found in [Bar,BB].

2. Varieties of (non-)associative algebras

The theory of varieties of (non-)associative algebras is one of the most developed. It has some special methods and very interesting results. Therefore, we devote a separate section to the theory of these varieties.

2.1. *Classification of identities*

Let k be a ground field of a characteristic $p \geqslant 0$ and V be a variety of (non-)associative k-algebras. Take the free V-algebra $F = F_V(X)$ over X and an identity $f = 0$ in some subvariety W in V, $f \in F$. Then f as an element of F is a finite sum of monomials

$$g = ax_{i_1} \dots x_{i_m} \tag{***}$$

with some bracketing, where $a \in k$, $x_{i_j} \in X$. Let M be the additive monoid of all functions $h : X \to \mathbb{N} \cup 0$ such that $h(x) = 0$ for almost all x in X. For a function $h \in M$ and a given element $f \in F$ denote by f_h the sum of all monomials g in (***) such that for each element $x_d \in X$

$$h(x_d) = \delta_{d,i_1} + \dots + \delta_{d,i_m},$$

where δ_{**} is the Kroneker δ-function. The number $h(x_d)$ is exactly the number of *occurences of a variable x_d* in a monomial g. Thus

$$f = \sum_{h \in M} f_h.$$

If $f = f_h$ for some $h \in M$, then f is called a *multihomogeneous identity*.

THEOREM 59. *Let k be an infinite field. Then any identity is equivalent to a system of identities $f_h = 0$, $h \in M$.*

COROLLARY 60. *Let V be a variety over an infinite field k. If F is a free V-algebra over a set X, then F is a graded algebra*

$$F = \bigoplus_{h \in M} F_h,$$

where each F_h is the linear span of elements $f = f_h$. Moreover, $F_h F_t \subseteq F_{h+t}$, for all $h, t \in M$.

Consider now an identity $f = 0$, where f is an element of the free V-algebra $F_V(X)$ with the standard countable base $X = x_1, x_2, \dots$.

DEFINITION 61. An identity $f = 0$ is *multilinear* of degree n if $f = f_h$, where $h(x_1) = \cdots = h(x_n) = 1$, $h(x_j) = 0$ for $j > n$.

For instance, the Jacobi identity $J(x, y, z) = x(yz) + y(zx) + z(xy) = 0$ and the associative law $x(yz) - (xy)z = 0$ are examples of multilinear identities of a degree three.

THEOREM 62 (A.I. Mal'cev). *Let* char $k = 0$. *Then any identity is equivalent to a system of multilinear identities. If* char $k \neq 0$, *then every identity implies a system of multilinear identities.*

The proof of this theorem is based on a linearization process. Let $f = f_h$ be a multihomogeneous element in the free V-algebra $F_V(X)$. Assume for example, that $h = h(x_1) \geqslant 1, \ldots, h(x_j) = 0$ for $j > n$. Then $f = f(x_1, \ldots, x_n)$, and for every monomial g in

$$f(x_1 + x_{n+1}, x_2, \ldots, x_n) - f(x_1, x_2, \ldots, x_n) - f(x_{n+1}, x_2, \ldots, x_n)$$

the number of occurences of x_1, x_{n+1} are less than h, and the number of occurences of the other variables did not change. Repeating this process we can finally obtain multilinear consequences of f. If char $k = 0$ then an identity $f = 0$ is equivalent to its multilinear consequences.

Denote by F_n a subspace in $F_V(X)$ spanned by all monomials (***) such that $h(x_1) = \cdots = h(x_n) = 1$, $h(x_j) = 0$, $j > n$. Then the symmetric group S_n has a natural representation in F_n. This group acts by permutation of the variables $x_1, \ldots, x_n$. If W is a subvariety in V denote by $W(F)$ a set of all elements f in F such that $f = 0$ is an identity in W. Thus $W(F)$ in this sense is the ideal associated to the verbal congruence in F associated to a subvariety W in the sense of Section 1.1 (see Theorem 9). Therefore, $W(F)$ is an ideal in $F = F_V(X)$ stable under all endomorphisms of F. The ideal $W(F)$ is called a *verbal ideal* (or T-*ideal*) *associated with* W. According to Theorem 62 in characteristic zero the ideal $W(F)$ is generated by a family of subspaces $W(F) \wedge F_n$, $n \geqslant 2$. Each subspace $W(F) \wedge F_n$ is invariant under permutation of variables, since $W(F)$ is stable under all endomorphisms of F. Hence, each space $W(F) \wedge F_n$, $n \geqslant 2$, is a kS_n-submodule in F_n. The lattice of subvarieties $L(V)$ is embeddable in the direct product of lattices of kS_n-submodules in F_n, $n \geqslant 2$.

Each kS_n-module F_n has finite dimension. Thus instead of a free spectra $s(n)$ in a theory of (non-)associative algebras it is necessary to study the behaviour of the function $\dim F_n$ and properties of the generating series

$$f(t) = \sum_{n \geqslant 0} (\dim F_n) t^n.$$

2.2. *Varieties of associative algebras*

The theory of associative algebra with a non-trivial identity (PI-algebras) is one of the most advanced. A systematic exposition of this theory can be found in [J,Pr,Ro], [S, Ch. III].

One of the most important identities in associative algebras is the *standard identity*

$$S_n(x_1, \dots, x_n) = \sum_{s \in S_n} (-1)^s x_{s1} \dots x_{sn}.$$

If A is an algebra of a finite dimension n then A satisfies $S_{n+1} = 0$.

THEOREM 63 (Amitsur–Levitzki [Pr, p. 22]). *Let R be a commutative algebra over a field k. Then a matrix algebra* $\mathrm{Mat}(n, R)$ *satisfies $S_{2n} = 0$. This is an identity of a minimal degree in* $\mathrm{Mat}(n, R)$.

THEOREM 64 (Amitsur [Pr, p. 43]). *Let A be a* PI-*algebra. Then there exists positive integers n, m such that A satisfies an identity $S_n(x_1, \dots, x_n)^m = 0$. In particular, if A is semiprime, then A satisfies an identity $S_n(x_1, \dots, x_n) = 0$ for some n.*

An algebra in which every element is algebraic of a bounded degree is a PI-algebra [J, pp. 236–237].

THEOREM 65 [J, pp. 242–243]. *Let A be an algebraic* PI-*algebra. Then A has a finite dimension.*

In contrast we have

THEOREM 66 (E. Golod). *Let k be a field and $d \geqslant 2$ an integer. There exists a non-nilpotent graded algebra $A = \bigoplus_{n \geqslant 0} A_n$ in which every $(d-1)$-generated subalgebra is nilpotent.*

Thus an assumption of a polynomial identity in Theorem 65 is essential.

THEOREM 67 (Nagata–Higman [J, p. 274]). *Let* $\mathrm{char}\, k = p \geqslant 0$ *and let A be a k-algebra with an identity $x^n = 0$, where $p > n$ provided $p > 0$. Then A is nilpotent of class at most $2^n - 1$.*

THEOREM 68 (Yu.P. Razmyslov). *Let $p = 0$ in the Theorem* 66. *Then A is nilpotent of a class at most n^2.*

Note that lower bound for the class of nilpotency of an algebra A in Theorems 67, 68 $(p = 0)$ is $\frac{1}{2}n(n+1)$ [JSSS].

THEOREM 69 (I. Kaplansky). *Let R be a primitive* PI-*algebra. Then R is a simple algebra of a finite dimension over its center C. If R satisfies an identity of degree d, then*

$$\dim_C R \leqslant [d/2]^2.$$

DEFINITION 70. An element $f(x_1, \dots, x_m) \neq 0$ of the free associative algebra is a *central polynomial* in the matrix algebra $\mathrm{Mat}(n,k)$ if $f(x_1, \dots, x_m)x_{m+1} - x_{m+1}f(x_1, \dots, x_m) = 0$ is an identity in $\mathrm{Mat}(n,k)$ and $f(a_1, \dots, a_m) \neq 0$ for some elements $a_1, \dots, a_m$ in $\mathrm{Mat}(n,k)$.

THEOREM 71 (Yu. Razmyslov, E. Formanek [BO, pp. 159–160]). *For every integer $n \geqslant 2$ there exists a central multilinear polynomial of a degree $4n^2 - 2$ in* $\mathrm{Mat}(n,k)$.

In order to present the explicit form of this polynomial consider the *Capelli element*

$$\begin{aligned} &C_{n^2}(x_1, \dots, x_{n^2}, y_1, \dots, y_{n^2-1}) \\ &\quad = \sum_{s \in S_{n^2}} (-1)^s x_{s1} y_1 x_{s2} y_2 \dots x_{s(n^2-1)} y_{n^2-1} x_{s(n^2)}. \end{aligned}$$

Consider the linear operator A on the set of multilinear polynomials defined by $A(ax_{n^2}b) = bx_{n^2}a$, where a, b are polynomials. Now, in

$$A\big(C_{n^2}(x_1, \dots, x_{n^2}, y_1, \dots, y_{n^2})\big)$$

substitute for x_i a commutator $[x_i, x_i']$, and for y_i a commutator $[y_i, y_i']$. The polynomial obtained in this manner is a central polynomial in $\mathrm{Mat}(n,k)$ of a degree $4n^2 - 2$.

THEOREM 72 (Yu. Razmyslov, A. Kemer, Braun [BO]). *Let A be a* PI*-algebra which is finitely generated over its central Noetherian subalgebra. Then the nil-radical $N(A)$ of A is nilpotent.*

We have already mentioned the fundamental result of Kemer that in characteristic zero every variety of associative algebras is finitely based. The next theorem is also due to A. Kemer [K].

THEOREM 73. *Let F be a finitely generated relatively free k-algebra, where k is an infinite field. Then F is embeddable in a matrix algebra* $\mathrm{Mat}(n, C)$ *for some n and for some commutative k-algebra C.*

Let k be an infinite field and A a finitely generated associative k-algebra. Then there exists a finite-dimensional k-algebra B such that $\mathrm{HSP}(A) = \mathrm{HSP}(B)$. In particular, A is finitely based (A. Kemer [K]). Some new interesting on the subject were recently obtained by young Moscow mathematicients A.Ya. Belov, V.V. Shchigolev and A.V. Grishin.

THEOREM 74 (A.Ya. Belov [B1]). *Let A be a finitely generated* PI*-algebra over any commutative ring. Then A is embeddable in an algebra B which is a finitely generated module over its center. In particular, A is finitely based.*

THEOREM 75 (A.Ya. Belov [B2], V.V. Shchigolev [Shch2], A.V. Grishin [Gri]). *Let k be a field of a positive characteristic. Then there exists a non-finitely based variety of associative k-algebras.*

For the case of characteristic 2 Theorem 75 was proved in [Gri].

A subspace L in a free associative algebra $F(X)$ with a base X is a T-*space* if L is closed under all endomorphisms of $F(X)$. Observe that if T-space is an ideal then it is a T-ideal. T-space L is *finitely generated* if there exist finitely many elements $g_1, \ldots, g_m \in L$ such that L is the least T-space containing $g_1, \ldots, g_m$.

THEOREM 76 (V.V. Shchigolev [Shch1]). *Let k be a field of characteristic zero. Then any T-space in a free associative k-algebra is finitely generated.*

THEOREM 77 (V.V. Shchigolev [Shch2], A.V. Grishin [Gri]). *Let k be a field of a positive characteristic and $F(X)$ a free associative k-algebra with a base X, $|X| \geqslant 2$. Then there exists a non-finitely generated T-space in $F(X)$.*

The case of characteristic 2 was considered in [Gri], an arbitrary case – in [Shch2].

Another aspect of the theory of PI-algebras is consedered in papers by A. Giambruno and M. Zaicev. These investigations are inspired by

THEOREM 78 (A. Regev [R]). *Let L be a T-ideal in a free associative algebra $F(X)$ over a field k where $X = \{x_1, x_2, \ldots\}$ is a countable base. Let $F_n, n \geqslant 1$, be as in Theorem* 59 *and $L \cap F_d \neq 0$ for some $d \geqslant 1$. Then for all $n \geqslant 1$*

$$\dim\bigl(F_n/(F_n \cap L)\bigr) \leqslant (d-1)^{2n}.$$

THEOREM 79 (A. Giambruno, M. Zaicev [GZ]). *Let k be a field of characteristic zero and L, F_n from the previous theorem. Then*

$$\lim_{n\to\infty} \sqrt[n]{\dim_k(F_n/(L \cap F_n))}$$

exists and it is an integer.

Group algebras and a universal enveloping algebras are important examples of associative algebras which play a role in the representation theory of groups and Lie algebras.

THEOREM 80 (V. Latyshev). *Let k be a field of characteristic zero. The universal enveloping algebra $U(L)$ of a Lie k-algebra L is a* PI-*algebra if and only if L is Abelian.*

THEOREM 81 (Yu. Bahturin [BO]). *Let k be a field of a characteristic $p > 0$. The universal enveloping algebra $U(L)$ is a* PI-*algebra if and only if L has an Abelian ideal I of a finite codimension, such that the adjoint representation of L is algebraic.*

THEOREM 82 (D. Passman). *Let k be a field of a characteristic $p \geqslant 0$. Then a group algebra kG is a* PI-*algebra if and only if G has a subgroup A of a finite index such that*

(i) *A is Abelian if $p = 0$;*

(ii) *$[A, A]$ is a finite p-group if $p > 0$.*

A similar result for restricted enveloping algebras of Lie p-algebras was obtained by Passman and Petrogradsky [R, p. 72].

An important role in PI-algebras is played by a combinatorial theorem of Shirshov on heights. Let A be an algebra and X be a subset in A. If $a \in A$ and

$$a = bx_{i_1}^{m_1} \ldots x_{i_n}^{m_n}, \quad b \in k, \ x_{i_j} \in X, \ m_i \geqslant 1,$$

then the *height* $h(a)$ is equal to at most n with respect to a set X.

THEOREM 83 (A. Shirshov [JSSS, pp. 128–129]). *Let A be an algebra generated by elements $a_1, \ldots, a_d$ and satisfying a multilinear identity of degree n. Consider the subset X in A consisting of all monomials*

$$a_{i_1} \ldots a_{i_r}, \quad r < n, \ i_j \in \{1, \ldots, d\}.$$

Then there exists a positive integer h such that every element in A is a sum of monomials a in X as above with $h(a) < h$.

This was a very brief review of the main result in the theory of PI-algebras. For further details consult [Pr,Ro,BO,BLH].

2.3. *Varieties of non-associative algebras*

A theory of varieties of Lie algebras was influenced by the Burnside problem on local finiteness of groups with an identity $x^n = 1$ (see [Ko]). The latest fundamental results on this topic were obtained by E. Zel'manov (Siberian Math. J. **30** (6) (1989)).

THEOREM 84. *Let k be a field. A Lie k-algebra with an identity* $(\operatorname{ad} x)^n = 0$ *is nilpotent if* $\operatorname{char} k = 0$ *and is locally nilpotent if* $\operatorname{char} k > 0$.

We have already mentioned that a finite-dimensional Lie algebra is finitely based.

DEFINITION 85. A variety V of Lie algebras over a finite field is called a *just-non-Cross variety* if

(i) V is not generated by a finite algebra;
(ii) any proper subvariety in V is generated by a finite algebra.

THEOREM 86 (Yu. Bahturin, A. Ol'shansky [BO], [Ba, Ch. 7]). *There exists a unique soluble just-non-Cross variety of Lie algebras over a finite field. It is defined by the identities*

$$\big[[x, y], [z, t]\big] = (\operatorname{ad} x)^p (\operatorname{ad} y) z = (\operatorname{ad} x)^{p-1} (\operatorname{ad} y)^p x = 0,$$

where $p = \operatorname{char} k$.

This variety first appeared in [A2] as one of the varieties with a chain of subvarieties. Note that over an infinite field the lattice of subvarieties in the variety of all metabelian (i.e. satisfying an identity $[[x, y], [z, t]] = 0$) Lie algebras is distributive [Ba, Ch. 4.7]. There is a series of publications on special varieties of Lie algebras (see [Ba, Ch. 6]). *Speciality* of a variety V means that V is generated by a Lie algebra L such that there exists an associative PI-algebra A and an embedding of L into a Lie algebra $A^{(-)}$ which is defined on A by a multiplication $[x, y] = xy - yx$.

In connection with Theorem 86 it is necessary to mention

THEOREM 87 (Yu. Razmyslov). *There exists a non-soluble just-non-Cross variety of Lie algebras $E_{p-2,p}$ over a field of a characteristic $p \geqslant 5$. This variety is defined by the one identity* $(\operatorname{ad} x)^{p-2} y = 0$.

Note that by [A1] the lattice of subvarieties in $E_{3,5}$ is a chain and any proper subvariety in $E_{3,5}$ is nilpotent.

THEOREM 88 (A. Krasil'nikov [BO]). *Let k be a field of a characteristic zero. Then any subvariety in $N_c A$ is finitely based.*

Here $N_c A$ is the product of the nilpotent variety N_c of all nilpotent Lie algebras of a degree $\leqslant c$ and an Abelian variety A. Note that in the class of all Lie algebras the Mal'cev product UV of two subvarieties U, V consists of all Lie algebras L such that there exists an ideal I in L for which $L/I \in V$ and $I \in U$. For a more detailed review on this topic consult [Ba,BO].

Free alternative algebras F have been mainly studied by I. Shestakov, K. Jevlakov, A. Philippov, A. Il'tyakov, S. Pchelincev and others. For elements x, y, z in F put $(x, y, z) = (xy)z - x(yz)$.

THEOREM 89 [S, Vol. I, pp. 403–404]. *If $x, y, z, t, u \in F$, then*

(i) $([x, y]^4, z, t) = (z, [x, y]^4, t) = (z, t, [x, y]^4) = 0$ (*Kleinfeld*);

(ii) $((x, y, z)^4, t, u) = (t, (x, y, z)^4, u) = (t, u, (x, y, z)^4) = (x, y, z)^4, t = t, (x, y, z)^4 = 0$.

The quasiregular radical of F is locally nilpotent. It consists of all nil elements in F and coincides with the ideal of all identities in the Cayley–Dickson ring $O(Z)$ that belong to the ideal in F generated by all elements $(x, y, z), x, y, z \in F$. If the rank of a free alternative ring G is sufficiently large then G has elements whose additive order is equal to 3.

THEOREM 90 [JSSS]. *Let A be an alternative algebraic* PI-*algebra. Then A is locally finite dimensional. In particular, alternative algebras with an identity $x^n = 0$ are locally nilpotent, and are soluble.*

THEOREM 91 (I. Shestakov). *The quasiregular radical in a finitely generated alternative* PI-*algebra over a field is nilpotent* (*see Theorem* 71).

THEOREM 92 (I. Shestakov). *Let* Alt(n) *be the variety of alternative algebras generated by the free alternative algebra of rank* n. *Then* $\mathrm{Alt}(n) \neq \mathrm{Alt}(2^n + 1)$. *In particular, the variety of all alternative algebras has infinite basic rank.*

THEOREM 93 (S.V. Pcelincev [P]). *A soluble alternative algebra over a field of characteristic* $\neq 2, 3$ *belongs to a variety* $N_k A \cap N_3 N_m$, *where as above* A *is the variety of all algebras with zero multiplication,* N_m *is a variety of all nilpotent algebras of a class* $\leqslant m$.

THEOREM 94 (U.U. Umirbaev [Um]). *Every variety* $N_k A \cap N_3 N_m$ *of alternative algebras is a Specht variety.*

Let $J(X)$ be the free Jordan algebra over X. We shall assume that the characteristic of the base field k is not equal to 2.

THEOREM 95 (A. Shirshov). *If* $|X| \leqslant 2$, *then* $J(X)$ *is special, that is* $J(X)$ *is isomorphic to a Jordan subalgebra in a free associative algebra* Ass(X) *generated by* X *with the respect to the multiplication* $x \circ y = \frac{1}{2}(xy + yx)$.

If $|X| \geqslant 3$, then $J(X)$ is not special and contains zero divizors (Yu. Medvedev). It follows from Theorem 95 that any two-generated Jordan algebra is special (i.e. embeddable in an algebra $A^{(+)}$ defined in associative algebra A by the multiplication $x \circ y = \frac{1}{2}(xy + yx)$) (P.M. Cohn).

Denote by $SJ(X)$ the Jordan subalgebra in Ass(X) generated by X with the respect to the multiplication $x \circ y$ as above. There exists a surjective homomorphism $p : J(X) \to SJ(X)$. The elements of $\operatorname{Ker} p$ are called s-identities. The Glenny identity $G(x, y, z) = K(x, y, z) - K(y, x, z)$, where

$$K(x, y, z) = 2\{\{y\{xzx\}y\}z(xy)\} - \{y\{x\{z(xy)z\}x\}y\},$$

is an example of an s-identity. Here $\{abc\} = (ab)c + (cb)a - b(ac)$.

DEFINITION 96. A Jordan algebra is a PI-*algebra* if it satisfies an identity which is not an s-identity.

THEOREM 97 (E. Zel'manov). *Jordan nil* PI-*algebras are locally nilpotent.*

THEOREM 98 (A. Veis, E. Zel'manov [I, pp. 42–51]). *Let* A *be a finitely generated Jordan* PI-*algebra. Then any subvariety in* HSP(A) *is finitely based.*

A variety of Jordan algebras is *special* if all its members are special. Let N_m be the variety of Jordan nilpotent algebras of a class at most m and M_t the variety of Jordan algebras generated by the Jordan algebra of symmetric matrices in Mat(t, k).

THEOREM 99 (A. Slin'ko). *The variety* N_m *is special if and only if* $m \leqslant 5$. *The variety* M_t *is special if and only if* $t \leqslant 2$.

THEOREM 100 (K. Patil, M. Racine [BSS]). *The Formanek central polynomial* (*see Theorem* 70) *as an element in* $\mathrm{Ass}(X)$ *belongs to* $SJ(X)$.

DEFINITION 101. *A Mal'cev algebra* is a skew-commutative algebra with satisfying the identity $J(x, y, xz) = J(x, y, z)x$.

In every Mal'cev algebra any two-generated subalgebra is a Lie algebra.

THEOREM 102 (V. Philippov). *Let* A *be a Mal'cev algebra satisfying the identity* $(\operatorname{ad} x)^n y = 0$. *Then* A *is locally nilpotent.*

Assume now that $\operatorname{char} k \neq 2, 3$ and let $M(n)$ be the free Mal'cev algebra of rank n. V. Philippov introduced the element

$$g(x, y, z, t, u) = J\big([x, y, xz], t, u\big) + J\big([x, t, yz], x, u\big) \in M(5),$$

where $[x, y, z] = 3(xy)z - J(x, y, z)$. For $n \geqslant 5$ put

$$p_n = \big(\dots\big(g(x_1, \dots, x_5)x_6\big)\dots\big)x_n \in M(n).$$

THEOREM 103 (V. Philippov). *The algebra* $M(4)$ *satisfies the identity* $g(x_1, \dots, x_5) = 0$. *The algebra* $M(n)$, $n \geqslant 5$, *satisfies the identity* $p_{n+2} = 0$ *and does not satisfy the identity* $p_{n+1} = 0$. *In particular, the basic rank of the variety of all Mal'cev algebras is infinite.*

Note that $(g(x_1, \dots, x_5)x_6)x_6$ is a non-trivial central element in $M(n)$, $n \geqslant 6$.

Detailed expositions of the theory of varieties of non-associative algebras can be found in [JSSS,KS,BSS].

3. Varieties of groups, semigroups, lattices and other classes of algebras

3.1. *Varieties of groups*

The developement of the theory of varieties of groups was greatly influenced by the Burnside problem on periodic groups.

THEOREM 104 (S. Adyan [A]). *Let* $p \geqslant 665$ *be an odd prime. Then the free group of rank two in the variety* B_p *defined by the identity* $x^p = 1$, *is infinite. A free group of arbitrary finite rank in* B_p *is embeddable in the free group of a rank two.*

THEOREM 105 (S.V. Ivanov [Iv1]). *Let* $d \geqslant 2^4 8$ *and* d *is divisible by* 2^9 *if* d *is even. Then the free group of a rank two in the variety* B_d *is infinite.*

THEOREM 106 (S.V. Ivanov, A.Yu. Olshansky [IO]). *Let* d *be from the previous theorem and* $B(m, d)$ *a free* B_d*-group of a finite rank* m. *Then the word and conjugacy problems* *in $B(m,d)$ are solvable. Finite subgroups in $B(m,d)$ are isomorphic to direct products of dihedral groups*

$$D(2n_1) \times D(2n_2)^l.$$

The group $B(m,d)$ contains maximal locally finite subgroups that are not FC-groups. A subgroup G in $B(m,d)$ is locally finite if every 2-generator subgroup in G is finite.

THEOREM 107 (L.G. Lysenok [L]). *Let $n \geqslant 13$. Then the free group of a rank two in the variety B_{2^n} is infinite.*

A survey of recent results of Burnside-type theorems for groups is exposed in [Iv2].

THEOREM 108 (A. Kostrikin, E. Zel'manov). *Let $n \geqslant 5$ be a a positive integer. Then all locally finite groups with the identity $x^n = 1$ form a variety (provided there exists finitely many sporadic finite simple groups).*

Note that the groups with identity $x^2 = 1$ are Abelian and therefore, locally finite. Let $B(r)$ be the free group of a rank r in the variety defined by the identity $x^3 = 1$. Then the order of $B(r)$ is equal to $3^{m(r)}$,where

$$m(r) = \binom{r}{3} + \binom{r}{2} + r$$

(F. Levi, B. van der Waerden).

THEOREM 109 (I. Sanov). *A group with identity $x^4 = 1$ is locally finite.*

THEOREM 110 (Yu. Razmyslov [Ra]). *The variety of groups defined by an identity $x^4 = 1$ is not soluble.*

It was mentioned in Section 1 that every finite group is finitely based. There exists continuum many non-finitely based varieties of groups. One of the most simple examples of non-finitely based varieties was given by Yu. Kleiman. Let B_4 be the variety defined by an identity $x^4 = 1$ and A the variety of all Abelian groups. Then the product B_4A is a non-finitely based group variety. Note that the Mal'cev product UV of two varieties of groups U, V consists of all groups G which have a normal subgroup N such that $N \in U$ and $G/N \in V$. According to Theorem 45 the product UV is again a variety of groups.

DEFINITION 111. A variety of groups V is a just-non-Cross variety if V is not generated by a finite groups, while any proper subvariety in V is generated by some finite groups.

THEOREM 112 (A. Ol'shansky [BO]). *Here is a list of all soluble just-non-Cross varieties of groups (compare this with Theorem 85)*

(i) *the variety of all Abelian groups A;*

(ii) A_p^2, *where* A_p *is the variety of Abelian groups with identity* $x^p = 1$;
(iii) $A_p(B_q \wedge N_2)$, *where* $q \neq p$ *are primes,* q *is odd,* N_2 *is defined by the identity* $[x, y, z] = 1$ *and* B_q *by the identity* $x^q = 1$;
(iv) $A_p(B_4 \wedge N_2)$.

THEOREM 113 (Yu. Razmyslov [Ra]). *The variety of groups defined by the identities* $x^p = (\mathrm{ad}\, x)^{p-2} y = 1$, $p \geqslant 5$ *a prime, is a non-solvable just-non-Cross variety of groups.*

The Nielsen–Schreier theorem states that a subgroup of a free group is free.

THEOREM 114 (P. Neumann, J. Wiegold [N]). *The only varieties of groups satisfying this freeness property are the following*:
(i) *the variety of all groups*;
(ii) *the variety of all Abelian groups*;
(iii) *the trivial variety consisting of a one-element groups*;
(iv) *the variety of Abelian groups with an identity* $x^p = 1$, *where* p *is a prime.*

Let G be a group defined in the variety V of all soluble groups of a class at most n by k generators $x_1, \dots, x_k$ and by $1 < k$ relations $r_1 = \dots = r_1 = 1$. Then there exists elements $x_{i_1}, \dots, x_{i_{k-1}}$, $i_1 < i_2 < \dots < i_{k-1}$ which freely generate a free V-group (N. Romanovsky, Algebra and Logic **16** (1) (1977), 88–97).

The lattice of subvarieties $L(A^2)$ in A^2 has DCC. A lattice $L(V)$ is distributive in the following cases [S, Ch. II]:
(i) $V = A$;
(ii) $V = N_3$;
(iii) $V = B_p \wedge N_{p+1}$, p is a prime;
(iv) $V = A^2 \wedge B_p \wedge N_p$, p is a prime;
(v) $V = (A \wedge B_m)(A \wedge B_n)$, where (m, n) is square-free.

Here N_m is a variety of nilpotent groups of a class at most m.

Let K be a countable subclass of non-isomorphic groups in A^7. Then there exist countably many subvarieties V_a in A^7 such that $V_a \wedge K = V_b \wedge K$ for all a, b (Yu. Kleiman, Siberian Math. J. **23**(6) (1982), 117–132). There exists a non-Abelian variety of groups in which all finite groups are Abelian (A. Ol'shansky, Mat. Sb. **126** (1) (1985), 59–82).

For further details see [N], [S, Ch. II].

3.2. *Varieties of semigroups and lattices*

A survey of results on non-finitely based varieties of semigroups was published in [ShV]. We shall only recall that there exists a unique non-group variety V of inverse semigroups with the following properties:
(i) V is finitely based;
(ii) any proper subvariety in V is not finitely based.

This variety V is generated by the 6-element Brandt semigroup B_2^1 with external unity.

If in a variety of inverse semigroups W any proper subvariety is finitely based, then either $W = M \vee \mathrm{HSP}(B_2)$, or $W = M \vee C$, where M is the group variety with the same property, B_2 is the 5-element Brandt semigroup and C is the variety of semilattices. The variety generated by B_2 can be defined by the identities

$$x^3 = x^2, \qquad (xy)^2x = xyx, \qquad x^2y^2 = y^2x^2.$$

A variety of semigroups V is residually small if and only if V satisfies one of the following sets of identities and properties:

(i) $x^{n+1}y = xy = xy^{n+1}$, $xy^nz^nt = xz^ny^nt$, and all groups in V form a residually small variety of groups;

(ii) $x^{n+1}y = xy$, $xyz = yxz$;

(iii) $xy^{n+1} = xy$, $xyz = xzy$.

THEOREM 115. *For a variety of semigroups V the following are equivalent:*

(i) *every semigroup in V is residually finite;*

(ii) *V is residually small and locally finite.*

In this case the variety V is generated by a finite semigroups.

The class of semigroup varieties is not closed under Mal'cev product. The lattice of semigroups varieties has no coatoms and the unity in this lattice is not a union of a family of non-unity elements. This lattice satisfies a quasi-identity

$$(x \wedge y = 0)\&(x \wedge z = 0) \to \big((x \vee y) \wedge z = 0\big).$$

Further results on varieties of semigroups can by found in [S, Vol. II, Ch. IV], [ShV,I] and, for example, in papers of E. Pastijn (Notes Canad. Math. Soc. **20** (8) (1988)), M. Petrich (Semigroup Forum **25** (1–2) (1982), 153–169), L. Polak (Simigroup Forum **32** (1) (1985), 97–123; **36** (3) (1987), 253–284; **37** (1) (1988), 1–30).

The variety of distributive lattices D is the unique minimal variety of lattices. It is generated by the two-element lattice (0, 1) (theorem of Stone), and therefore is locally finite. The free distributive lattice of rank 3 has 18 elements. The free distributive lattice of rank 7 has 2 414 682 040 996 elements. The free modular lattice of rank 3 has 28 elements and that of rank 4 is infinite.

THEOREM 116. *There are only two varieties of lattices covering the minimal variety D. One of the is generated by the diamond M_3 and the other by the pentagon N_5.*

THEOREM 117. *Let M be the variety of all modular lattices. Then the lattice $L(M)$ of subvarieties in M is infinite.*

The variety M has one covering $M \vee \mathrm{HSP}(N_5)$. The variety $\mathrm{HSP}(N_5)$ has 16 coverings and it is finitely based. The variety $\mathrm{HSP}(M_3)$ has 3 coverings.

THEOREM 118 (I. Berman, B. Wolk). *The free lattice in* $\mathrm{HSP}(M_3)$ *of rank* 3 *has order* 28 *and that of rank* 4 *has order* 19 982. *The free lattice in* $\mathrm{HSP}(N_5)$ *of a rank* 3 *has order* 99 *and that of rank* 4 *has order* 540 792 672.

Detailed surveys of recent results on this topic can be found in [U].

References

[A] S.I. Adyan, *The Burnside Problem and Identities in Groups*, Nauka, Moscow (1975).

[A1] V.A. Artamonov, *Universal algebra*, Algebra. Topology. Geometry, Vol. 4, VINITI, Moscow (1976), 191–248.

[A2] V.A. Artamonov, *Lattices of varieties of linear algebras*, Uspekhi Mat. Nauk **33** (2) (1978), 135–167.

[A3] V.A. Artamonov, *Universal algebra*, Algebra. Topology. Geometry, Vol. 27, VINITI, Moscow (1989), 45–124.

[Ba] Yu.A. Bahturin, *Identities in Lie Algebras*, Nauka, Moscow (1985).

[Bar] T.M. Baranovich, *Universal algebra*, Algebra. Topology. Geometry, VINITI, Moscow (1966), 109–136.

[BB] T.M. Baranovich and M.S. Burgin, *Linear Ω-algebras*, Uspekhi Mat. Nauk **30** (4) (1975), 61–106.

[BO] Yu.A. Bahturin and A.Yu. Ol'shansky, *Identities*, Algebra 2, Fund. Napr., Vol. 18, VINITI, Moscow (1988), 117–240.

[B1] A.Ya. Belov, *Local representability of relatively free associative algebras*, Kurosh Algebraic Conference'98. Abstract of Talks, Moscow State University (1978), 143–144.

[B2] A.Ya. Belov, *On non-Specht varieties*, Fund. Appl. Math. **5** (1) (1999), 47–66.

[BLH] L.A. Bokut, I.V. L'vov and V.K. Harchenko, *Non-commutative rings*, Algebra 2, Fund. Napr., Vol. 18, VINITI, Moscow (1988), 5–116.

[BS] S. Burris and H.P. Sankappanavar, *A Course in Universal Algebra*, Springer, Berlin (1981).

[BSS] Yu.A. Bahturin, A.M. Slin'ko and I.P. Shestakov, *Non-associative rings*, Algebra. Topology. Geometry, Vol. 18, VINITI, Moscow (1981), 3–72.

[C] P.M. Cohn, *Universal Algebra*, Reidel, Dordrecht (1981).

[DF] A. Day and R. Freese, *A characterization of identities implying congruence-modularity I*, Canad. J. Math. **32** (5) (1980), 1140–1167.

[FM] R. Freese and R. McKenzie, *Commutator Theory for Congruence-Modular Varieties*, Lecture Notes Ser., Vol. 125, London (1987).

[GT] O.G. Garcia and W. Taylor, *The lattice of onterpretability of types of varieties*, Mem. Amer. Math. Soc. **50** (305) (1984).

[GZ] A. Giambruno and M.M. Zaicev, *Exponential growth of codimensions of* PI*-algebras*, Kurosh Algebraic Conference'98. Abstract of Talks, Moscow State University (1978), 54–55.

[Gr] G. Gratzer, *Universal Algebra*, Springer, Berlin (1979).

[Gri] A.V. Grishin, *Examples of non-finitely based T-spaces and T-ideals in characteristic* 2, Fund. Appl. Math. **5** (1) (1999), 101–118.

[HM] D. Hobby and R. McKenzie, *The structure of finite algebras*, Contemp. Math. **76** (1988).

[Il] A.V. Iltyakov, *On finite basis of identities of Lie algebra representations*, Nova J. Algebra Geometry **3** (1992).

[Iv1] S.V. Ivanov, *The free Burnside groups of sufficiently large exponents*, Int. J. Algebra Comp. **4** (1994), 1–308.

[Iv2] S.V. Ivanov, *On the Burnside problem for groups of even exponent*, Doc. Math. J. DMV. Extra volume ICM II (1998), 67–75.

[IO] S.V. Ivanov and A.Yu. Olshansky, *On finite and locally finite subgroups of free Burnside groups of large even exponent*, J. Algebra **195** (1997), 241–284.

[I] Izv. Vyssh. Uchebn. Zaved. Mat., No. 6 (1989) (a special issue on varieties).

[J] N. Jacobson, *Structure of Rings*, Amer. Math. Soc. Colloq. Publ., Vol. 37 (1968).

[Ji] G.I. Jitomirsky, *Basic Notion in Universal Algebra*, Saratov Univ., Saratov (1981).

[Jo] B. Jonsson, *Topics in Universal Algebra*, Lecture Notes in Math., Vol. 250, Springer, Berlin (1972).

[JSSS] K. Jevlakov, A.M. Slin'ko, I.P. Shestakov and A.I. Shirshov, *Rings Close to Associative*, Nauka, Moscow (1978).

[K] A.G. Kemer, *Ideals of Identities of Associative Algebras*, Transl. Math. Monogr., Vol. 87, Amer. Math. Soc. (1991).

[Kn] A. Knoebel, *The equational classes generated by single functionally complete algebra*, Mem. Amer. Math. Soc. **332** (1979).

[KC] P.H. Krauss and D.M. Clark, *Global subdirect products*, Mem. Amer. Math. Soc. **210** (1979).

[Ko] A.I. Kostrikin, *Around Burnside*, Nauka, Moscow (1987).
[Ku1] A.G. Kurosh, *Lectures in General Algebra*, Nauka, Moscow (1973).
[Ku2] A.G. Kurosh, *General Algebra: Lectures 1969–1970 Academic Year*, Nauka, Moscow (1974).
[KS] I.N. Kuz'min and I.P. Shestakov, *Non-associative structures*, Algebra 6, Fund. Napr., Vol. 57, VINITI, Moscow (1990), 179–266.
[L] I.G. Lysenok, *Infinite Burnside groups of even period*, Izv. RAN, Math. **60** (1996), 3–224.
[Mal1] A.I. Mal'cev, *Algebraic Systems*, Nauka, Moscow (1970).
[Mal2] A.I. Mal'cev, *Proceedings*, Vol. 2, Math. Logic and General Theory of Algebraic Systems, Nauka, Moscow (1976).
[Man] E.G. Manes, *Algebraic Theories*, Springer, Berlin (1976).
[Mc] G. McNulty, *An equational logic sampler*, Lecture Notes in Comput. Sci., Vol. 355 (1989), 234–262.
[MMT] R.N. McKenzie, G.F. McNulty and W.F. Taylor, *Algebras, Lattices, Varieties I*, Col. Math. Ser. & Software, The Wardsdorth & Brooks (1989).
[Mov] Yu.M. Movsisyan, *Introduction to the Theory of Algebras with Superidentities*, Erevan Univ., Erevan (1986).
[N] H. Neumann, *Varieties of Groups*, Springer, Berlin (1967).
[P] S.V. Pchelincev, *Solvability and nilpotence of alternative algebras and* $(-1, 1)$*-algebras*, Groups and Other Algebraic Systems with Finiteness Conditions, Nauka, Novosibirsk (1984), 81–101.
[Pi1] A.G. Pinus, *Congruence-Modular Varieties*, Irkutsk. Univ., Irkutsk (1986).
[Pi2] A.G. Pinus, *Congruence-distributive varieties of algebras*, Algebra. Topology. Geometry, Vol. 26, VINITI, Moscow (1988), 45–83.
[Pl] B.I. Plotkin, *Universal Algebra, Algebraic Logic and Data Bases*, Nauka, Moscow (1991).
[Po] S.V. Polin, *On identities in congruence lattices of universal algebras*, Mat. Zametki **22** (3) (1977), 443–451.
[Pr] C. Procesi, *Rings with Polynomial Identities*, M. Dekker (1973).
[R] Ring Theory 1989, Israel Math. Conf. Proc., Vol. 1 (1989).
[Ra] Yu. P. Razmyslov, *Identities of Algebras and Their Representations*, Nauka, Moscow (1989).
[Re] A. Regev, *Existence of identities in* $A \otimes B$, Israel J. Math. **11** (1972), 131–152.
[Ro] L. Rowen, *Polynomial Identities in Ring Theory*, Academic Press, London (1980).
[S] L.A. Skornyakov, ed., *General Algebra*, Vols 1, 2, Nauka, Moscow (1990–1991).
[Shch1] V.V. Shchigolev, *Finite base* T*-spaces over fields of characteristic zero*, Abstract of Talks of International Seminar Dedicated to the Memory of Prof. L.A. Skornyakov, Volgograd, Sept. 6–11, 1999, Volgograd (1999), 75–76.
[Shch2] V.V. Shchigolev, *Examples of non-finitele based* T*-ideals*, Fund. Appl. Math. **5** (1) (1999), 307–312.
[Sm] D.M. Smirnov, *Mal'cev Classes*, Nauka, Moscow (1992).
[ShV] L.N. Shevrin and M.V. Volkov, *Identities in semigroups*, Izv. Vyssh. Uchebn. Zaved. Mat. **11** (1985), 3–47.
[Sz] A. Szendrei, *Clones in Universal Algebra*, Montreal Univ., Montreal (1986).
[U] *Usporiadane Mnoziny a Zvazy*, Vol. 2, Unov. Komenskeho, Bratislava (1988).
[Um] U.U. Umirbaev, *Spechtness of varieties of soluble alternative algebras*, Algebra and Logic **24** (2) (1985), 226–239.

This Page Intentionally Left Blank

Section 4E
Lie Algebras

This Page Intentionally Left Blank

Infinite-dimensional Lie Superalgebras

Yuri Bahturin

Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NF, Canada A1C 5S7

Department of Mechanics and Mathematics, Moscow State University, Moscow, 119899, Russia

Alexander A. Mikhalev

Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong

Mikhail Zaicev

Department of Mechanics and Mathematics, Moscow State University, Moscow, 119899, Russia

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This survey paper describes an area which is comparatively new and there are only few references which can be used to recover the proofs of the results presented here. One of the main sources is the joint monograph [BMPZ], which has appeared with Walter de Gruyter Co, Berlin. Actually, we've made an extensive use of the monograph and in the case of any difficulty we advise that the reader should consult this source. Other references are given at the end of the present paper.

1. Basic definitions and results

1.1. *Defining identities*

In what follows K denotes a field of characteristic different from 2 or 3. All algebras to be considered are vector spaces over K unless otherwise specified. These vector spaces usually will be graded by elements of a commutative additive group G. A homogeneous mapping of degree g, $\varphi : V = \bigoplus_{g\in G} V_g \to V' = \bigoplus_{g\in G} V'_g$, is a linear mapping of vector spaces such that $\varphi(V_h) \subset V'_{g+h}$ for all $h \in G$. Some of these algebras will be G-graded associative algebras of the form

$$A = \sum_{g\in G} A_g$$

and then $A_g A_h \subset A_{g+h}$ for all $g, h \in G$. We will be endowing such algebras with a derived operation defined on homogeneous elements, i.e. those in the homogeneous components A_g, $g \in G$. The nature of these derived operations will be determined by an element from the collection of all bilinear alternating forms $\varepsilon : G \times G \to K^*$. Written elementwise, each of these forms satisfies the following relations: for any $g, h, k \in G$ one has

$$\begin{aligned} \varepsilon(g, h+k) &= \varepsilon(g,h)\varepsilon(g,k), && (1)\\ \varepsilon(g+h, k) &= \varepsilon(g,k)\varepsilon(h,k), && (2)\\ \varepsilon(h,g) &= \varepsilon(g,h)^{-1}. && (3) \end{aligned}$$

Given an ε as just above, the (ε)-commutator is given by

$$[x,y] = xy - \varepsilon(g,h)yx, \quad \text{where } x \in A_g,\ y \in A_h. \tag{4}$$

For other pairs of elements the commutator is defined by linearity. A simple verification shows that the following identities (better, homogeneous identities) hold:

$$\begin{aligned} [x,y] &= -\varepsilon(g,h)[y,x], && (5)\\ \big[[x,y],z\big] &= \big[x,[y,z]\big] - \varepsilon(g,h)\big[y,[x,z]\big], && (6) \end{aligned}$$

where $x \in A_g$, $y \in A_h$, $z \in A$. The same set of identities defines the class of *(ε)-colour Lie superalgebras*. By such an algebra we mean a G-graded vector space

$$L = \sum_{g \in G} L_g,$$

where $[\,,\,]$ is a multiplication (not necessarily derived) satisfying the homogeneous identities (5), (6) for any $x \in L_g$, $y \in L_h$, $z \in L$. The identity (6) is the colour Jacobi identity.

1.2. *Ordinary Lie superalgebras*

The 'classical' case is that of $G = \mathbf{Z}_2$, $\varepsilon(1,1) = -1$, $\varepsilon(i,j) = 1$ for $(i,j) \neq (1,1)$. Such an algebra is called an (ordinary) Lie superalgebra without the adjective 'colour'. Ordinary Lie superalgebras find their origin in algebraic topology and later found some applications in various fields of mathematics and (even!) physics ([Ber]). Certain investigations in physics have led to an extension of the basic concept of an ordinary Lie superalgebra ([RW2]). It should be noted, however that in a paper of M. Scheunert [Sc1] it was shown that by modifying the operation of a colour Lie superalgebra one can reduce some questions about these algebras to those about ordinary Lie superalgebras. Such a modification is performed by setting

$$[x, y]_\sigma = \sigma(g, h)[x, y], \quad x \in L_g, \ y \in L_h, \ \sigma \in Z^2(G, K^*), \tag{7}$$

where $Z^2(G, K^*)$ stands for the set of all 2-cocycles of the group G with values in the multiplicative group of the ground field K. Now the grading group G splits into the disjoint union $G = G_+ \cup G_-$ where

$$\begin{aligned} G_+ &= \{g \in G \mid \varepsilon(g, g) = 1\}, \\ G_- &= \{g \in G \mid \varepsilon(g, g) = -1\}. \end{aligned}$$

(It is immediate from the last equation in (3) that 1 and -1 are the only values allowed for ε.) If we set

$$L_+ = \sum_{g \in G_+} L_g, \qquad L_- = \sum_{g \in G_-} L_g \tag{8}$$

then $L = L_+ \oplus L_-$ is a $\mathbf{Z}_2$-grading and $[\,,\,]_\sigma$ makes this graded algebra into an (ordinary) Lie superalgebra. It is customatry to call the elements in L_+ *even* and those in L_- *odd*. The following are the identities valid for the homogeneous $x \in L_+$ and $y \in L_-$:

$$[y, [y, x]] = 2[y^2, x], \qquad [y, [y, y]] = 0. \tag{9}$$

The identities (9) are exactly the place where the restrictions on the characteristic of the ground field arise. (Actually, if we want to work in either of these characteristic we must adjoin (9) to the defining laws (6).)

In the above Scheunert's work the grading group G was assumed finitely generated. Recently in [BMo] it was shown that G can be taken arbitrary. Similar result was proved in [Pa2].

1.3. *Homogeneity*

An important remark about the general concepts of the theory is that they are meant to be *homogeneous*. Thus, any subalgebra $M \subset L$ is a G-graded subspace $M = \sum_{g\in G} M_g$ where $M_g = M \cap L_g$ for any $g \in G$. Also, a homomorphism $\varphi: L \to L'$ is an ordinary operation-preserving linear operator but with $\varphi(L_g) \subset L'_g$ for all $g \in G$. In this setting a *representation* $\rho: L \to \operatorname{End} V$, $V = \sum_{g\in G} V_g$, is a linear mapping such that

$$\rho(L_g) \subset (\operatorname{End} V)_g = \{\mathcal{A}: V \to V \mid \mathcal{A}(V_h) \subset V_{g+h},\ \forall h \in G\}$$

for any $g \in G$, and

$$\rho([x, y]) = \rho(x)\rho(y) - \varepsilon(g, h)\rho(y)\rho(x), \quad \forall x \in L_g,\ \forall y \in L_h.$$

A related notion, that of an L-module, provides us with a vector space $V = \sum_{g\in G} V_g$ and a bilinear pairing $(x, v) \mapsto xv \in V_{g+h}$ for any $x \in L_g$, $v \in V_h$. The *derivation algebra* $D = \operatorname{Der} P$ of an arbitrary G-graded algebra $P = \sum_{g\in G} P_g$, which is, in general, neither associative, nor Lie, is the linear hull of all the spaces D_g whose elements are the homogeneous mappings of degree g such that

$$d(xy) = (dx)y + \varepsilon(g, h)x(dy) \tag{10}$$

with $x \in P_h$ and y arbitrary in P. There is a natural homomorphism $\operatorname{ad}: L \to \operatorname{Der} L$, L a colour Lie superalgebra, defined by $(\operatorname{ad} x)(y) = [x, y]$ for any $x, y \in L$. The derivations in $\operatorname{ad} L$ are called *adjoint* or *inner*.

1.4. *Enveloping algebras*

Consider an associative G-graded algebra $A = \bigoplus_{g\in G} A_g$. As previously, A becomes a colour Lie superalgebra if we introduce a "colour" commutator: $[x, y] = xy - \varepsilon(g, h)yx$, where x, y are homogeneous elements $d(x) = g$, $d(y) = h$. We denote this superalgebra by $[A]$ or, more precisely, $[A]_\varepsilon$. A homomorphism f of a colour Lie superalgebra L into A is defined as a homomorphism L into $[A]$. If f is injective then we say that A is an enveloping associative algebra for L.

One of the most important concepts which is a powerful tool in the study of Lie superalgebras is the (universal) *enveloping algebra* $U(L)$ of a colour Lie superalgebra L. In this subsection we will be using terminology "enveloping algebra" to denote the universal enveloping algebra. We hope there will be no confusion where the enveloping algebras appear in the extended sense as in the beginning of this subsection. The (universal) enveloping algebra can be defined as the universal object in the category of arrows $L \xrightarrow{\pi} A$ where A is an associative G-graded algebra with unity 1, G the same grading group as for L, and π a homomorphism of colour Lie superalgebras from L into $[A]_\varepsilon$ (see Subsection 1.1). Thus $U(L)$ is a G-graded associative algebra with a Lie superalgebra homomorphism $\iota : L \to U(L)$ such that for any Lie superalgebra homomorphism $\pi : L \to A$, A as above, there exists a unique homomorphism $f : U(L) \to A$ (which is also a Lie superalgebra homomorphism) with $f(1) = 1$ making the following diagram commutative:

$$\begin{tikzcd} L \arrow[r, "\iota"] \arrow[dr, "\pi"'] & U(L) \arrow[d, "f"] \\ & A \end{tikzcd}$$

The existence and the uniqueness of $U(L)$ being standard, one has an extension of the classical Poincaré–Birkhoff–Witt Theorem to the case of colour Lie superalgebras (we formulate it in the case of fields although it is of a more general nature, e.g., for algebras which are free modules over an arbitrary commutative ground ring K).

THEOREM 1.1. *Let $L = \sum_{g \in G} L_g$ be an (ε)-Lie superalgebra over the ground field K and $E = \bigcup_{g \in G} E_g$ a basis of L which is the union of bases E_g for each L_g, $g \in G$. Set $E_+ = \bigcup_{g \in G_+} E_g$, $E_- = \bigcup_{g \in G_-} E_g$ and make E totally ordered. Then a basis of $U(L)$ is formed by 1 jointly with the set of all monomials of the form*

$$\iota(x_1) \cdots \iota(x_m)\iota(y_1) \cdots \iota(y_n), \tag{11}$$

where $x_1 \leqslant \cdots \leqslant x_m$, $x_i \in E_+$, $y_1 < \cdots < y_n$, $y_j \in E_-$.

It follows that ι is a monomorphism which enables us to identify L with its image in $U(L)$ under ι.

The importance of $U(L)$ is explained by a simple result saying that every L-module is a left $U(L)$-module, the converse being also true. Moreover, all properties of an L-module are the same as those of $U(L)$-module. Thus, a module over an Abelian Lie superalgebra $L = L_0 \oplus L_1$ is, in fact, a module over its enveloping algebra

$$U(L) \cong K[L_0] \otimes \Lambda(L_1),$$

where $K[L_0]$ means the polynomial (or symmetric) algebra over L_0 and $\Lambda(L_1)$ the Grassmann (exterior) algebra of L_1. This latter construction generalizes to the case of arbitrary ε-colour Lie superalgebras, $\varepsilon : G \times G \to F$, in the following way.

A corollary to the above theorem relates to the *associated graded algebra* $\operatorname{gr} U(L)$. This is given by the natural ascending filtration $\{U_n(L)\}_{n=-1,0,1,\ldots}$ of $U(L)$ given by

$$U_{-1} = 0, \ U_0 = K \cdot 1, \ U_1 = U_0 + L, \ldots, U_{n+1} = U_n + U_n L, \ldots.$$

Set

$$\mathrm{gr}_n U(L) = U_n / U_{n-1}, \quad n \geqslant 0.$$

Then

$$\mathrm{gr}\, U(L) = \sum_{n=0}^{\infty} \mathrm{gr}_n(L)$$

becomes an algebra in an obvious natural way:

$$(a + U_{k-1})(b + U_{l-1}) = ab + U_{k+l-1}, \quad \text{if } a \in U_k,\ b \in U_l.$$

Let us also define $K_\varepsilon[L_+]$ and $\Lambda_\varepsilon(L_-)$ as follows. Given a G-homogeneous basis $e_1, \ldots, e_m, \ldots$ of L_+ and $f_1, \ldots, f_n, \ldots$ of L_- we define $K_\varepsilon[L_+]$ as the associative algebra generated by $e_1, \ldots, e_m, \ldots$ subject to the defining relations $e_i e_j = \varepsilon(g, h) e_j e_i$ where $e_i \in L_g$, $e_j \in L_h$. If $g = h$ then $\varepsilon(g, h) = 1$ and we have usual relations of polynomial algebras. Also $\Lambda_\varepsilon(L_-]$ is given in the same way, i.e. by generators $f_1, \ldots, f_n, \ldots$ and defining relations $f_i f_j = \varepsilon(g, h) f_j f_i$, $f_i \in L_g$, $f_j \in L_h$. If $g = h$ then $\varepsilon(g, h) = -1$ and we have usual relations of an exterior algebra.

We also need an $(\varepsilon - G)$-graded version of the tensor product $A \otimes_\varepsilon B$ which is an ordinary vector space tensor product where the multiplication is determined by setting

$$(a \otimes b)(a' \otimes b') = \varepsilon(g, h) aa' \otimes bb'$$

if $b \in B_g$, $a' \in A_h$. Now the corollary in question reads as follows:

PROPOSITION 1.1. *Let K be a field,* $\mathrm{char}\, K \neq 2, 3$.

(1) *Suppose that L is a colour Lie superalgebra with a G-homogeneous basis $E = E_+ \cup E_-$ and that $U(L)$ is its universal enveloping algebra. Then* $\mathrm{gr}\, U(L) \cong K[E_+, \varepsilon] \otimes_\varepsilon \Lambda[E_-, \varepsilon]$.
(2) *Let $p = \mathrm{char}\, K > 3$. If L is a colour Lie p-superalgebra and $u(L)$ is its restricted enveloping algebra then for any G-homogeneous linear basis $E = E_+ \cup E_-$ in L we have* $\mathrm{gr}\, u(L) \cong (K[E_+, \varepsilon]/I) \otimes_\varepsilon \Lambda[E_-, \varepsilon]$.
(3) *The universal enveloping algebra $U(L)$ of a colour Lie algebra L (i.e. $L = L_+$) has no zero divisors.*
(4) *Let* $\dim L < \infty$. *Then the algebra $U(L)$ is left and right Noetherian.*
(5) *If $L = L_+$, then the Jacobson radical $J(U(L))$ is equal to zero.*

1.5. *Restricted Lie superalgebras*

In the case where the ground field K is of characteristic $p > 3$ one can define the so called *restricted* Lie superalgebras (or p-Lie-superalgebras). By this we mean a colour Lie superalgebra $L = \sum_{g\in G} L_g = L_+ \oplus L_-$ in which an additional unary operation is defined, acting on homogeneous even elements only. Thus, with each element $x \in L_g$, $g \in G_+$, one associates an $x^{[p]} \in L_{pg}$ such that the following axioms hold:

$$(\lambda x)^{[p]} = \lambda^p x^{[p]}, \quad \lambda \in K, \tag{12}$$

$$\left[x^{[p]}, y\right] = (\operatorname{ad} x)^p(y), \quad y \in L, \tag{13}$$

$$(x+y)^{[p]} = x^{[p]} + y^{[p]} + \sum_{i=0}^{p} s_i(x, y), \quad y \in L_g, \tag{14}$$

where $is_i(x, y)$ is the coefficient of t^{i-1} in the polynomial $(\operatorname{ad}(tx+y))^{p-1}(x)$. If $G = \{0\}$ then we arrive at a standard definition of restricted Lie algebras.

The relations just listed are motivated by an observation that they are trivially satisfied in the Lie superalgebras of the form $[A]$, A an associative superalgebra for the operation $x^{[p]} = x^p$, where the right hand side term denotes the usual raising to the p-th power in A. The restricted algebra thus obtained is denoted by $[A]_p$.

Similar to the notion of the enveloping algebra one can define *restricted enveloping algebras* $u(L)$ each being the quotient of $U(L)$ modulo the relations of the form

$$x^{[p]} - x^p, \quad x \in L_g,\ g \in G_+.$$

The importance of $u(L)$ is that this is exactly the associative algebra which is generated by the image of L in End V under a *restricted* representation of L by linear operators of a vector space V. By such a representation of L in End V we mean an ordinary representation $\rho : L \to [\operatorname{End} V]$ with additionally $\rho(x^{[p]}) = \rho(x)^p$ for $x \in L_g$, $g \in G_+$. The restricted version of the PBW-Theorem (usual abbreviation for Poincaré–Birkhoff–Witt), says that, given, e.g., a finite basis $e_1, \ldots, e_m, f_1, \ldots, f_n$ of L with e_i homogeneous even and f_j homogeneous odd, a basis of $u(L)$ is formed by all monomials of the form

$$e_1^{k_1} \cdots e_m^{k_m} f_1^{l_1} \cdots f_n^{l_n}, \quad 0 \leqslant k_i < p,\ 0 \leqslant l_j < 1,$$

where, as usual, $x^0 = 1$. Again, $L \subset u(L)$ and any p-Lie-superalgebra mapping of L into an associative algebra B extends to a homomorphism of associative algebras with unit.

1.6. *Classes of colour Lie superalgebras*

Some classes of colour Lie superalgebras are of special interest. Among them we mention just few, such as Abelian, nilpotent, soluble, simple, semisimple and free Lie superalgebras. To define these we call a Lie superalgebra of dimension greater than 1 *simple* if it does not contain proper nontrivial ideals, i.e. those different from the zero subalgebra and the superalgebra itself. The classification of simple finite-dimensional ordinary Lie superalgebras over an algebraically closed field of characteristic zero has been given by V.G. Kac (see [Kac1,Sc2]). Now we have defined *Abelian* superalgebras as those with trivial multiplication $[x, y] = 0$ for all x, y. To pass to further classes it is convenient to introduce the mutual commutator $[M, N]$ of two subspaces of a Lie superalgebra L as the linear span of all commutators of the form $[x, y]$ where $x \in M$, $y \in N$. If M, N are two ideals of L (to specify this we write: $M \lhd L$, $N \lhd L$, then the mutual commutator is also an ideal of L. Thus the following are all ideals of the ambient Lie superalgebra. Let us define the powers L^n, $n = 1, 2, \dots$, of a Lie superalgebra L inductively by setting $L^1 = L$ and, for $n > 0$, by $L^n = [L^{n-1}, L]$. There results descending series of ideals in L:

$$L = L^1 \supset L^2 \supset \cdots \supset L^n \supset \cdots \tag{15}$$

called the *lower central series* of L. The terminology is explaind by the fact that (15) is the 'least' series with the property that each L^n/L^{n+1} lies in the center of L/L^{n+1}. Also, by the Jacobi identity (6) it is easy to verify $[L^k, L^l] \subset L^{k+l}$ for all $k, l \in \mathbf{N}$. It may well happen, of course, that for some superalgebras we always have $L = L^1 = L^2 = \cdots$, as, e.g., for the simple superalgebras. However, in a number of cases of importance for the theory we have $L^n = \{0\}$ for some natural n. In this case we say that L is *nilpotent* and the least c with $L^{c+1} = \{0\}$ is called the *nilpotency class* of L. It should be noted that the same algebras may have different nilpotency classes depending on their being considered as ordinary Lie algebras or as Lie superalgebras. For instance, the Grassmann algebra $\Lambda(V)$ of a vector space V with $\dim V > 1$ (see Subsection 1.4) is nilpotent in both cases but it is nilpotent of the nilpotency class 2 as a Lie algebra and of class 1 as a Lie superalgebra. In a number of other cases we never have $L^n = \{0\}$ but $\bigcap_{n=1}^{\infty} L^n = \{0\}$. In this case we say that L is *residually nilpotent*. One of the examples of residually nilpotent Lie superalgebras is supplied by the free Lie superalgebras (see below).

The derived series $\{L^{(n)}\}$, $n = 0, 1, \dots$, of a Lie superalgebra L is defined by induction if one sets $L^{(0)} = L$ and, for $n > 0$, $L^n = [L^{(n-1)}, L^{(n-1)}]$. Again there arises a descending series

$$L = L^{(0)} \supset L^{(1)} \supset \cdots \supset L^{(n)} \supset \cdots. \tag{16}$$

A Lie superalgebra L with $L^{(n)} = \{0\}$ is called *solvable* and the least $l \geqslant 0$ with $L^{(l)} = \{0\}$ is called the *solvability length* of L. Being solvable is equivalent to being a finite extension of a sequence of Abelian superalgebras. Nilpotent superalgebras are obviously solvable. In every finite-dimensional superalgebra there exists a largest nilpotent ideal and a largest solvable ideal called its nilpotent and solvable *radicals*, respectively. A finite-dimensional ordinary Lie superalgebra $L = L_0 \oplus L_1$ is solvable if, and only if this is true for L_0. Note that in the case of usual Lie algebras any finite-dimensional solvable Lie algebra over a field K of characteristic zero is isomorphic to a Lie subalgebra of an algebra of triangular matrices of some order over K. The structure of infinite-dimensional Lie superalgebras is far more difficult and some information will be given later in Section 5. There is a class of solvable Lie superalgebras, however, where the progress is more visible. These are solvable Lie superalgebras of solvability length 2, called *metabelian*. A common way of obtaining such superalgebras is via *semidirect products*. Given two Lie superalgebras M and N and a homomorphism φ of M into the derivation algebra of N, one defines an operation on $M \oplus N$ by setting

$$[m+n, m'+n'] = [m, m'] + \varphi(m)(n') - \varepsilon(g,h)\varphi(m')(n), \quad \text{where } n \in N_g,\ m' \in M_h. \tag{17}$$

For instance, if M is a Lie superalgebra and N a module of it then considering N as an Abelian Lie superalgebra and φ as the representation of M in End N we obtain a Lie algebra $L = M\lambda N$ where M is a subalgebra and N is an Abelian ideal. Of special importance is the case where N is a *free* M-module with free generating set X. Denoting by A the Abelian Lie superalgebra with basis X we call the product as just above the *wreath product* of M and A which is denoted by $A \operatorname{wr} M$. This terminology comes from group theory (see [Ne]).

Wider classes of Lie superalgebras can be obtained by considering Lie superalgebras which are *locally* in one of the classes defined above. We say that a Lie superalgebra L is locally in a class $\mathcal{K}$ if every finite set of elements of L lies in a subalgebra $M \in \mathcal{K}$. Thus, every Lie superalgebra has a locally nilpotent radical, i.e. the largest locally nilpotent ideal. Some other so called *finiteness conditions* will be considered in Section 5.

A Lie superalgebra L is called *semisimple* if it has no non-zero solvable ideals (this is equivalent to having non-zero Abelian ideals). Obviously, any simple Lie superalgebra is semisiple. In the case of finite-dimensional Lie *algebras* it is true that any semisimple Lie algebra is the direct product of its simple ideals. This is not true already in the case of Lie algebras with positive characteristic, not mentioning Lie superalgebras where it is wrong already for finite-dimensional ordinary Lie superalgebras over a field of characteristic zero.

Finally, we define free Lie algebras. Given a non-empty G-graded set $X = \bigcup_{g\in G} X_g$ we define *the free colour Lie superalgebra* $L(X)$ with *free generating set* X as a Lie superalgebra generated by X whose grading is compatible with that of X and such that any graded mapping $\varphi : X \to M = \sum_{g\in G} M_g$ extends to a homomorphism $\overline{\varphi} : L(X) \to M$ such that $\overline{\varphi}\,|_X = \varphi$. Every Lie superalgebra is isomorphic to a quotient algebra of a suitable free Lie superalgebra. If $M \cong L(X)/I$ is such an isomorphism with the ideal I generated by a subset R of L then we write $M = \langle X \mid R\rangle$ and say that M is given *in terms of generators and defining relations* and that $\langle X \mid R\rangle$ its *presentation*. Effectively determining properties of a Lie superalgebra from its presentation is often impossible. For example, there is no algorithm to decide whether such a presentation determines the zero algebra (already in the case of ordinary Lie algebras).

1.7. *Simple colour Lie superalgebras*

The classification problem of simple algebras is one of the most important tasks for any class of algebras. For Lie algebras and ordinary Lie superalgebras this classification is complete in the finite-dimensional case over any algebraically closed fields of characteristic zero (see, for example, [Kac3] or [Sc2] for Lie superalgebras). In the general case we illustrate this problem with one recent result.

THEOREM 1.2 [Pa2]. *Suppose A is a color commutative G-graded K-algebra and let D be a nonzero color commutative G-graded K-vector space of color derivations of A. Let* $\dim KD = 1$ *and assume that either* $\operatorname{char} K = 2$ *or* $\operatorname{char}^- K \neq 2$ *and* $D = D_-$. *Then $A \otimes D = AD$ is a simple Lie color algebra if and only if A is graded D-simple and* $D(A) = A$.

Recall that $V_- = \sum_{g \in G_-} V_g$ for any G-graded vector space V and $G_- = \{x \in G\colon \varepsilon(x,x) = -1\}$.

2. Bases of colour Lie superalgebras

The main sources of the theory of free Lie algebras are in the papers of P. Hall [Hall P.], W. Magnus [Ma1,Ma2], and E. Witt [Wi1] devoted to the study of free groups. The construction of linear basis for free Lie algebras has been suggested in the article [Hall M.] of M. Hall (so called "Hall basis"). Later A.I. Shirshov [Shi2,Shi5] has given some other constructions of free linear bases. Lyndon basis was introduced in [CFL] by K.T. Chen, R.H. Fox, and R.C. Lyndon. A.I. Shirshov [Shi3] also invented the composition method for Lie algebras which became an extremely useful tool in ring theory for solving quite a few algorithmic problems; he considered in [Shi4] free products of Lie algebras with amalgamated subalgebra.

2.1. *Free colour Lie (p-)superalgebras*

Let $X = \bigcup_{g \in G} X_g$ be a G-graded set, i.e. $X_g \cap X_f = \emptyset$ for $g \neq f$, $d(x) = g$ for $x \in X_g$; let also $\Gamma(X)$ be the groupoid of nonassociative monomials in the alphabet X, $u \circ v = (u)(v)$ for $u, v \in \Gamma(X)$, and $S(X)$ be the free semigroup of associative words with the bracket removing homomorphism $\bar{\ }: \Gamma(X) \to S(X)$; $[u]$ being an arrangement of brackets on the word $u \in S(X)$. For $u = x_1 \dots x_n \in S(X)$, $x_i \in X$ we define the word length as $l(u) = n$, and the degree as $d(u) = \sum_{i=1}^n d(x_i) \in G$. For $z \in \Gamma(X)$ we set $l(z) = l(\bar{z})$ and $d(z) = d(\bar{z})$.

Let K be a commutative associative ring with 1, $A(X)$ and $F(X)$ the free associative and nonassociative K-algebras, respectively,

$$S(X)_g = \{u \in S(X),\ d(u) = g\}, \qquad \Gamma(X)_g = \{v \in \Gamma(X),\ d(v) = g\},$$

let $A(X)_g$ and $F(X)_g$ be the K-linear spans of the subsets $S(X)_g$ and $\Gamma(X)_g$ respectively, $A(X) = \bigoplus_{g \in G} A(X)_g$ and $F(X) = \bigoplus_{g \in G} F(X)_g$ being the free G-graded associative and nonassociative algebras respectively. Let now I be the G-graded ideal of $F(X)$ generated by the homogeneous elements of the form

$$a \circ b + \varepsilon(a,b) b \circ a$$

and

$$(a \circ b) \circ c - a \circ (b \circ c) + \varepsilon(a,b) b \circ (a \circ c),\ e \circ e,\ f \circ (f \circ f),$$

where $a, b, c, e, f \in \Gamma(X)$, $d(e) \in G_+$, $d(f) \in G_-$; then $L(X) = F(X)/I$ is the free colour Lie K-superalgebra (i.e. every G-map φ of degree zero from X into any colour Lie K-superalgebra R with the same group G and form ε $(d(\varphi(x)) = d(x), x \in X)$ can be uniquely extended to a colour Lie superalgebra homomorphism $\overline{\varphi}: L(X) \to R$). It is clear that this universal property defines the free colour Lie superalgebra $L(X)$ uniquily up to a colour Lie superalgebra isomorphism. For $u \in F(X)$ we set $\widetilde{u} = u + I \in L(X)$.

Suppose that the set $X = \bigcup_{g \in G} X_g$ is totally ordered and the set $S(X)$ is ordered lexicographically, i.e. for $u = x_1 \ldots x_r$ and $v = y_1 \ldots y_m$ where $x_i, y_j \in X$ we have $u < v$ if either $x_i = y_i$ for $i = 0, 1, \ldots, t-1$ and $x_t < y_t$ or $x_i = y_i$ for $i = 1, 2, \ldots, m$ and $r > m$. Then $S(X)$ is linearly ordered.

A word $u \in S(X)$ is said to be *regular* if for any decomposition $u = ab$, where $a, b \in S(X)$, we have $u > ba$. The word $w \in S(X)$ is said to be *s-regular* if either w is a regular word or $w = vv$ with v a regular word and $d(v) \in G_-$.

A monomial $u \in \Gamma(X)$ is said to be *regular* if either $u \in X$ or:

(a) from $u = u_1 \circ u_2$ if follows that u_1, u_2 are regular monomials with $\overline{u}_1 > \overline{u}_2$;

(b) from $u = (u_1 \circ u_2) \circ u_3$ it follows that $\overline{u}_2 \leqslant \overline{u}_3$.

Observe that if $x \in X$, v is a regular monomial and $x > \overline{v}$ then $x(v)$ is a regular monomial.

A monomial $u \in \Gamma(X)$ is said to be *s-regular* if either u is a regular monomial or $u = (v)(v)$ with v a regular monomial and $d(v) \in G_-$.

If u is an s-regular monomial, then $\overline{u}$ is an s-regular word. Conversely, if u is an s-regular word, then there is a unique arrangement of brackets $[u]$ on u such that $[u]$ is an s-regular monomial.

THEOREM 2.1 [Mik1]. *Let K be a commutative ring with* 1; $2 \in K^*$, *and let $L(X) = F(X)/I$ be a free colour Lie superalgebra. Then the cosets in $F(X)/I$ whose representatives are s-regular monomials give us a basis of $L(X)$ as a free G-graded K-module. Moreover, if $[X]$ is the subalgebra of $[A(X)]$ generated by X and $\pi : L(X) \to [X]$ is an homomorphism such that $\pi(x + I) = x$ for all $x \in X$, then π is an isomorphism.*

COROLLARY 2.1.1 (An analogue of the Witt formula). *Let $2 \in K^*$, $X = \bigcup_{g \in G} X_g$, $X_+ = \{x_1, \ldots, x_t\}$, $X_- = \{x_{t+1}, \ldots, x_{t+s}\}$, and let $L(X)$ be the free colour Lie K-superalgebra, $\mu(l)$ the Möbius function, and $W(\alpha_1, \ldots, \alpha_k)$ the rank of the free module of elements of multidegree $\alpha = (\alpha_1, \ldots, \alpha_k)$ in the free Lie algebra of rank k, then*

$$W(\alpha_1, \ldots, \alpha_k) = \frac{1}{|\alpha|} \sum_{e|\alpha_i} \mu(e) \frac{(|\alpha|/e)!}{(\alpha_1/e)! \cdots (\alpha_k/e)!},$$

where $|\alpha| = \sum_{i=1}^{k} \alpha_i$ (Witt formula). Let $SW(\alpha_1, \ldots, \alpha_{t+s})$ be the rank of the free module of elements of multidegree $\alpha = (\alpha_1, \ldots, \alpha_{t+s})$ in the free colour Lie superalgebra $L(X)$ of rank $t + s$. Then

$$SW(\alpha_1, \ldots, \alpha_{t+s}) = W(\alpha_1, \ldots, \alpha_{t+s}) + \beta W\left(\frac{\alpha_1}{2}, \ldots, \frac{\alpha_{t+s}}{2}\right),$$

where

$$\beta = \begin{cases} 0 & \textit{if there exists an } i \textit{ such that } \alpha_i \textit{ is odd,} \\ & \textit{or if } \frac{1}{2}\sum_{i=t+1}^{t+s} \alpha_i \cdot d(\alpha_i) \in G_-; \\ 1 & \textit{otherwise.} \end{cases}$$

Some other analogues of the Witt formula can be found in the papers [Bab,Kang,Kant, MT,Ree].

For $w \in L(X)$ we define the length $l_X(w)$ as the greatest length of the s-regular monomials in the presentation of w as a linear combination of s-regular monomials. If $\pi(\widetilde{w}) = \sum_{i=1}^{k} \alpha_i u_i$ where $0 \neq \alpha_i \in K$, $u_i \in S(X)$, $u_i \neq u_j$ for $i \neq j$, then $l_X(w) = \max_{1 \leqslant i \leqslant k}\{l_X(u_i)\}$. Therefore $l_X(w)$ does not depend on the ordering of X.

Let now w_0 denote the leading part of $w \in L(X)$ (i.e. the sum of those summands whose length is equal to $l_X(w)$ in the s-regular presentation of w). If $w = w_0$ then we say that w is l-homogeneous. By analogy, we define the length $l_x(w)$, the leading part and property of being homogeneous relative to $x \in X$. We say that an element $w \in L(X)$ is multihomogeneous if w is homogeneous relative to all $x \in X$. The greatest monomial of the leading part of w will be called the leading term of w.

The cardinality $|X|$ is called *the rank* of the free colour Lie superalgebra $L(X)$. It follows from Theorem 2.1 that

$$\operatorname{rank} L(X) = \dim L(X)/\big[L(X), L(X)\big];$$
$$|X|_g = \dim\big(L(X)/\big[L(X), L(X)\big]\big)_g$$

and therefore, $\operatorname{rank} L(X)$ and $|X_g|$ do not depend on the choice of a set of free generators in $L(X)$.

Consider now bases of free color Lie p-superalgebras. Let K be a field with $\operatorname{char} K > 2$, $X = \bigcup_{g \in G} X_g$ a G-graded set and $A(X)$ the free G-graded associative algebra. We say that an element $w \in A(X)$ is an *as-regular monomial* relative to the alphabet X if $w = \pi(\widetilde{u})$ where u is an s-regular monomial relative to X, and, respectively, a *ps-regular monomial*, if either w is an as-regular monomial or

$$w = \underbrace{\big(\ldots\big(\big(\pi(\widetilde{v})\big)^p\big)^p \ldots\big)^p}_{k \text{ times}} = \big(\pi(\widetilde{v})\big)^{p^k},$$

where $k \in \mathbf{N}$, and v is an s-regular monomial with $d(v) \in G_+$. By definition, ps-regular monomials are just the associative presentations of s-regular monomials and their p-th powers for even s-regular monomials.

Observe that if we consider the operation $a \to a^p$ for even homogeneous element $a \in [A(X)]$ then we have a colour Lie p-superalgebra which will be denoted by $[A(X)]^p$. Let $L^p(X)$ be the subalgebra of $[A(X)]^p$ generated by X. The set of ps-regular monomials give us a linear basis in the colour Lie p-superalgebra $L^p(X)$.

The colour Lie p-superalgebra $L^p(X)$ constructed above is the free colour Lie p-superalgebra with the set $X = \bigcup_{g \in G} X_g$ of free generators (i.e. X generates the algebra $L^p(X)$

and any mapping $\varphi: X \to R$ where R is any colour Lie p-superalgebra with the same G and ε, can be extended to a homomorphism $\overline{\varphi}: L^p(X) \to R$ of colour Lie p-superalgebras).

The number $\operatorname{rank} L^p(X) = |X|$ is called *the rank* of the free colour Lie p-superalgebra. It is clear that the rank of $L^p(X)$ does not depend on the choice of a set of free generators.

2.2. *Free metabelian colour Lie superalgebras*

The problem of finding linear bases is important not only in the case of free Lie algebras but also in the case of free algebras in varieties other than the variety of all Lie superalgebras. What results is a canonical form of the elements in the superalgebra via its generating set. However interesting may this be, a complete solution exists only exceptionally. Here we present the results in the case of metabelian Lie superalgebras.

If in a colour Lie superalgebra L any two commutators commute, i.e. $[[x, y], [z, t]] = 0$ for any $x, y, z, t \in L$, we say that L is *a metabelian colour Lie superalgebra.*

The free metabelian colour Lie superalgebra with a set X of free generators is the algebra $M(X) = L(X)/L^{(2)}(X)$ where $L(X)$ is the free colour Lie superalgebra, $L^{(2)} = [L^{(1)}(X), L^{(1)}(X)]$, $L^{(1)}(X) = [L(X), L(X)]$.

THEOREM 2.2 [BD]. *Let $X = \bigcup_{g\in G} X_g = X_+ \cup X_-$ be a G-graded set, $M(X)$ the free metabelian colour Lie superalgebra in X over a field K,* $\operatorname{char} K \neq 2, 3$. *Then a basis in $M(X)$ is formed by all monomials of the form*

$$
\begin{aligned}
&x_i \in X_+,\ y_j \in X_-,\\
&[x_{i_1}, x_{i_2}, \dots, x_{i_p}, y_{j_1}, \dots, y_{j_q}], \quad i_1 > i_2 \leqslant \dots \leqslant i_p,\ j_1 < j_2 < \dots < j_q,\ p > 1,\\
&[y_{j_1}, y_{j_2}, \dots, y_{j_q}], \quad j_1 \geqslant j_2 < \dots < j_q,\ q > 1,\\
&[y_{j_1}, x_{i_1}, \dots, x_{i_p}, y_{j_2}, \dots, y_{j_q}], \quad i_1 \leqslant \dots \leqslant i_p,\ j_2 < \dots < j_q,\ p, q \geqslant 1.
\end{aligned}
$$

Using this theorem it is possible to compute the Hilbert series of $M(X)$ (see Section 5).

2.3. *Free products with amalgamated subalgebra*

Let K be a field, T a set, and H^0, H_α, $\alpha \in T$, colour Lie K-superalgebras (with the same group G and skew symmetric bilinear form ε), H_α^0 a G-homogeneous subalgebra of H_α, $\delta_\alpha: H^0 \to H_\alpha^0$ an isomorphism of colour Lie superalgebras. We say that a colour Lie K-superalgebra Q (with the group G and the form ε) is *the free product of colour Lie superalgebras H_α, $\alpha \in T$, with amalgamated homogeneous subalgebra* H^0 (notation: $Q = \prod^*_{H^0, \alpha\in T} H_\alpha$) if the following conditions are satisfied:

(1) there are homogeneous subalgebras H'_α in Q, isomorphisms $\sigma_\alpha : H_\alpha \to H'_\alpha$ of colour Lie superalgebras such that $\sigma_\alpha\delta_\alpha(h) = \sigma_\beta\delta_\beta(h)$ for all $\alpha, \beta \in T$ and $h \in H^0$;

(2) if M is a colour Lie K-superalgebra (with the same group G and form ε), $\gamma_\alpha : H_\alpha \to M$ is a homomorphism of colour Lie superalgebras, $\gamma_\alpha \delta_\alpha(h) = \gamma_\beta \delta_\beta(h)$ for all $\alpha, \beta \in T$ and $h \in H^0$, then there is one and only one homomorphism

$$\psi : \prod\nolimits^*_{\substack{H^0 \\ \eta \in T}} H_\eta \to M$$

of colour Lie superalgebras such that $\psi \delta_\alpha = \gamma_\alpha$ for all $\alpha \in T$.

It is clear that the colour Lie superalgebra $\prod^*_{H^0, \alpha \in T} H_\alpha$ is defined uniquely (up to an isomorphism of colour Lie superalgebras).

Let $B^0 = \{f_\beta,\ \beta \in T^0\}$ be a G-homogeneous basis in H^0, $B^0_\alpha = \{e_{\alpha\beta},\ e_{\alpha\beta} = \delta_\alpha(f_\beta),\ \alpha \in T,\ \beta \in T^0\}$, and let B_α be a homogeneous basis in H_α, $B^0_\alpha \subset B_\alpha$ and $B_\alpha = \{e_{\alpha\beta},\ \beta \in T_\alpha\}$ for all $\alpha \in T$. Then the multiplication in H_α for all $\alpha \in T$ can be given using structure constants $p^\tau_{\alpha\beta\gamma} \in K$:

$$[e_{\alpha\beta}, e_{\alpha\gamma}] = \sum_\tau p^\tau_{\alpha\beta\gamma} e_{\alpha\tau}, \quad \beta, \gamma \in T_\alpha.$$

For any $\alpha \in T$ consider the G-graded set $X_\alpha = \{x_{\alpha\beta},\ \beta \in T_\alpha,\ d(x_{\alpha\beta}) = d(e_{\alpha\beta})\}$.

Let X be the set obtained from $\bigcup_{\alpha \in T} X_\alpha$ by the identification $x_{\alpha\gamma} = x_{\beta\gamma}$ for $\gamma \in T^0$. Consider T^0 with a total order and a total order extension of T^0 to T_α for all $\alpha \in T$ such that $\beta > \gamma$ for $\beta, \gamma \in T_\alpha$, $\beta \notin T^0$, $\gamma \in T^0$. Let J be the ideal of the free colour Lie superalgebra $L(X)$ generated by the following subset of G-homogeneous elements

$$D = \{d_{\alpha\beta\gamma}\} = [x_{\alpha\beta}, x_{\alpha\gamma}] - \sum_\tau p^\tau_{\alpha\beta\gamma} x_{\alpha\tau}, \quad \alpha \in T,\ \beta, \gamma \in T_\alpha,$$

where $\beta > \gamma$ if $x_{\alpha\beta} \in X_+$ and $\beta \geqslant \gamma$ if $x_{\alpha\beta} \in X_-$.

It is easy to see that the quotent-algebra $L(X)/J$ is isomorphic to the algebra $\prod^*_{H_0, \alpha \in T} H_\alpha$.

Now we order X in the following way: for $x_{\alpha\beta} \neq x_{\gamma\delta}$ we set $x_{\alpha\beta} > x_{\gamma\delta}$ if $\alpha > \gamma$ or $\alpha = \gamma$ and $\beta > \delta$.

We say that a regular monomial $w \in \Gamma(X)$ is *special* if the word $\overline{w}$ does not contain subwords $x_{\alpha\beta}x_{\alpha\gamma}$, $\beta > \gamma$, and also subwords $x_{\alpha\beta}x_{\alpha\gamma}$, where $\beta \geqslant \gamma$ for $x_{\alpha\beta} \in X_-$. We say that a monomial $w \in \Gamma(X)$ is called *s-special* if either w is a special monomial or $w = [v, v]$ where v is a special monomial, $d(v) \in G_-$, $l_X(v) > 1$.

THEOREM 2.3 [Mik2]. *Let K be a field,* char $K \neq 2, 3$. *Then the cosets whose representatives are s-special monomials form a linear basis of the algebra $L(X)/J = \prod^*_{H^0, \alpha \in T} H_\alpha$.*

If, in the definition of the free product of colour Lie superalgebras with amalgamated subalgebra, in place of colour Lie superalgebras we consider colour Lie p-superalgebras then we have the definition of *the free product of colour Lie p-superalgebras with amalgamated homogeneous subalgebra.* It is clear that such product is defined uniquely (up to an isomorphism of colour Lie p-superalgebras). Let J be the ideal of the free colour Lie p-superalgebra $L^p(X)$ generated by elements $d_{\alpha\beta\gamma}$ and $h_{\alpha\beta} = x^p_{\alpha\beta} - \sum_\tau q^\tau_{\alpha\beta} x_{\alpha\tau}$ where

$\alpha \in T$, $\beta, \tau \in T_\alpha$, $q^\tau_{\alpha\beta} \in K$, $x_{\alpha\beta} \in X_+$ and $e^p_{\alpha\beta} = \sum_\tau q^\tau_{\alpha\beta} e_{\alpha\tau}$ in H_α. It is clear that the algebra $L^p(X)/J$ is the free product of colour Lie p-superalgebras H_α with amalgamated homogeneous subalgebra H^0.

We say that a regular monomial $w \in \Gamma(X)$ is *p-special* if the word $\overline{w}$ does not contain subwords of the form $x_{\alpha\beta}x_{\alpha\gamma}$, $\beta > \gamma$; $x_{\alpha\beta}x_{\alpha\gamma}$, $\beta \geqslant \gamma$ with $x_{\alpha\beta} \in X_-$; or $x^p_{\alpha\beta}$ with $x_{\alpha\beta} \in X_+$; or $[x_{\alpha\beta}, x_{\alpha\beta}]^p$ with $x_{\alpha\beta} \in X_-$. A monomial $w \in \Gamma(X)$ is called *ps-special* if either w is a p-special monomial or $w = [u,u]^{p^t}$, $t \geqslant 0$, where u is a p-special monomial, $d(u) \in G_-$, $l_X(u) > 1$, or $w = v^{p^t}$ where $t \in \mathbf{N}$, v is a p-special monomial, $d(v) \in G_+$, $l_X(v) > 1$.

THEOREM 2.4 [BMPZ]. *Let K be a field,* $\operatorname{char} K = p > 3$. *Then the cosets, whose representatives are ps-special monomials, form a linear basis of $L^p(X)/J$.*

2.4. *The composition lemma*

Consider a weight function $\varphi : X \to \mathbf{N}$, let $\varphi(x_1 \ldots x_n) = \sum_{i=1}^n \varphi(x_i)$ for $x_1, \ldots, x_n \in X$ and let $\widehat{a}$ be the leading term in $a \in A(X)$ first relative to the weight φ and then lexicographically.

LEMMA 2.1. *Let u, v be regular words.*

(1) *If $u = avb$ $(a, b \in S(X))$ then on u there is an arrangement of brackets $[u] = [a[v]b]$ such that $[v]$ is a regular monomial, $\widehat{[u]} = u$ in $L(X)$.*

(2) *Let $u = avvb$, $d(v) \in G_-$, $a, b \in S(X)$. Then on u there is an arrangement of brackets $[u]$ such that $[u] = [a[v,v]b]$ where $[v,v]$ is an s-regular monomial, $\widehat{[u]} = 2u$ in $L(X)$.*

THEOREM 2.5 [Mik5]. *Any Lie superalgebra (Lie p-superalgebra) over a field with countable set of generators can be isomorphically embedded in a Lie superalgebra (in a Lie p-superalgebra) with two generators over the same field.*

The original Lie algebra version of this theorem is due to A.I. Shirshov [Shi2].

DEFINITION. Let a, b be two G-homogeneous elements in $L(X)$. Assume that their leading terms a_0 and b_0 in the s-regular presentation (first relative to the weight φ and then lexicographically) have coefficients which are equal to 1 and $\overline{a_0} = e_1e_2$, $\overline{b_0} = e_2e_3$ where $e_1, e_2, e_3 \in S(X)$. Then the word $e_1e_2e_3$ is regular. Now we define the element $(a,b) \in L(X)$ which is called *the composition of the elements $a, b \in L(X)$ (relative to a subword e_2)*:

(1) in the case where $a_0 \neq b_0$ we set $(a,b) = [(a)e_3] - \alpha[e_1(b)]$ where $[(a)e_3]$ is obtained from the monomial $[[a_0]e_3]$ with the arrangement of brackets described in Lemma 2.1 by replacing a_0 by a and analogously $[e_1(b)]$ is obtained from $[e_1[b_0]]$ by replacing b_0 by b; and α is such that $\widehat{[(a)e_3]} = \alpha\widehat{[e_1(b)]}$;

(2) for $a = b$, $\overline{a_0} = uu$, where u is a regular word and $d(u) \in G_-$ we set $(a,a) = [a,[u]] = [a - [u,u],[u]]$ where $[u]$ is a regular monomial.

A G-homogeneous subset $S \in L(X)$ is called *stable* if the following conditions are satisfied:

(a) the coefficients of the leading monomials in the s-regular presentation for all elements from S are equal to 1;

(b) all words $\widehat{s}$, where $s \in S$, are distinct and do not contain the other ones as subwords;

(c) if $a, b \in S$ and the composition (a, b) does exist then either $(a, b) \in S$ (up to a coefficient from K) or after expressing the elements associatively we have $(a, b) = \sum \alpha_i A_i s_i B_i$, where $\alpha_i \in K$, $s_i \in S$, $A_i, B_i \in S(X)$, $\alpha_1 A_1 \widehat{s_1} B_1 = \widehat{(a, b)}$, $A_i \widehat{s_i} B_i < A_1 \widehat{s_1} B_1$ $(i \neq 1)$ for all i.

THEOREM 2.6 (The Composition Lemma, [Mik4]). *Let K be a field,* $\operatorname{char} K \neq 2, 3$, *$L(X)$ a free colour Lie superalgebra, S a stable set,* $\operatorname{id}(S)$ *the ideal in $L(X)$ generated by S. Then for any element $a \in \operatorname{id}(S)$ the word $\widehat{a}$ contains a subword $\widehat{b}$, $b \in S$. Furthermore, s-regular monomials u such that $\widehat{u}$ does not contain subwords $\widehat{b}$, $b \in S$, form a linear basis of $L(X)/\operatorname{id}(S)$.*

An analogous theorem holds for colour Lie p-superalgebras (see [BMPZ]).

As corollaries one can derive a number of algorithms to solve algorithmic problems in Lie superalgebras, for example, an algorithm for solving equality in colour Lie p-superalgebras with stable or φ-homogeneous finite sets of defining relations (see [BMPZ, MZ3,MZ5]). Examples of symbolic computation in Lie superalgebras one may find in [GK1,GK2,MZ3,MZ5]. Also from the Composition Lemma one can deduce the theorems from Section 2.3.

For any ideal I of $L(X)$ (of $L^p(X)$) there exists a stable system of relations. Such a system of generators of I can be called a complete system of relations. Furthermore, there exists such complete system S of relations of I that for any $(p)s$-regular monomial w in the $(p)s$-regular presentation of $a \in S$ the word $\overline{w}$ does not contain subwords $\widehat{b}$, $b \in S \setminus a$. Such complete system of relations of I can be called *the Gröbner–Shirshov basis* of the ideal I of $L(X)$ (of $L^p(X)$).

EXAMPLE. Consider the case of $L(X)$. Let $a \in L(X)$, $\widehat{a} = 2uu$, where u is a regular word, $d(u) \in G_-$, $a = [u, u] + \sum \alpha_i v_i$ where $\alpha_i \in K$, v_i is an s-regular monomial, $\overline{v_i} < uu$, $\overline{v_1} > \overline{v_i}$ for $i \neq 1$ (first by weight φ and then lexicographically), $\overline{v_1} \neq bu$ for $\overline{v_1} > u$, $\overline{v_1} \neq ub$ for $u > \overline{v_1}$, $b \in S(X)$. Then $\operatorname{id}(a) = \operatorname{id}(S)$, $S = \{a, (a, a)\}$ where (a, a) is the composition, S is a stable set. Therefore, for $\operatorname{id}(a)$, we have an algorithmic solution of the equality problem.

More information on Gröbner–Shirshov bases of ideals of Lie superalgebras one may find in [BMPZ,BKLM,Mik8,MZ3,MZ5].

3. Subalgebras of free colour Lie $(p\text{-})$superalgebras

The theorem about freeness of subalgebras in free Lie algebras was proved by A.I. Shirshov [Shi1] and in free Lie rings and p-algebras by E. Witt [Wi2]. The theorem on the finiteness of the rank for the intersection of two subalgebras of finite rank is due to G.P. Kukin [Ku].

We say that a subset M of G-homogeneous elements of $L(X)$ is *independent* if M is a set of free generators of the subalgebra generated by M in $L(X)$.

A subset $M=\{a_i\}$ of G-homogeneous elements in $L(X)$ is called *reduced* if for any i the leading part a_i^0 of the element a_i does not belong to the subalgebra generated by the set $\{a_j^0,\ j\neq i\}$.

Let $S=\{s_\alpha,\ \alpha\in I\}$ be a G-homogeneous subset of $L(X)$. We say that a mapping $\omega: S\to L(X)$ is *an elementary transformation* if $\omega(s_\alpha)=s_\alpha$ for all $\alpha\in I\setminus\beta$, $\omega(s_\beta)=\lambda s_\beta+\omega(s_{\alpha_1},\dots,s_{\alpha_t})$, where

$$0\neq\lambda\in K,\ \alpha_1,\dots,\alpha_t\neq\beta,\ d\big(\omega(s_{\alpha_1},\dots,s_{\alpha_t})\big)=d(s_\beta),$$
$$\omega(y_{\alpha_1},\dots,y_{\alpha_t})\in L(y_1,y_2,\dots),\ d(y_i)=d(s_i).$$

Such elementary transformation is called *triangular* if $l_X(\omega(s_{\alpha_1},\dots,s_{\alpha_t}))\leqslant l_X(s_\beta)$. It is clear that elementary transformations of the set X of free generators induce automorphisms of the free colour Lie superalgebra $L(X)$.

THEOREM 3.1 [Mik1]. *Let K be a field,* $\operatorname{char} K\neq 2,3$. *Then any reduced subset of a free colour Lie superalgebra $L(X)$ is independent, and, therefore, any G-homogeneous subalgebra of $L(X)$ is free.*

The case when K is a ring was considered in [Mik6].

Now we pass to the corollaries of the theorem about the freeness of subalgebras.

THEOREM 3.2 (An analogue of P.M.Cohn's theorem on generators of the automorphism group of a free Lie algebra of finite rank, [BMPZ]). *The set of elementary automorphisms generates the automorphism group of a free colour Lie superalgebra of finite rank.*

From these theorems we derive an algorithm to decide whether some finite set of G-homogeneous slements of a free colour Lie superalgebra is an independent set (it is necessary to transform these elements using triangular elementary transformations to a reduced set) and an algorithm to solve the entry problem for finite generated subalgebras of free colour Lie superalgebra, see [BMPZ].

GENERATORS OF THE DERIVED ALGEBRA. *Let X be a G-graded set, $|X|\geqslant 2$. Suppose that X is totally ordered in such way that from $x\in X_+$, $y\in X_-$ it follows that $y>x$. The derived algebra $L^2(X)=[L(X),L(X)]$ is a free colour Lie superalgebra, and, as a free set of generators, we can choose the left-normed monomials $[x_1x_2\dots x_{s-1}x_s]=[[x_1x_2\dots x_{s-1}],x_s]$ where $x_i\in X$, $s\geqslant 2$, $x_1>x_2\leqslant x_3\leqslant\cdots\leqslant x_s$ and $x_i\neq x_{i+1}$ for $x_i\in X_-$, and also if $x_1,\dots,x_s\in X_-$ then $x_1\geqslant x_2<x_3<\cdots<x_s$.*

Now we consider the cardinality of ranks of some subalgebras of $L(X)$.

THEOREM 3.3 (An analogue of O. Schreier's Formula, [Mik1]). *Let K be a field,* $\operatorname{char} K \neq 2, 3$, *$L = L(X) = L_+ \oplus L_-$ the free colour Lie superalgebra,* $\operatorname{rank} L = |X| = N < \infty$, *and let M be a G-homogeneous subalgebra, $M = M_+ \oplus M_-$, $M_+ = L_+$,* $\dim L_-/M_- = s < \infty$. *Then* $\operatorname{rank} M = 2^s(N-1) + 1$.

Using formal power series V.M. Petrogradsky obtained in [Pe4] more general Schreier type formulas.

COROLLARY 3.0.1. *Let K be a field,* $\operatorname{char} K \neq 2, 3$, *$L = L_+ \oplus L_-$ the finitely generated colour Lie superalgebra, and let $M = M_+ \oplus M_-$ be a G-homogeneous subalgebra of L such that $M_+ = L_+$ and* $\dim L_-/M_- < \infty$. *Then M is a finitely generated colour Lie superalgebra.*

THEOREM 3.4 [BMPZ]. *Let $|X_-| \geqslant 1$, $L(X) = L_+ \oplus L_-$ be a free colour Lie superalgebra, $M = M_+ \oplus M_-$ a G-homogeneous subalgebra in $L(X)$ such that $M_- = L_-$ and* $\dim L_+/M_+ = s < \infty$. *Then* $\operatorname{rank} M = \infty$, *$M = L(Y)$ and $|Y_-| = \infty$.*

PROPOSITION 3.1. *Any nonzero ideal $H = H_+ \oplus H_-$ of the free colour Lie superalgebra $L(X) = L_+ \oplus L_-$ such that $H \neq 0$, $|X| > 1$ and $H_+ \neq L_+$, is a free colour Lie superalgebra of infinite rank.*

The following theorem shows that any subalgebra of finite rank of $L(X)$ is completely defined by its (nontrivial) even component.

THEOREM 3.5 [BMPZ]. *Let K be a field,* $\operatorname{char} K \neq 2, 3$, *$|X_-| \geqslant 1$, and let A, B be G-homogeneous subalgebras of finite rank in the free colour Lie superalgebra $L(X)$, $A = A_+ \oplus A_-$, $B = B_+ \oplus B_-$. If $A_- = B_- \neq \{0\}$, then $A = B$.*

Now we give some examples concerning this theorem.

(1) Let $X_+ = \{x\}$, $X_- = \{y\}$, $A = L(X)$, $B = L(Z(x))$ where $Z(x) = \{y, ad^k x(y) \mid k \in \mathbf{N}\}$. Then since $x \in A_+$ and $x \notin B_+$, we see that $A_+ \neq B_+$; but at the same time $A_- = B_-$ (here $\operatorname{rank} B = \infty$).

(2) In general, from $A_+ = B_+$ it does not follow that $A = B$. Indeed, let $X_+ = \{x\}$, $X_- = \{y\}$, $A = L(X)$, $B = L(W(y))$, where $W(y) = \{x, [x, y], [y, y]\}$. Then $\operatorname{rank} A < \infty$, $\operatorname{rank} B < \infty$ and $A_+ = B_+$. Since $y \in A_-$ and $y \notin B_-$, we have at the same time $A_- \neq B_-$.

Now we consider the action of a finite group of linear automorphisms on the free colour Lie superalgebra $L(X)$. In the Lie algebra case this problem was considered in [Br] and [Dr]. Let V_g be the free K-module spanned by the set X_g, $\mathrm{GL}(V_g)$ the group of all automorphisms of the K-module V_g, $F = \prod_{g \in G} \mathrm{GL}(V_g)$, H a finite subgroup in F. Then $L(X)$ is a KH-module, moreover, any element of H acts as an automorphism of the algebra $L(X)$. Denote by L^H the subalgebra of all fixed points in $L(X)$ relatively the action of the group H.

THEOREM 3.6 [Mik7]. *Let K be a principal ideal domain,* $\operatorname{char} K \neq 2$, $2 \leqslant |X| < \infty$ *and the group H be finite and nontrivial. Then L^H is a free colour Lie superalgebra of infinite rank.*

As a corollary we have that in the case where $|X| > 1$, $|X_-| \geqslant 1$, the even component L_+ of $L(X) = L_+ \oplus L_-$ has infinite rank (see also [BMPZ]).

Now we pass to subalgebras of the free colour Lie p-superalgebra $L^p(X)$. As in the case of $L(X)$, we have

THEOREM 3.7 [Mik3]. *Let K be a field, $p = \operatorname{char} K > 3$, $L^p(X)$ the free colour Lie p-superalgebra. Then any reduced subset of $L^p(X)$ is independent, and therefore, any G-homogeneous subalgebra H of $L^p(X)$ is a free colour Lie p-superalgebra.*

From this theorem one can deduce a number of properties of subalgebras of $L^p(X)$ (see [BMPZ]). For example, we have the following

THEOREM 3.8 [Mik3]. *Let K be a field, $p = \operatorname{char} K > 3$, $|X| = N < \infty$. Suppose that $H = H_+ \oplus H_-$ is a G-homogeneous subalgebra of the free colour Lie p-superalgebra $L^p(X)$, $\dim L_+^p / H_+ = t < \infty$ and $\dim L_-^p / H_- = s < \infty$. Then $\operatorname{rank} H = 2^s p^t (N-1) + 1$.*

Now we consider the intersection of two finitely generated subalgebras of $L(X)$ (of $L^p(X)$) over a field K with $\operatorname{char} K > 3$.

THEOREM 3.9 [Mik3]. *Let K be a field, $p = \operatorname{char} K > 3$, $L^p(X)$ the free colour Lie p-superalgebra, and let A and B be two G-homogeneous finitely generated subalgebras of $L^p(X)$. Then the subalgebra $A \cap B$ is also finitely generated.*

We conclude this section by formulating a theorem on the finite separability of free colour Lie superalgebras. It should be recalled that an algebra is called *finitely separable* if for any element not in a finitely generated subalgebra there exists a homomorphism onto a finite-dimensional algebra where this property; i.e. the image of the element is not in the image of the subalgebra.

THEOREM 3.10 [BMPZ]. *If a be a G-homogeneous element of L, B a G-homogeneous finitely generated subalgebra of L, $a \notin B$, then there exist a finite dimensional colour Lie $(p\text{-})$superalgebra H and an epimorphism $\varphi : L \to H$ such that $\varphi(a) \notin \varphi(B)$.*

In the case of ordinary Lie algebras this is a theorem of U. Umirbaev [Um].

Some other properties concerning the algebras $L(X)$ and $L^p(X)$ and their subalgebras can be found in [MY,MZ1,MZ2,MZ3,MZ4]: associative support of $L(X)$ (of $L^p(X)$); annihilators of elements of $L(X)$; images of inner derivations; Jacobi criterion for an epimorphism of $L(X)$ to be an automorphism; matrix criteria for system of elements to be primitive and to have given rank; test elements and automorphic orbits.

4. Universal enveloping algebras

4.1. *The Jacobson radical*

A description of the Jacobson radical for universal enveloping algebra of Lie superalgebra is rather a difficult problem. Some assumptions, for example, $L_- = 0$, yield the equality $J(U(L)) = 0$. We will record certain partial results.

THEOREM 4.1. *Let L be a colour Lie superalgebra over an arbitrary field. Suppose that $M \subset L_-$ is a homogeneous L_+-submodule such that $[M, L_-] = 0$. Then $U(L) \cdot M \subset J(U(L))$. Moreover, $U(L)M$ is a locally nilpotent algebra.*

In particular, one can easily conctruct a superalgebra L with nonzero Jacobson radical in $U(L)$.

COROLLARY 4.1.1. *Let L be a colour Lie superalgebra over an arbitrary field. Suppose that $[L_-, L_-] = 0$. Then $J(U(L)) = U(L)L_-$ and it is a locally nilpotent algebra.*

On the other hand, having L_- nonzero does not need to imply $J(U(L)) \neq 0$. For example, if $L = L_0 \oplus L_1 = \Gamma_n(1, 1)$ is the Heisenberg Lie superalgebra with $L_0 = \langle z \rangle$, $L_1 = \langle x_1, \dots, x_n, y_1, \dots, y_n \rangle$, $[x_i, y_j] = \delta_{ij} z$, then $J(U(L)) = 0$.

Another important condition for $U(L)$ is the primeness. For a simple finite-dimensional Lie superalgebras the following is known.

THEOREM 4.2 [Wil]. *Let L be a finite-dimensional simple Lie superalgebra over an algebraically closed field of characteristic zero. Then L satisfies Bell's criterion (so that $U(L)$ is prime) if and only if L is not of one of: $b(n)$ for $n \geqslant 3$; $W(n)$ for odd $n \geqslant 5$; $S(n)$ for odd $n \geqslant 3$.*

It is known that $U(b(n))$ and $U(S(n))$ are not semiprime.

4.2. *Hopf algebra structure on $U(L)$*

Consider a colour Lie superalgebra L over a field K, $\operatorname{char} K \neq 2, 3$, and its universal enveloping algebra $U(L)$. We define the diagonal mapping δ from L to the colour tensor product $U(L) \otimes_\varepsilon U(L)$ by setting $\delta(a) = a \otimes 1 + 1 \otimes a$ for $a \in L$. If we put $\delta(1) = 1 \otimes 1$ then this mapping has a unique extension to a homomorphism $\delta : U(L) \to U(L) \otimes_\varepsilon U(L)$ called the co-product on the universal enveloping algebra $U(L)$. The co-product is co-associative, i.e. $\varphi_1 = \varphi_2$ where

$$\varphi_1 = (\mathrm{id}_U \otimes \delta) \circ \delta, \qquad \varphi_2 = (\delta \otimes \mathrm{id}_U) \circ \delta,$$
$$\varphi_1, \varphi_2 : U(L) \to U(L) \otimes_\varepsilon U(L) \otimes_\varepsilon U(L).$$

Consider the unique homomorphism $e : U(L) \to K$ such that $e(L) = 0$ and $e(1) = 1$. Then e is the co-unit of $U(L)$, i.e. $(e \otimes \mathrm{id}) \circ \delta = \mathrm{id} = (\mathrm{id} \otimes e) \circ \delta$ (we identify canonically

$K \otimes_\varepsilon U(L)$ and $U(L) \otimes_\varepsilon K$ with $U(L)$). If τ is the automorphism of $U(L) \otimes_\varepsilon U(L)$ defined by $\tau(x \otimes y) = \varepsilon(x, y) y \otimes x$ with homogeneous $x, y \in U(L)$, then $\tau \circ \delta = \delta$. Hence, $U(L)$ is a co-commutative bialgebra.

We define the mapping $\bar{e} : U(L) \to U(L)$ by setting $\bar{e}(x) = e(x) \cdot 1_U$ for all $x \in U(L)$. Let η be the linear mapping on $U(L)$ such that $\eta(1) = 1, \eta(x) = -x$ for all $x \in L$ and $\eta(xy) = \varepsilon(x, y)\eta(y)\eta(x)$ for all homogeneous $x, y \in U(L)$. Then $\eta^2 = \mathrm{id}_U$ and $\mu \circ (\eta \otimes \mathrm{id}_U) \circ \delta = \mu \circ (\mathrm{id}_U \otimes \eta)\circ = \bar{e}$ where the linear mapping $\mu : U(L) \otimes_\varepsilon U(L) \to U(L)$ is induced by the multiplication on $U(L)$. So, η is an antipode, and our colour bialgebra $U(L)$ with its antipode η is a co-commutative colour Hopf algebra.

For a colour Hopf algebra we define the set of all its primitive elements by $\mathcal{P}(H) = \{a \in H \mid \delta(a) = a \otimes 1 + 1 \otimes a\}$.

THEOREM 4.3. *Let K be a field of characteristic zero. Then the set of primitive elements in the universal enveloping algebra $U(L)$ of the colour Lie superalgebra L coincides with L.*

Using the same method one can define the structure of a co-commutative colour Hopf algebra on the restricted universal enveloping algebra provided that $\operatorname{char} K = p > 0$ and L is a colour Lie p-superalgebra.

THEOREM 4.4. *Let K be a field,* $\operatorname{char} K = p > 3$, *$L$ a colour Lie superalgebra, M a colour Lie p-superalgebra over K. Then*

(1) $\mathcal{P}(u(M)) = M$;

(2) *$\mathcal{P}(U(L))$ coincides with the colour Lie p-superalgebra generated by L in $U(L)$.*

4.3. *Regularity and self-injectivity of $u(L)$*

For a ring R with unity the following properties are equivalent:

- For any $a \in R$ there exists $x \in R$ with $axa = a$.
- Each left principal ideal I of R is generated by an idempotent (i.e. I is a direct summand of R as a left R-module).
- Each finitely generated left ideal is generated by idempotent.

Such rings are called *(von Neumann) regular.*

An associative algebra A is called a (left) *self-injective* if it is injective as the (left) module over itself.

The restricted universal enveloping algebra $u(L)$ plays the same role with respect to L as the group ring $K[H]$ for a group H. It is known that $K[H]$ is regular if and only if H is a locally finite group without elements of order p, where $p = \operatorname{char} K > 0$ [Vi]. For colour Lie superalgebras we have the following result.

THEOREM 4.5. *Let G be a finite group and L a colour Lie p-superalgebra. The following conditions are equivalent:*

(1) *The ring $u(L)$ is regular.*

(2) *The trivial one-dimensional $u(L)$-module K is injective.*

(3) *All simple $u(L)$-modules are injective.*
(4) $L = L_+$ *is Abelian, locally finite-dimensional, and possesses a non-degenerate p-map in the following sense:* $H = \langle x^{[p]} \mid x \in H_g,\ g \in G_+ \rangle$ *for any homogeneous restricted subalgebra* $H \subset L$.

For group rings it is known that $K[H]$ is self-injective if and only if H is finite [Re].

THEOREM 4.6. *Let L be a colour Lie p-superalgebra. Then $u(L)$ is self-injective if and only if* $\dim L < \infty$.

4.4. *Identities in enveloping algebras*

Let $L = \bigoplus_{g \in G} L_g$ be a colour Lie superalgebra over K. We say that L is a *special* superalgebra if it has an associative enveloping algebra A with nontrivial polynomial identity. For example, if $\dim L < \infty$ then it is special by the Ado–Iwasawa Theorem. We will give necessary and sufficient conditions for a colour Lie superalgebra over a field under which its restricted or universal enveloping algebra satisfies a non-trivial associative identity.

THEOREM 4.7. *Let G be a finite group and $L = \bigoplus L_g$ a restricted colour Lie superalgebra over a field of positive characteristic $p > 2$. Then the restricted enveloping algebra $u(L)$ satisfies a non-trivial identity if and only if there exist restricted homogeneous ideals $Q \subset R \subset L$ such that*
(1) $\dim L/R < \infty$, $\dim Q < \infty$;
(2) $R^2 \subset Q$, $Q^2 = 0$;
(3) *Q has a nilpotent p-map (i.e. for each even homogeneous $x \in G_g$, $g \in G_+$, we have $x^{[p^n]} = 0$ for some integer n).*

THEOREM 4.8. *Let G be a finite group and $L = \bigoplus L_g$ a colour Lie superalgebra over a field of positive characteristic $p > 2$. Then its universal enveloping algebra $U(L)$ satisfies a non-trivial identity if and only if there exist homogeneous ideals $B \subset A \subset L$ such that*
(1) $\dim L/A < \infty$, $\dim B < \infty$;
(2) $A^2 \subset B$;
(3) $B = B_-$ *(i.e.* $B_+ = 0$);
(4) *All inner derivations* $\operatorname{ad} x$, *$x \in L_g$, $g \in G_+$, defined on the whole of the superalgebra, are algebraic and their degrees are bounded by some constant.*

THEOREM 4.9. *Let G be a finite group and $L = \bigoplus L_g$ a colour Lie superalgebra over a field of zero characteristic. Then the universal enveloping algebra $U(L)$ satisfies a non-trivial identity if and only if there exists a homogeneous L_+-submodule $M \subset L_-$ such that*
(1) $\dim L_-/M < \infty$, $\dim[L_+, M] < \infty$;
(2) *L_+ is Abelian and* $[M, M] = 0$.

Necessary and sufficient conditions for $U(L)$ being PI-algebra for a Lie algebra L were obtained by V.N. Latyshev [La] and Yu.A. Bahturin [Bah6]. For restricted enveloping algebras the same problem was solved by V.M. Petrogradsky [Pe1] and D.S. Passman [Pa1].

Finiteness of the group G in the previous theorems is essential since there exists examples of Abelian colour Lie superalgebras with infinite group G whose enveloping algebras do not satisfy any identity. Suppose that $\lambda \in K$ is not a root of unity. Set $G = \mathbf{Z} \times \mathbf{Z}$ and $\alpha = (1,0)$, $\beta = (0,1)$. Then the condition $\varepsilon(\alpha,\beta) = \lambda$, $\varepsilon(\alpha,\alpha) = \varepsilon(\beta,\beta) = 1$ uniquely determines a bilinear alternating form $\varepsilon: G \times G \to K^*$. Consider the two-dimensional Abelian colour Lie algebra $L = \langle x \rangle \oplus \langle y \rangle$ with $d(x) = \alpha \in G, d(y) = \beta \in G$. Then $U(L)$ does not satisfy any identity. If $\operatorname{char} K = p > 0$ then the subspace

$$\widetilde{L} = \bigoplus_{i=0}^{\infty} \left(\langle x^{p^i} \rangle \oplus \langle y^{p^i} \rangle\right) \subset U(L)$$

is naturally equipped with a structure of an Abelian colour Lie p-algebra such that $u(\widetilde{L}) \cong U(L)$. Hence, $u(\widetilde{L})$ is not a PI-algebra.

5. Varieties of colour Lie superalgebras

5.1. *Graded identities and varieties*

Let $X = \bigcup_{g \in G} X_g$ be a G-graded set, $L(X)$ the free colour Lie superalgebra with free generating set X over an associative and commutative ring K with 1. If $f = f(x_1, \ldots, x_n)$ is an element in $L(X)$ then an expression of the form $f(x_1, \ldots, x_n) = 0$ is called *an identical relation* (or an *identity*). Suppose also that $R = \bigoplus_{g \in G} R_g$ is a colour Lie superalgebra over K. We say that $f = 0$ is satisfied in R if $f(a_1, \ldots, a_n) = 0$ for any choice of $a_1, \ldots, a_n \in R$ such that all a_i are G-homogeneous and $d(a_i) = d(x_i)$, $i = 1, \ldots, n$. Since $L(X)$ can be also graded by the degrees in X or some fixed $x \in X$, we can talk about homogeneous and multihomogeneous identities in colour Lie superalgebras. Finally, we say that an identity is *multilinear* if it is multihomogeneous and its degree with respect to any variable occuring is equal to 1. One can use 'non-graded' multilinear identities, namely, a colour Lie superalgebra $R = \bigoplus_{g \in G} R_g$ satisfies multilinear (non-graded) identity $f(y_1, \ldots, y_n) = 0$ if $f(a_1, \ldots, a_n) = 0$ for any $a_1, \ldots, a_n \in R$. Obviously, a non-graded multilinear identity is equivalent to some family of graded multilinear identities. It enables us to define Abelian, nilpotent and solvable colour Lie superalgebras by means of identities, as usual.

The class of all G-graded K-algebras satisfying a fixed system of identical relations is called a *variety*. The least variety containing a given colour Lie superalgebra $R = \bigoplus_{g \in G} R_g$ will be denoted by $\operatorname{var} R$. It is easy to see that Birkhoff's Theorem remains valid for varieties of graded algebras. As usual, the set of all elements $f \in L(X)$ such that $f = 0$ is an identity of a fixed variety $\mathcal{V}$ forms a homogeneous ideal V of the algebra $L(X)$. The quotient-algebra $L(X, \mathcal{V}) = L(X)/V$ is the (relatively) free algebra of the variety $\mathcal{V}$.

In [BZ2] the authors prove a theorem about the existence of a nontrivial identity in a superalgebra graded by a finite group provided such an identity exists in the neutral component of the grading (this is new even in the case of ordinary Lie algebras). It follows also that such an identity exists in a color Lie superalgebra over a finite group if a nontrivial Lie identity exists in the neutral component.

5.2. *Hilbert series and growth functions*

Now let K be a field, $L = L(X, \mathcal{V})$ the free algebra of a variety $\mathcal{V}$ generated by a finite set $X = X_+ \cup X_-$, $X_+ = \{x_1, \dots, x_m\}$, $X_- = \{y_1, \dots, y_n\}$. Given two rows of integers $\gamma = (\gamma_1, \dots, \gamma_m)$, $\delta = (\delta_1, \dots, \delta_n)$ denote by $L_{\gamma,\delta}$ the following multihomogeneous component of L:

$$L_{\gamma,\delta} = \{w \in L \mid \deg_{x_i}(w) = \gamma_i,\ \deg_{y_j}(w) = \delta_j\}.$$

(We assume that the ideal of identities of $\mathcal{V}$ is multihomogeneous. This holds, for example, if K is infinite or if $\mathcal{V}$ is a variety defined by a system of multilinear identities. Then $\deg_{x_i}(w)$ is a well-defined notation.)

The formal series:

$$H(L) = H(t_1, \dots, t_m, u_1, \dots, u_n) = \sum (\dim L_{\gamma,\delta}) t_1^{\gamma_1}, \dots, t_m^{\gamma_m}, u_1^{\delta_1}, \dots, u_n^{\delta_n}$$

is called the *Hilbert series* of L; this is an important invariant of the variety $\mathcal{V}$. Similarly, the Hilbert series $H(L, t)$ in one variable t is

$$H(L, t) = \sum (\dim L_s) t^s, \quad \text{where } L_s = \sum_{|\gamma| + |\delta| = s} L_{\gamma,\delta}.$$

For the simplest varieties the Hilbert series can be computed explicitly. This is possible, for example, for the metabelian variety which is defined by the (non-graded) identity $[[x, y], [z, t]] = 0$.

THEOREM 5.1. *Let $X_+ = \{x_1, \dots, x_m\}$, $X_- = \{y_1, \dots, y_n\}$, $M(X)$ the free metabelian algebra on X. Then*

$$\begin{aligned} H(M(X)) &= 1 + (t_1 + \cdots + t_m + u_1 + \cdots + u_n) \\ &\quad + \frac{(t_1 + \cdots + t_m + u_1 + \cdots + u_n - 1)(1 + u_1) \cdots (1 + u_n)}{(1 - t_1) \cdots (1 - t_m)}, \\ H(M(X), t) &= 1 + (m + n)t + ((m + n)t - 1) \frac{(1 + t)^n}{(1 - t)^m}. \end{aligned}$$

Now denote by F a free non-associative algebra generated by a countable set $X = \{x_1, x_2, \dots\}$ over a field K, $\operatorname{char} K = 0$, and by $\mathcal{V}$ a variety of K-algebras. Then the family of identities of $\mathcal{V}$ forms an ideal I in F. Let P_n be the subspace of all multilinear polynomials in $x_1, \ldots, x_n$ in F. We set

$$c_n(\mathcal{V}) = \dim \frac{P_n}{P_n \cap I}.$$

The sequence $(c_n(\mathcal{V}))$, $n = 1, 2, \ldots$, is an important numerical characteristic of variety $\mathcal{V}$. The growth of $\mathcal{V}$ is said to be polynomial if $c_n(\mathcal{V}) < n^k$ for some k and it is exponential if $a^n < c_n(\mathcal{V}) < b^n$ for some $a, b > 1$.

If $\mathcal{V}$ is a variety of Lie superalgebras over K then another numerical characteristic can be considered. Namely, denote by $L_0 \oplus L_1$ a free Lie superalgebra over K with the same set X of even generators and the set $Y = \{y_1, y_2, \ldots\}$ of odd generators. Let R be a non-homogeneous subalgebra in L generated by $z_i = x_i + y_i$, $n = 1, 2, \ldots$, and let $\overline{P}_n$ denote a subspace of all multilinear polynomials on $z_1, \ldots, z_n$ in R. Any map $x_i \mapsto a_i \in H_0$, $y_i \mapsto b_i \in H_1$ with the image in some Lie superalgebra $H = H_0 \oplus H_1$ can be uniquelly extended to a Lie superalgebras homomorphsm $\varphi\colon R \to H$. If H is a free algebra of countable rank of a variety $\mathcal{V}$ then an intersection J of kernels of all such φ can also be considered as the ideal of identities of $\mathcal{V}$, and codimension of $\overline{P}_n \cap J$ in $\overline{P}_n$ can be taken instead of $c_n(\mathcal{V})$. However, these characteristics coincide since

$$\pi(P_n) = \overline{P}_n, \qquad \pi(I) = J, \qquad \pi^{-1}(J) = I$$

for the canonical epimorphism $\pi : F \to R$ where $\pi(x_i) = z + i$.

The theory of codimension growth is well-developed in cases of associative and Lie algebras. For Lie superalgebras there is a description of varieties of relatively slow growth. Recall that $\mathcal{N}_t\mathcal{A}$ means a variety of Lie superalgebras which satisfy a non-graded identity $[z_1, z_2], \ldots, [z_{2t+1}, z_{2t}] = 0$.

THEOREM 5.2 [ZM]. *A variety $\mathcal{V}$ of Lie superalgebras over a field of characteristic zero has a polynomial codimension growth if and only if the following conditions hold:*

(1) $\mathcal{V} \subset \mathcal{N}_c\mathcal{A}$;

(2) *there exist an integer k such that any multilinear polynomial containing at least k even and k odd variables is equal to zero identically in $\mathcal{V}$;*

(3) *for any $j = 0, 1, \ldots, c$ identity of the form*

$$\begin{aligned}
&[z_1, z_2, x_1, \ldots, x_N], t_1, \ldots, t_r, [z_3, z_4, x_{N+1}, \ldots, x_{2N}] \\
&\quad = \sum_\sigma a_\sigma^r [z_1, z_2, x_{\sigma(1)}, \ldots, x_{\sigma(N)}], t_1, \ldots, t_r, [z_3, z_4, x_{\sigma(N+1)}, \ldots, x_{\sigma(2N)}]
\end{aligned}$$

holds in $\mathcal{V}$ where the summation extends only over those permutations $\sigma in S_{2n}$ for which $\sigma(\{1, \ldots, N\}) \not\subset \{1, \ldots, N\}$.

A sequence of integers a_n is said to be of *intermediate growth* if a_n is asymptotically less than d^n for any $d > 1$ and asymptotically greater than n^t for any integer t. As in an usual Lie or associative case we have

THEOREM 5.3 [ZM]. *There are no varieties of intermediate growth in the class of Lie superalgebras.*

5.3. *Symmetric group action*

Let $L(X)$, $X = \bigcup_{g\in G} X_g$, be a free colour Lie superalgebra over a field K, $\operatorname{char} K = 0$, and each X_g is countable. For any function $\gamma : G \to \mathbf{N}$ which is zero almost everywhere denote by L_γ the multilinear component of $L(X)$ of length $\gamma(x)$ in the variables X_g with this variable fixed. The Cartesian product of the symmetric groups $S_\gamma = \prod_{g\in G} \mathrm{Sym}(\gamma(g))$ acts on L_γ by permuting variables of the same G-degree. If we denote by $M(d)$ the irreducible $\mathrm{Sym}(m)$-module associated with the Young diagram d then any irreducible S_γ-submodule in L_γ is isomorphic to $M(d_1)\otimes M(d_2)\otimes\cdots\otimes M(d_s)$ where $d_1, d_2, \ldots, d_s$ are partitions of the numbers $\gamma(g)$ which are different from zero. As an example of the uses of symmetric group techniques consider the free metabelian colour Lie superalgebra $M(X)$. The action of the group S_γ on L_γ induces a similar action on the set P_γ of multilinear elements in the free metabelian superalgebra $M(X)$. In order to describe P_γ as an S_γ-module we assume that the set of elements in G is totally ordered.

THEOREM 5.4. *Let γ have the form*

$$\gamma = \big(\gamma(g_1), \ldots, \gamma(g_r), 0, 0, \ldots, \gamma(h_1), \ldots, \gamma(h_s), 0, 0, \ldots\big),$$

where $g_1 < \cdots < g_r$, $h_1 < \cdots < h_s$, $g_i \in G_+$, $h_j \in G_-$, and $\gamma(g_i) \neq 0$, $\gamma(h_j) \neq 0$. Then P_γ is the direct sum of the folowing irreducible S_γ-modules.

(1) $M_1(f) = M(\gamma(f)-1, 1)\otimes(\bigotimes_{g\neq f} M(\gamma(g))\otimes(\bigotimes_h M(1^{\gamma(h)}))$, *where $f \in G_+$, $\gamma(f) > 1$, and the first factor in the tensor product corresponds to $f \in G$.*

(2) $M_2(f) = M(2, 1^{\gamma(f)-2})\otimes(\bigotimes_g M(\gamma(g))\otimes(\bigotimes_{h\neq f} M(1^{\gamma(h)}))$, *where $f \in G_-$, $\gamma(f) > 1$, and the first factor in the tensor product corresponds to $f \in G$.*

(3) $r + s - 1$ *isomorphic copies of the module*

$$M_3 = \Big(\bigotimes_g M\big(\gamma(g)\big)\Big)\otimes\Big(\bigotimes_h M\big(1^{\gamma(h)}\big)\Big).$$

5.4. *Subvarieties of the metabelian variety*

Let K be a field of zero characteristic, G a finite group.

THEOREM 5.5. *Any proper subvariety $\mathcal{V}$ of the metabelian variety of colour Lie superalgebras over K is multinilpotent, i.e. there exists $\gamma = (\gamma_1, \ldots, \gamma_r)$ such that $P_\delta(\mathcal{V}) = 0$ as soon as the multidegree $\delta = (\delta_1, \ldots, \delta_r)$ satisfies $\delta_i \geqslant \gamma_i$, $i = 1, \ldots, r$, where $r = |G|$.*

Here $P_\delta(\mathcal{V})$ is the multilinear component of multidegree δ in the free superalgebra $L(X, \mathcal{V})$.

From the previous theorem it follows that any subvariety $\mathcal{V}$ in the metabelian variety can be defined by a finite set of its identities $f_1 = 0, \ldots, f_n = 0$.

5.5. *Grassmann envelopes*

Let $L = \bigoplus_{g \in G} L_g$ be a colour Lie superalgebra defined by a form $\varepsilon : G \times G \to K^*$. Consider a G-graded space $V = \bigoplus_{g \in G} V_g$ as an Abelian ε^{-1}-colour Lie superalgebra. Let $P = U(V)$ be its universal enveloping algebra. Then the tensor product $L \otimes P$ is a nonassociative G-graded algebra. Its zero component $\mathcal{G}(L)$ is equal to direct sum of the subspaces $L_g \oplus P_{-g}$, $g \in G$. Easy calculations show that $\mathcal{G}(L)$ is a usual Lie algebra. If L is a Lie superalgebra over K, $G = \mathbf{Z}_2$, then $\varepsilon^{-1} = \varepsilon$ on G, $V = V_0 \oplus V_1$ and $P = K[V_0] \otimes \Lambda(V_1)$, i.e. P is the tensor product of the polynomial ring $P = K[V_0]$ and the Grassmann algebra $\Lambda(V_1)$. An important particular case is that of $V_0 = 0$, $V = V_1$, where $P = P_0 \oplus P_1$ is the Grassmann algebra of the vector space V_1. In this case $\mathcal{G}(L) = L_0 \otimes P_0 \oplus L_1 \otimes P_1$ is a $\mathbf{Z}_2$-graded Lie algebra. Moreover, if $L = L_0 \oplus L_1$ is some $\mathbf{Z}_2$-graded nonassociative algebra then $\mathcal{G}(L)$ will be a Lie algebra only if L is a Lie superalgebra. In all cases to follow we assume that for $G = \mathbf{Z}_2$, the space V is odd, i.e. $V = V_1$, and $\dim V = \infty$ if K is a field.

5.6. *Application of Lie superalgebras*

Recall that a variety $\mathcal{V}$ of Lie algebras over a field K is named a *special* one if $\mathcal{V} = \operatorname{var} L$ where L is a special Lie algebra, i.e. it has an associative enveloping algebra with polynomial identity. Let $\operatorname{char} K = 0$.

THEOREM 5.6 [Wa]. *For any special variety $\mathcal{V}$ of Lie algebras there exists a certain finitely generated special Lie superalgebra $L = L_0 \oplus L_1$ such that $\mathcal{V} = \operatorname{var} \mathcal{G}(L)$.*

The basic rank $r_b(\mathcal{V})$ of a variety $\mathcal{V}$ is the least integer t such that $\mathcal{V} = \operatorname{var} R$ and the algebra R is generated by t elements, if such a number exists. Otherwise $r_b(\mathcal{V}) = \infty$.

THEOREM 5.7 [Za6]. *Let $\mathcal{V}$ be a special variety of Lie algebras, $\mathcal{V} = \operatorname{var}(\mathcal{G}(L))$ where $L = L_0 \oplus L_1$ is a special finitely generated Lie superalgebra. Suppose that $\mathcal{W}$ is the variety of Lie superalgebras generated by L. Then $r_b(\mathcal{V}) < \infty$ if and only if L satisfies a graded identity of the form $(\operatorname{ad} x)^n = 0$, $x \in L_1$, and $\mathcal{W}$ does not contain any finite-dimensional simple Lie superalgebra $H = H_0 \oplus H_1$ with $H_1 \neq 0$.*

THEOREM 5.8 [Za6]. *The unique nonsolvable special variety of Lie algebras $\mathcal{V}$ such that $r_b(\mathcal{V}) = \infty$ and $r_b(\mathcal{U}) < \infty$ for any proper subvariety $\mathcal{U}$ in $\mathcal{V}$ is* $\operatorname{var}(\mathcal{G}(\operatorname{osp}(1,2)))$.

Here $\operatorname{osp}(1,2) = L_0 \oplus L_1$ is the five-dimensional simple Lie superalgebra over K with even component $L_0 \cong \operatorname{sl}(2, K)$, L_1 being the standard two-dimensional L_0-module.

Generalizing the notion of basic rank of a Lie variety $\mathcal{V}$ one can define a *superrank* $sr_b(\mathcal{V})$ as the least integer t such that $\mathcal{V} = \operatorname{var} G(L)$ where L is a Lie superalgebra with t homogeneous generators. It was mentioned that any special Lie variety has a finite superrank. On the other hand, the superrank of almost every polynilpotent Lie variety is infinite.

THEOREM 5.9 [Za8]. *Let $s \geqslant 2$. Then the superrank of any product $\mathcal{N}_{t_1} \cdots \mathcal{N}_{t_s}$ of nilpotent Lie varieties is infinite except for the case $s = 2$, $t_2 = 1$.*

In particular, any solvable Lie variety $\mathcal{A}^n$, $n \geqslant 3$, is of infinite superrank.

6. Finiteness conditions

6.1. *Types of finiteness conditions*

One of the most typical finiteness condition for a rings is being Noetherian. A colour Lie superalgebra is called *Noetherian* if any properly ascending chain of ideals $H_1 \subset H_2 \subset \cdots$ in it has finite length. We shall consider also superalgebras with the Hopf property. A colour Lie superalgebra L possesses Hopf property if any epimorphism of L (i.e. surjective homomorphism) is an automorphism. In other words, if φ is an endomorphism of L then the equality $\varphi(L) = L$ implies $\operatorname{Ker}\varphi = 0$. Any Noetherian superalgebra possesses the Hopf property. The infinite-dimensional Abelian Lie algebra L with linear basis $a_1, a_2, \ldots$ and endomorphism $\varphi: L \to L$ such that $\varphi(a_1) = 0$, $\varphi(a_i) = a_{i-1}$, $i > 1$, is an example of a superalgebra without Hopf's property. Note that any free colour Lie superalgebra $L(X)$ over an infinite field K is not Noetherian but it satisfies the Hopf condition whenever $1 < |X| < \infty$.

Another type of finiteness condition is *matrix representability*. Let $L = \bigoplus_{g\in G} L_g$ be a colour Lie superalgebra over a field K graded by a group G with a given bilinear form ε. We say that L is a matrix representable superalgebra if there exist an extension $\widetilde{K}$ of the ground field K and a finite-dimensional $\widetilde{K}$-algebra $\widetilde{L} = \bigoplus_{g\in G} \widetilde{L}_g$ with the same G and ε such that L is embeddable in $\widetilde{L}$ as a K-algebra.

A Lie superalgebra L over a field K is called *residually finite* if for any $x \neq 0$, $x \in L$ there exists a homomorphism $\varphi: L \to H$ onto a finite-dimensional over K superalgebra H such that $\varphi(x) \neq 0$. It is easy to see that L is residually finite if, and only if, it may be embedded in Cartesian product $\prod_\alpha L_\alpha$ where any L_α is of finite dimension. Residual finiteness and representability are closely related.

THEOREM 6.1. *Let K be an infinite field, $A = \bigoplus_{g\in G} A_g$ be a G-graded K-algebra generated by a finite set of elements, G being a finite commutative semigroup. Then A is representable if and only if it can be embedded in Cartesian product $\prod_\alpha A_\alpha$ of n-dimensional G-graded K-algebras A_α for some n.*

The following examples illustrate the above definitions. Let G be an arbitrary Abelian group with alternating bilinear form ε and let g, h be two elements of G (possibly not distinct). An example of a colour Lie superalgebra which is not residually finite is given by the Heisenberg superalgebra $\Gamma = \Gamma(g,h) = \bigoplus_{r\in G} \Gamma_r$. All components Γ_r, except Γ_g, Γ_h and Γ_{g+h}, are equal to zero. A basis of this superalgebra is formed by the set $\{a_i, b_i, c \mid i \in \mathbf{Z}\}$ where $a_i \in \Gamma_g$, $b_i \in \Gamma_h$, $c \in \Gamma_{g+h}$. The commutator is given by the formula $[a_i, b_i] = -\varepsilon(g,h)[b_i, a_i] = c$, $i \in \mathbf{Z}$, with all remaining commutators being zero. It is not residually finite since any ideal of finite codimension includes a nontrivial linear combination $\lambda_1 a_1 + \cdots + \lambda_n a_n$, thus also c.

To obtain an example of a finitely generated colour Lie superalgebra it is necessary to adjoin to $\Gamma(g,h)$ a binding derivation $\delta : \Gamma(g,h) \to \Gamma(g,h)$ and to place it into the zero component of the newly born algebra. As a result, we obtain a superalgebra $B = B(g,h)$ all of whose components, except zero, are the same as in $\Gamma(g,h)$ and $B_0 = \Gamma_0 \oplus \langle\delta\rangle$ with commutators given by

$$\begin{aligned} [\delta, a_i] &= \varepsilon(g,h)a_{i+1} - \varepsilon(h,g)a_{i-1}, \\ [\delta, b_i] &= \varepsilon(h,g)b_{i+1} - \varepsilon(g,h)b_{i-1}. \end{aligned}$$

The superalgebra $B(g,h)$ is not residually finite (and, as a concequence, is not representable) since it contains $\Gamma(g,h)$.

6.2. *Finiteness of irreducible representations*

The assertions of the two following theorems are similar to conditions of existence of polynomial identities in universal enveloping algebra.

THEOREM 6.2. *Let G be a finite group and $L = \bigoplus_{g\in G} L_g$ a colour Lie superalgebra over an algebraically closed field K of zero characteristic. Then the dimensions of all irreducible L-representations are bounded by some finite number if and only if there exists a homogeneous L_+-submodule $M \subset L_-$ such that*

(1) $\dim L_-/M < \infty$;
(2) L_+ *is Abelian and* $[M, M] = 0$;
(3) $\dim_K L_+ < |K|$.

THEOREM 6.3. *Let G be a finite group and $L = \bigoplus_{g\in G} L_g$ be a colour Lie superalgebra over an algebraically closed field K of positive characteristic. Then the dimensions of all irreducible L-representations are bounded by some finite number if and only if there exists a homogeneous ideal $R \subset L$ such that*

(1) $\dim L/R < \infty$;
(2) $R^2 \subset R_-$;
(3) *all inner derivations* $\operatorname{ad} x|_{L_+}$, $x \in L_g$, $g \in G_+$, *are algebraic of bounded degree*;
(4) $\dim_K L_+ < |K|$.

6.3. *Identities, maximality and Hopf conditions*

We present some sufficient conditions of maximality of Lie superalgebras in terms of their identities.

THEOREM 6.4. *Let $L = \bigoplus_{g\in G} L_g$ be a finitely generated solvable colour Lie superalgebra over a field K which is graded by a finite group G. If for any $g_1, g_2 \in G$, $h \in G_+$ an idenity of the form*

$$[x, y^{(n)}, z] = \sum_{j=1}^{n} \alpha_j [y^{(j)}, x, y^{(n-j)}, z], \tag{18}$$
$$d(x) = g_1, \qquad d(z) = g_2, \qquad d(y) = h, \quad \alpha_1, \dots, \alpha_n \in K$$

holds on L and L_+^m for some $m \geqslant 1$ acts on L as a nilpotent space of transformations then L is a Noetherian L-module.

Here we use the notation $[a_1, a_2, \dots, a_m, a_{m+1}]$ for the left-normed commutator $[a_1, [a_2, \dots, [a_m, a_{m+1}] \dots]$. In case of $a_1 = \dots = a_m = b$ we write it as $[b^{(m+1)}, a_{m+1}]$. As usual, L_+^m denotes the m-th term of lower central series of algebra L_+.

Any colour Lie superalgebra satisfying the above conditions is Noetherian and possesses the Hopf property.

An example of a colour Lie superalgebra satisfying all assertions of this theorem, in particular, identities (18), is a metabelian algebra.

Now let L be a usual Lie superalgebra, $G = \mathbf{Z}_2$, K an infinite field. Then only identities of the form (18) guarantee the Noetherian and Hopf conditions provided that L is finitely generated and solvable.

THEOREM 6.5. *For a given locally soluble variety $\mathcal{V}$ of Lie superalgebras over an infinite field K the following conditions are equivalent:*

(1) *$\mathcal{V}$ is a locally Noetherian variety,*

(2) *any finitely generated Lie superalgebra $L = L_0 \oplus L_1$ in $\mathcal{V}$ is a Noetherian L_0-module,*

(3) *$\mathcal{V}$ is a locally Hopf variety,*

(4) *$\mathcal{V}$ satisfies three identities*

$$[x, y^{(n)}, z] = \sum_{j=1}^{n} \alpha_j^i [y^{(j)}, x, y^{(n-j)}, z] \tag{19}$$

with $\alpha_j^i \in K$, $i = 1, 2, 3$, where $d(x) = d(z) = 0$ for $i = 1$, $d(x) = 0$, $d(z) = 1$ for $i = 2$ and $d(x) = d(z) = 1$ if $i = 3$. The element y is even for all $i = 1, 2, 3$.

We say that a variety $\mathcal{V}$ has some property *locally* if any finitely generated algebra from $\mathcal{V}$ possesses this property.

6.4. *Residual finiteness and representability*

First note that residual finiteness of a colour Lie superalgebra L and its universal enveloping algebra $U(L)$ are equivalent.

THEOREM 6.6 [Mik6]. *Let K be a field,* $\operatorname{char} K \neq 2, 3$, *$L$ a colour Lie superalgebra over K. Then L is residually finite if, and only if, its universal enveloping algebra $U(L)$ is residual finite.*

In the case $\operatorname{char} K = p > 3$ *let M be a colour Lie p-superalgebra and $u(M)$ its restricted universal enveloping algebra. Then M is residually finite if, and only if, $u(M)$ is residually finite.*

In the case of Lie algebras this theorem is due to W. Michaelis [Mic1,Mic2].

Sufficient conditions for residual finiteness or representability of formulated algebra L may be both in terms of its ideals and in terms of identities on L.

THEOREM 6.7. *Let $L = \bigoplus_{g \in G} L_g$ be a finitely generated colour Lie superalgebra over an infinite field K and G a finite group. Suppose that L contains an Abelian ideal A of finite codimension and, in addition, L_+^2 lies in A in case* $\operatorname{char} K = 0$. *Then L is a representable algebra.*

THEOREM 6.8. *Let L be a finitely generated soluble colour Lie superalgebra over an infinite field K graded by a finite group G. If for any $g_1, g_2 \in G$, $h \in G_+$, L satisfies an identity of the form* (18), *L_+^m acts on L as a nilpotent space of transformations and, in addition, $m = 2$ if* $\operatorname{char} K = 0$ *then L is representable.*

The claim $|G| < \infty$ is necessary in the previous theorems. It can be illustrated by the following example. Let $L = P + M$ be a metabelian algebra where P is an Abelian algebra with basis $\{x, y, \}$ and M is an Abelian ideal in L spanned by $\{z_i \mid i \in \mathbf{Z}\}$. The multiplication in L is given by the following: $[x, z_i] = z_{i-1}$, $[y, z_i] = z_{i+1}$, $i \in \mathbf{Z}$. So, L is a $\mathbf{Z}$-graded algebra where $L_i = \langle z_i \rangle$ for $i \neq \pm 1$, $L_1 = \langle y, z \rangle$, $L_{-1} = \langle x, z_{-1} \rangle$. If we define $\varepsilon(m, n) = 1$ for all $m, n \in \mathbf{Z}$ then L becomes a colour Lie superalgebra with nonzero least homogeneous ideal M. So, L is a 3-generator monolithic metabelian algebra of infinite dimension. Hence it is not residually finite.

Now consider Lie superalgebras ($G = \mathbf{Z}_2$) over an infinite field field K. Then the sufficient conditions for residual finiteness and representability in terms of identities cannot be extended to a wider class of solvable varieties.

THEOREM 6.9. *Let $\mathcal{V}$ be a locally solvable variety of Lie superalgebras over an infinite field K,* $\operatorname{char} K > 2$. *Then the following conditions are equivalent*:

(1) *$\mathcal{V}$ is locally representable*;

(2) *$\mathcal{V}$ is locally residually finite*;

(3) *$\mathcal{V}$ satisfies three identities* (19).

Note that the nilpotency of the action of L_0^m on L for some m is a consequence of the identities (19) for usual Lie superalgebras. If $\operatorname{char} K = 0$, one can drop the restriction of locally solvability.

THEOREM 6.10. *For a variety of Lie superalgebras $\mathcal{V}$ over a field K of zero characteristic the following conditions are equivalent.*

(1) $\mathcal{V}$ *is locally representable*;
(2) $\mathcal{V}$ *is locally residually finite*;
(3) $\mathcal{V}$ *satisfies three identities* (19) *and, in addition, on any finitely generated Lie superalgebra in* $\mathcal{V}$ *the following identity holds*:

$$\begin{aligned} &\big[[x_1, y_1], \ldots, [x_m, y_y], z\big] = 0, \\ &d(x_i) = d(y_i) = 0, \quad i = 1, \ldots, m, \end{aligned} \tag{20}$$

for some m.

This result does not extend the class of locally representable varieties since the conditions of the theorem lead to locally solvability.

The class of locally representable varieties of colour Lie superalgebras over an infinite field K is rather wide. To illustrate this, denote by $\mathcal{N}_t$ the variety of all nilpotent K-superalgebras with nilpotency degree no greater than t, i.e. N_t is defined by the nongraded identity $[x_1, \ldots, x_{s+1}] = 0$. The product $\mathcal{W}\mathcal{V}$ of two varieties consists of all superalgebras L such that for some ideal H in L one has: $H \in \mathcal{W}$, $L/H \in \mathcal{V}$. Then any variety $\mathcal{A}\mathcal{N}_t$ where $\mathcal{A} = \mathcal{N}_1$ over an infinite field of positive characteristic is locally representable. Over a field of zero characteristic any variety $\mathcal{N}_s\mathcal{A} \cap \mathcal{A}\mathcal{N}_t$ is locally representable.

It is interesting that the identities of the form (20) in all finitely generated algebras from $\mathcal{V}$ do not imply the same relation in an arbitrary algebra $L \in \mathcal{V}$. So, consider the variety $\mathcal{V}$ of Lie algebras over a field K, $\operatorname{char} K = 0$, given by two identities: $[[x_1, x_2, x_3], [y_1, y_2, y_3]] = 0$, $[[x_1, x_2], [x_1, x_3], [x_2, x_4]] = 0$. It is known that $\mathcal{V}$ does not lie in $\mathcal{N}_t\mathcal{A}$ but any finitely generated Lie algebra from $\mathcal{V}$ has a nilpotent commutant. This variety is also locally residually finite and locally representable. Another significant property of $\mathcal{V}$ is that it is the unique solvable variety of infinite basic rank such that $r_b(\mathcal{W}) < \infty$ whenever $\mathcal{W} \subset \mathcal{V}$, $\mathcal{W} \neq \mathcal{V}$.

References

[Ab] E. Abe, *Hopf Algebras*, Cambridge University Press (1980).

[Bab] I.K. Babenko, *Analytic properties of Poincaré series of a loop space*, Mat. Zametki **27** (1980), 751–765.

[Bah1] Yu.A. Bahturin, *Approximations of Lie algebras*, Mat. Zametki **12** (1972), 713–716. English transl. Math. Notes **12** (1972), 868–870.

[Bah2] Yu.A. Bahturin, *Identities in the universal envelopes of Lie algebras*, J. Austral. Math. Soc. **18** (1974), 10–21.

[Bah3] Yu.A. Bahturin, *Lectures on Lie Algebras*, Akademie-Verlag, Berlin (1978).

[Bah4] Yu.A. Bahturin, *A remark on irreducible representations of Lie algebras*, J. Austral. Math. Soc. **27** (1979), 332–336.

[Bah5] Yu.A. Bahturin, *Identities in the universal enveloping algebra for a Lie superalgebra*, Mat. Sb. **127** (1985), 384–397. English transl. Math. USSR-Sb. **55** (1986), 383–396.

[Bah6] Yu.A. Bahturin, *Identical Relations in Lie Algebras*, VNU Science Press, Utrecht (1987).

[Bah7] Yu.A. Bahturin, *On degrees of irreducible representations of Lie superalgebras*, C. R. Math. Acad. Sci. Canada **10** (1987), 19–24.

[Bah8] Yu.A. Bahturin, *Lie superalgebras with bounded degrees of irreducible representations*, Mat. Sb. **180** (1989), 195–206. English transl. Math. USSR-Sb. **66** (1990), 199–209.

[BD] Yu.A. Bahturin and V.S. Drensky, *Identities of the soluble colour Lie superalgebras*, Algebra i Logika **26** (1987), 403–418. English transl. Algebra and Logic **26** (1987) 229–240.

[BMPZ] Yu.A. Bahturin, A.A. Mikhalev, M.V. Zaitsev and V.M. Petrogradsky, *Infinite Dimensional Lie Superalgebras*, Walter de Gruyter, Berlin (1992).

[BMo] Yu. Bahturin and S. Montgomery, *PI-envelopes of Lie superalgebras*, Proc. Amer. Math. Soc. **127** (1999), 2829–2839.

[BZ1] Yu.A. Bahturin and M.V. Zaicev, *Residual finiteness of colour Lie superalgebras*, Trans. Amer. Math. Soc. **337** (1993), 159–180.

[BZ2] Yu.A. Bahturin and M.V. Zaicev, *Identities of graded algebras*, J. Algebra **205** (1998), 1–12.

[Beh] E.J. Behr, *Enveloping algebras of Lie superalgebras*, Pacif. J. Math. **130** (1987), 9–25.

[Ber] F.A. Berezin, *Introduction to Algebra and Analysis with Anticommuting Variables*, Izdat MGU, Moscow (1983).

[BP] J. Bergen and D.S. Passman, *Delta methods in enveloping rings*, J. Algebra **133** (1990), 277–312.

[Berg] G.M. Bergman, *The diamond lemma for ring theory*, Adv. in Math. **29** (1978) 178–218.

[Bo] L.A. Bokut', *Embeddings into simple associative algebras*, Algebra i Logika **15** (1976), 117–142. English transl. Algebra and Logic **15** (1976), 73–90.

[BK] L.A. Bokut' and G.P. Kukin, *Algorithmic and Combinatorial Algebra*, Kluwer Academic Publishers, Dordrecht (1994).

[BKLM] L.A. Bokut', S.-J. Kang, K.-H. Lee and P. Malcolmson, *Gröbner–Shirshov bases for Lie superalgebras and their universal enveloping algebras*, J. Algebra **217** (1999), 461–495.

[Br] M.R. Bryant, *On the fixed points of a finite group acting on a free Lie algebra*, J. London Math. Soc. (2) **43** (1991), 215–224.

[CFL] K.T. Chen, R.H. Fox and R.C. Lyndon, *Free differential calculus IV. The quotient groups of the lower central series*, Ann. Math. (2) **68** (1958), 81–95.

[Co] P.M. Cohn, *Subalgebras of free associative algebras*, Proc. London Math. Soc. **14** (1964), 618–632.

[Di] J. Dixmier, *Algebres Enveloppantes*, Gauthier-Villars, Paris (1974).

[Dr] V. Drensky, *Fixed algebras of residually nilpotent Lie algebras*, Proc. Amer. Math. Soc. **120** (1994), 1021–1028.

[GK1] V.P. Gerdt and V.V. Kornyak, *Construction of finitely presented Lie algebras and superalgebras*, J. Symbolic Comput. **21** (1996), 337–349.

[GK2] V.P. Gerdt and V.V. Kornyak, *A program for constructing a complete system of relations, basis elements and tables of commutators of finitely presented Lie algebras and superalgebras*, Programmirovanie No. 3 (1997), 58–71. English transl. Program. Comput. Software **23** (3) (1997), 164–172.

[Hall M.] M. Hall, *A basis for free Lie algebras and higher commutators in free groups*, Proc. Amer. Math. Soc. **1** (1950), 575–581.

[Hall P.] P. Hall, *A contribution to the theory of groups of a prime power order*, Proc. London Math. Soc. **36** (1934), 29–95.

[Han] K.C. Hannabuss, *An obstruction to generalized superalgebras*, J. Phys. A **20** (17) (1987), L1135–L1138.

[He] I.N. Herstein, *Noncommutative Rings*, Carus Math. Monograph 15, MMA (1968).

[Ja] N. Jacobson, *Lie Algebras*, Wiley-Interscience, New York (1962).

[Kac1] V.G. Kac, *Lie superalgebras*, Adv. in Math. **26** (1977), 8–96.

[Kac2] V.G. Kac, *A sketch of Lie superalgebras theory*, Comm. Math. Phys. **53** (1977), 31–64.

[Kac3] V.G. Kac, *Classification of simple* $\mathbf{Z}$*-graded Lie superalgebras*, Comm. Algebra **5** (1977), 1375–1400.

[Kac4] V.G. Kac, *Classification of infinite-dimensional simple linearly compact Lie superalgebras*, Adv. Math. **139** (1998), 1–55.

[Kang] S.-J. Kang, *Graded Lie superalgebras and the superdimension formula*, J. Algebra **204** (1998), 597–655.

[Kant] I.L. Kantor, *An analogue of E. Witt's formula for the dimensions of homogeneous components of free Lie superalgebras*, Differential Geometry and Lie Algebras (Moscow Region Pedagogical Institute, 1983, in Russian), Deposited VINITI, 17th April 1984, No. 2384-84 Dep.

[Ku] G.P. Kukin, *On subalgebras of free Lie p-algebras*, Algebra i Logika **11** (1972), 535–550. English transl. Algebra and Logic **11** (1972), 33–50.

[La] V.N. Latyshev, *Two remarks on PI-algebras*, Sibirsk. Mat. Zh. **4** (1963) 1120–1121.

[Le] D.A. Leites, *Lie superalgebras*, VINITI, Itogi Nauki i Tekhniki, Modern Problems of Math. Recent Progress, Vol. 25 (1984), 3–50. English transl. J. Soviet Math. **30** (1985), 2481–2512.

[Ma1] W. Magnus, *Beziehungen zwischen Gruppen und Idealen in einem speziellen Ring*, Math. Ann. **111** (1935), 259–280.

[Ma2] W. Magnus, *Über Beziehungen zwischen höheren Kommutatoren*, J. Reine Angew. Math. **177** (1937), 105–115.

[MR] J.C. McConnell and J.C. Robson, *Noncommutative Noetherian Rings*, Wiley, New York (1987).

[Mic1] W. Michaelis, *Properness of Lie algebras and enveloping algebras*, I, Proc. Amer. Math. Soc. **101** (1987), 17–23.

[Mic2] W. Michaelis, *Properness of Lie algebras and enveloping algebras*, II, Lie algebras and related topics. Canad. Math. Soc. Conf. Proc. **5** (1986), 265–280.

[Mik1] A.A. Mikhalev, *Subalgebras of free colour Lie superalgebras*, Mat. Zametki **37** (1985), 653–661. English transl. Math. Notes, 356–360.

[Mik2] A.A. Mikhalev, *Free colour Lie superalgebras*, Dokl. Akad. Nauk SSSR **286** (1986), 551–554. English transl. Soviet Math. Dokl. **33** (1986), 136–139.

[Mik3] A.A. Mikhalev, *Subalgebras of free Lie p-superalgebras*, Mat. Zametki **43** (1988), 178–191. English transl. Math. Notes, 99–106.

[Mik4] A.A. Mikhalev, *The junction lemma and the equality problem for colour Lie superalgebras*, Vestnik Moskov. Univ. Ser. 1. Mat. Mekh., No. 5 (1989), 88–91. English transl. Moscow Univ. Math. Bull., 87–90.

[Mik5] A.A. Mikhalev, *Embedding of Lie superalgebras of countable rank into a two generated Lie superalgebras*, Uspekhi Mat. Nauk SSSR **45** (1990), 139–140. English transl. Russian Math. Surveys **45** (1990), 162–163.

[Mik6] A.A. Mikhalev, *Ado–Iwasawa theorem, graded Hopf algebras and residual finiteness of colour Lie (p-)superalgebras and their universal enveloping algebras*, Vestnik. Moscov. Univ. Ser. 1. Mat. Mekh., No. 5 (1991), 72–74. English transl. Moscow Univ. Math. Bull. **46** (1991), No. 5, 51–53.

[Mik7] A.A. Mikhalev, *On fixed points of a free colour Lie superalgebra under the action of a finite group of linear automorphisms*, Uspekhi Mat. Nauk **47** (1992), 205–207. English transl. Russian Math. Surveys. **47** (1992), No. 4, 220–221.

[Mik8] A.A. Mikhalev, *A.I. Shirshov' composition techniques in Lie superalgebras (non commutative Gröbner bases)*, Trudy Sem. Petrovsk. **18** (1995), 277–289. Enlish transl. J. Math. Sci. **80** (1996), 2153–2160.

[MY] A.A. Mikhalev and J.-T. Yu, *Test elements, retracts and automorphic orbits of free algebras*, Internat. J. Algebra Comput. **8** (1998), 295–310.

[MZ1] A.A. Mikhalev and A.A. Zolotykh, *Rank and primitivity of elements of free color Lie (p-)superalgebras*, Intern. J. Algebra Comput. **4** (1994), 617–656.

[MZ2] A.A. Mikhalev and A.A. Zolotykh, *Automorphisms and primitive elements of free Lie superalgebras*, Comm. Algebra **22** (1994), 5889–5901.

[MZ3] A.A. Mikhalev and A.A. Zolotykh, *Combinatorial Aspects of Lie Superalgebras*, CRC Press, Boca Raton, New York (1995).

[MZ4] A.A. Mikhalev and A.A. Zolotykh, *Ranks of subalgebras of free Lie superalgebras*, Vestnik Mosk. Univ. Ser. 1. Mat. Mekh, No. 2 (1996), 36–40. English transl. Moscow Univ. Math. Bull. **51** (2) (1996), 32–35.

[MZ5] A.A. Mikhalev and A.A. Zolotykh, *Complex of algorithms for computation in Lie superalgebras*, Programmirovanie, No. 1 (1997), 12–23. English transl. Program. Comput. Software **23** (1) (1997), 8–16.

[MT] A.I. Molev and L.M. Tsalenko, *Representation of the symmetric group in the free Lie (super) algebra and in the space of harmonic polynomials*, Funktsional. Anal. i Prilozhen. **20** (1986), 76–77. English transl. Funktional Anal. Appl. **20** (1986), 150–152.

[Mo] S. Montgomery, *Constructing simple Lie superalgebras from associative graded algebras*, J. Algebra **195** (1997), 558–579.

[Mu] I.M. Musson, *Some Lie superalgebras associated to the Weyl algebras*, Proc. Amer. Math. Soc. **127** (1999), 2821–2827.

[Ne] H. Neumann, *Varieties of Groups*, Springer, Berlin (1967).

[Pa1] D.S. Passman, *Enveloping algebras satisfying polynomial identity*, J. Algebra **134** (1990), 469–490.

[Pa2] D.S. Passman, *Simple Lie color algebras of Witt type*, J. Algebra **208** (1998), 698–721.

[Pe1] V.M. Petrogradsky, *On restricted enveloping algebras for Lie p-algebras*, Proc. 5-th Siberian Conf. on Varieties of Algebraic Systems, Barnaul (1988), 52–55.

[Pe2] V.M. Petrogradsky, *On the existence of identity in restricted enveloping algebras*, Mat. Zametki **49** (1991), 84–93. English transl. Math. Notes, 60–66.

[Pe3] V.M. Petrogradsky, *Identities in the enveloping algebras for modular Lie superalgebras*, J. Algebra **145** (1992), 1–21.

[Pe4] V.M. Petrogradsky, *Exponential Schreier's formula for free Lie algebras and its applications*, J. Math. Sci. **93** (1999), 939–950.

[Ree] R. Ree, *Generalized Lie elements*, Canad. J. Math. **12** (1960), 493–502.

[Re] G. Renault, *Sur les anneaux de groups*, C. R. Acad. Sci. Paris **273** (1971), 84–87.

[Reu] C. Reutenauer, *Free Lie Algebras*, Clavendon Press, Oxford (1993).

[RW1] V. Rittenberg and D. Wyler, *Sequences of $\mathbf{Z}_2 \oplus \mathbf{Z}_2$-graded Lie algebras and superalgebras*, J. Math. Phys. **19** (1978), 2193–2200.

[RW2] V. Rittenberg and D. Wyler, *Generalized superalgebras*, Nuclear Phys. B **139** (1978), 189–202.

[Sc1] M. Scheunert, *Generalized Lie algebras*, J. Math. Phys. **20** (1979), 712–720.

[Sc2] M. Scheunert, *The Theory of Lie Superalgebras. An Introduction*, Lecture Notes in Math., Vol. 716 (1979).

[Shi1] A.I. Shirshov, *Subalgebras of free Lie algebras*, Mat. Sb. **33** (1953), 441–452.

[Shi2] A.I. Shirshov, *On free Lie rings*, Mat. Sb. **45** (1958), 113–122.

[Shi3] A.I. Shirshov, *Some algorithmic problems about Lie algebras*, Sibirsk. Mat. Zh. **3** (1962), 292–296.

[Shi4] A.I. Shirshov, *About one conjecture in the theory of Lie algebras*, Sibirsk. Mat. Zh. **3** (1962), 297–301.

[Shi5] A.I. Shirshov, *On bases of free Lie algebras*, Algebra i Logika **1** (1962), 14–19.

[Shi6] A.I. Shirshov, *Collected Works. Rings and Algebras*, Nauka, Moscow (1984).

[Sht] A.S. Shtern, *Free Lie superalgebras*, Sibirsk. Mat. Zh. **27** (1986), 170–174.

[Sw] M.E. Sweedler, *Hopf Algebras*, Benjamin, New York (1969).

[Um] U.U. Umirbaev, *On the approximation of free Lie algebras with respect to entry*, Sci. Notes of Tartu Univ. **878** (1990), 147–152.

[Va] M.A. Vasiliev, *De Sitter supergravity with positive cosmological constant and generalized Lie superalgebras*, Class. and Quantum Gravity **2** (1985), 645–652.

[Vi] O.E. Villamayor, *On weak dimensions of algebras*, Pasific J. Math. **9** (1959), 941–951.

[Wa] A. Wais, *On special varieties of Lie algebras*, Algebra i Logika **28** (1988), 29–40.

[Wil] M.C. Wilson, *Bell's primeness criterion and the simple Lie superalgebras*, J. Pure Appl. Algebra **133** (1998), 214–260.

[Wi1] E. Witt, *Treu Darstellung Liescher Ringe*, J. Reine Angew. Math. **177** (1937), 152–160.

[Wi2] E. Witt, *Die Unterringe der freien Lieschen Ringe*, Math. Z. **64** (1956), 195–216.

[Za1] M.V. Zaicev, *Locally residually finite varieties of Lie algebras*, Mat. Zametki **44** (1988), 352–361. English transl. Math. Notes **44** (1988), 674–680.

[Za2] M.V. Zaicev, *Residual finiteness and the Noetherian property of finitely generated Lie algebras*, Mat. Sb. **136** (1988), 500–509. English transl. Math. USSR-Sb. **64** (1989), 495–504.

[Za3] M.V. Zaicev, *Locally Hopf varieties of Lie algebras*, Mat. Zametki **45** (1989), 56–61. English transl. Math. Notes **45** (1989), 469–472.

[Za4] M.V. Zaicev, *Locally representable varieties of Lie algebras*, Mat. Sb. **180** (1989), 798–808. English transl. Math. USSR-Sb. **67** (1990), 249–259.

[Za5] M.V. Zaicev, *Locally representable varieties of Lie superalgebras*, Mat. Zametki **52** (5) (1992), 33–41. English transl. Math. Notes **52** (1992), 1101–1106.

[Za6] M.V. Zaicev, *Basic rank of varieties of Lie algebras*, Mat. Sb. **184** (9) (1993), 21–40. English transl. Sbornik: Math. **80** (1995), 15–31.

[Za7] M.V. Zaicev, *On an extremal variety of Lie superalgebras*, Vestnik Moskov. Univ. Ser. 1 Mat. Mekh. No. 6 (1997), 18–20. English transl. Moscow Univ. Math. Bull. **52** (6) (1997), 11–14.

[Za8] M.V. Zaicev, *A superrank of varieties of Lie algebras*, Algebra i Logika **37** (1998), 394–412. English transl. Algebra and Logic **37** (1998), 223–233.

[ZM] M.V. Zaicev and S.P. Mishchenko, *A criterion for the polynomial growth of varieties of Lie superalgebras*, Izv. Ross. Akad. Nauk Ser. Mat. **62** (5) (1998), 103–116. English transl. Izv. Math. **62** (1998), 953–967.

Nilpotent and Solvable Lie Algebras

Michel Goze and Yusupdjan Khakimdjanov

University of Haute Alsace, Faculty of Sciences, 4, rue des Frères Lumière, 68093 Mulhouse Cédex, France

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

In this chapter we consider finite-dimensional solvable and nilpotent Lie algebras. These classes of Lie algebras play an important role in the structure theory. A Lie algebra over a field of characteristic 0 can be presented as a semidirect sum of a semisimple subalgebra and its radical. The classical results of Cartan and Killing give a complete classification of semisimple Lie algebras. Thus the structure theory of Lie algebras is reduced to the study of solvable Lie algebras. The class of nilpotent Lie algebras is a very important subclass in the class of the solvable ones. We can say that the problem of their description is reduced to a description of the nilpotent ones. Unfortunately we are very far from the complete classification of these classes of Lie algebras.

The first non-trivial classifications of some classes of low-dimensional nilpotent Lie algebras are due to Umlauf. In his thesis (Leipzig, 1891), he presented the list of nilpotent Lie algebras of dimension $m \leqslant 6$. He gave also the list of nilpotent Lie algebras of dimension $m \leqslant 9$ admitting a basis $(X_0, X_1, \ldots, X_{m-1})$ with $[X_0, X_i] = X_{i+1}$ for $i = 1, \ldots, m-2$ (now, the nilpotent Lie algebras with this property are called filiform Lie algebras). Umlauf's list of filiform Lie algebras is exact only for dimension $m \leqslant 7$; in dimension 8 and 9 this list contains errors and it is incomplete. The list of nilpotent Lie algebras of dimension $\leqslant 6$ also contains errors.

The importance of nilpotent Lie algebras in classification problems and many of their applications in different domains of Mathematics and Physics have contributed over the last years to go back to study this class of Lie algebras. Unfortunately, many of these papers are based on direct computations and the complexity of those computations frequently leads to errors. The first exact classification of nilpotent Lie algebras in dimension $\leqslant 6$ is probably due to Morosov. All existing classifications in dimension 7 probably contain some errors and are incomplete.

The results presented here give the main classical theorems and some directions for research.

1. Generalities

1.1. *Derived series, central series*

Let $\mathfrak{g}$ be a Lie algebra over a field K. We put $D^0\mathfrak{g} = \mathfrak{g}$; $D^1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ and more generally $D^{k+1}\mathfrak{g} = [D^k\mathfrak{g}, D^k\mathfrak{g}]$ for $k \geqslant 0$. Each of these subspaces is an ideal of $\mathfrak{g}$ and we have the following decreasing series, called the *derived* series of $\mathfrak{g}$.

$$\mathfrak{g} = D^0\mathfrak{g} \supset D^1\mathfrak{g} \supset \cdots \supset D^k\mathfrak{g} \supset \cdots.$$

The *descending central series* of $\mathfrak{g}$:

$$\mathfrak{g} = C^0\mathfrak{g} \supset C^1\mathfrak{g} \supset \cdots \supset C^k\mathfrak{g} \supset \cdots$$

is defined by the following ideals:

$$C^0\mathfrak{g} = \mathfrak{g}, \quad C^{k+1}\mathfrak{g} = [C^k\mathfrak{g}, \mathfrak{g}], \quad k = 0, 1, 2, \ldots.$$

Note that this series decreases more slowly that the derived sequence and the Lie algebra $C^{i+1}\mathfrak{g}/C^{i+2}\mathfrak{g}$ is an ideal of $C^{i}\mathfrak{g}/C^{i+2}\mathfrak{g}$ contained in its center. The *ascending central series*

$$\{0\} = C_0(\mathfrak{g}) \subset C_1(\mathfrak{g}) \subset \cdots \subset C_k(\mathfrak{g}) \subset \cdots$$

is defined by:

$$C_0(\mathfrak{g}) = 0, \quad C_k(\mathfrak{g}) = \{x \in \mathfrak{g}\colon [x, \mathfrak{g}] \subset C_{k-1}(\mathfrak{g})\} \quad \text{for } k \geqslant 1.$$

1.2. *Definition of solvable Lie algebras*

DEFINITION 1. A Lie algebra $\mathfrak{g}$ is called *solvable* if there exists an integer k such that $D^k\mathfrak{g} = \{0\}$.

EXAMPLES OF SOLVABLE LIE ALGEBRAS.

(1) An Abelian Lie algebra.
(2) The Lie algebra of the affine group of the straight line: it is generated by two independent vectors x and y satisfying: $[x, y] = x$.
(3) The subalgebra of $\mathrm{gl}(n, K)$ formed by the upper triangular matrices.
(4) Let V be a vector space of dimension n and let $D = (0 = V_0 \subset V_1 \subset \cdots \subset V_n)$ be a flag of V, that is a sequence of increasing vector subspaces such that $\lim V_i = i$. Let $\mathfrak{b}(D) = \{f \in \operatorname{End} V\colon\ f(V_i) \subset V_i \text{ for all } i \geqslant 0\}$. Then $\mathfrak{b}(D)$ is a solvable Lie algebra for the bracket:

$$[f, g] = f \circ g - g \circ f.$$

The matrices of the elements of $\mathfrak{b}(D)$ relative to an adapted basis of the flag are triangular.

PROPOSITION 1. *A Lie algebra $\mathfrak{g}$ is solvable if and only if there exists a descending series of ideals of $\mathfrak{g}$*

$$\mathfrak{g} = I_0 \supset I_1 \supset \cdots \supset I_k = \{0\}$$

such that $[I_j, I_j] \subset I_{j+1}$ for $0 \leqslant j \leqslant k-1$.

PROPOSITION 2. *Let $\mathfrak{g}$ be a Lie algebra. Then*

(a) *If $\mathfrak{g}$ is solvable, each subalgebra and each quotient algebra of $\mathfrak{g}$ are solvable.*
(b) *If I is a solvable ideal of $\mathfrak{g}$ such that $\mathfrak{g}/I$ is solvable, then $\mathfrak{g}$ is also solvable.*
(c) *If I and J are solvable ideals of $\mathfrak{g}$, then $I + J$ is also solvable.*

COROLLARY 3. *$\mathfrak{g}$ contains a maximal solvable ideal $\mathfrak{r}$ which contains every solvable ideal.*

The maximal solvable ideal $\mathfrak{r}$ of $\mathfrak{g}$ is called the *radical* of $\mathfrak{g}$.

1.3. *Definition of nilpotent Lie algebras*

DEFINITION 2. A Lie algebra $\mathfrak{g}$ is called *nilpotent* if there exists an integer k such that $C^k\mathfrak{g} = \{0\}$. The smallest integer k such that $C^k\mathfrak{g} = \{0\}$ is called the *nilindex* (or the *nilpotency class*) of $\mathfrak{g}$.

EXAMPLES OF NILPOTENT LIE ALGEBRAS.

(1) An Abelian Lie algebra.

(2) The Heisenberg algebra H_k defined on the basis $(e_1, e_2, \ldots, e_{2k+1})$ by $[e_{2i-1}, e_{2i}] = e_{2k+1}$, $i = 1, \ldots, k$ (all other brackets being zero). Its nilindex is equal to 2.

(3) The n-dimensional algebra defined on the basis $(e_1, \ldots, e_n)$ by the brackets

$$[e_1, e_i] = e_{i+1}, \quad 2 \leqslant i \leqslant n-1.$$

Its nilindex is equal to $n-1$.

(4) The subalgebra of $\mathrm{gl}(n, K)$ formed by strictly upper triangular matrices (these are the matrices $A = (a_{ij})$ with $a_{ij} = 0$ for $i \leqslant j$).

Let V be a vector space and $D = (V_0 \subset V_1 \subset \cdots \subset V_n = V)$ a flag of length n. The Lie algebra

$$\mathfrak{n}(D) = \{f \in \mathrm{End}(V)\colon\ f(V_i) \subset V_{i-1} \text{ for all } i \geqslant 1\}$$

is nilpotent. We note that in a basis adapted to D the Lie algebra $\mathfrak{n}(D)$ consists of those endomorphisms whose matrices are strictly upper triangular.

PROPOSITION 4. *Let $\mathfrak{g}$ be a Lie algebra. Then*

(a) *Each subalgebra and each quotient algebra of a nilpotent Lie algebra $\mathfrak{g}$ are also nilpotent.*

(b) *Let $Z(\mathfrak{g})$ be the center of $\mathfrak{g}$. If $\mathfrak{g}/Z(\mathfrak{g})$ is nilpotent, then $\mathfrak{g}$ is also nilpotent.*

(c) *If $\mathfrak{g}$ is nilpotent, then $Z(\mathfrak{g}) \neq \{0\}$.*

(d) *If I and J are nilpotent ideals of $\mathfrak{g}$, then $I \cap J$ and $I + J$ also are nilpotent.*

COROLLARY 5. $\mathfrak{g}$ *contains a unique maximal nilpotent ideal $\mathfrak{n}$ called the nilradical of* $\mathfrak{g}$.

PROPOSITION 6. *A nilpotent Lie algebra is solvable.*

In fact, $C^i\mathfrak{g} \subset D^i\mathfrak{g}$.

PROPOSITION 7. *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if there exists a descending sequence of ideals*

$$\mathfrak{g} = I_0 \supset I_1 \supset \cdots \supset I_k = \{0\}$$

such that $[\mathfrak{g}, I_j] \subset I_{j+1}$, $0 \leqslant j \leqslant k-1$.

PROPOSITION 8. *A Lie algebra* $\mathfrak{g}$ *is nilpotent with a nilindex* $\leqslant k$ *if and only if*

$$\big[x_0, \big[x_1, \big[x_2, \cdots [x_{k-1}, x_k]\big]\big] \cdots \cdots\big] = 0$$

for all $x_0, x_1, \ldots, x_k \in \mathfrak{g}$.

1.4. *The Engel Theorem*

Let $\mathfrak{g}$ be a Lie algebra and let $x \in \mathfrak{g}$. Recall that $\operatorname{ad} x$ is the endomorphism of $\mathfrak{g}$ defined by

$$\operatorname{ad} x(y) = [x, y].$$

THEOREM 9. *A Lie algebra* $\mathfrak{g}$ *is nilpotent if and only if* $\operatorname{ad} x$ *is nilpotent for every* x *in* $\mathfrak{g}$.

The proof is based on the following lemma (the Engel lemma):

LEMMA 10. *Let* V *be a nontrivial space over* K *and* $\mathfrak{g}$ *a subalgebra of* $\mathrm{gl}(V)$. *Suppose that every* x *in* $\mathfrak{g}$ *is a nilpotent endomorphism of* V. *Then there exists a* $v \neq 0 \in V$ *such that* $x(v) = 0$, *for all* $x \in \mathfrak{g}$.

PROOF. We use an induction on the dimension n of $\mathfrak{g}$. The lemma is true for $n = 1$. Let $n \geqslant 2$. Suppose that it is true for the dimension less than n. We begin by showing that $\mathfrak{g}$ contains an ideal of codimension 1. Let $\mathfrak{h}$ be a subalgebra of dimension $m < n$ of $\mathfrak{g}$. Let $x \in \mathfrak{h}$; it is a nilpotent endomorphism of V. This implies that the endomorphism $\operatorname{ad} x$ of $\mathfrak{h}$ is also nilpotent. Indeed the terms of the decomposition of $(\operatorname{ad} x)^p(y)$ have the form $x^i y x^k$ with $i + k = p$. As $x^p = 0$ for some power p, there exists an integer l such that $(\operatorname{ad} x)^l(y) = 0$. Now let $\sigma(x) : \mathfrak{g}/\mathfrak{h} \to \mathfrak{g}/\mathfrak{h}$ the endomorphism defined from $\operatorname{ad} x$ by projection on the quotient. As $\operatorname{ad} x$ is nilpotent, the endomorphism $\sigma(x)$ is also nilpotent. Then, by hypothesis, there is an element $Y \in \mathfrak{g}/\mathfrak{h}$ such that $\sigma(x)(Y) = 0$ for all x in $\mathfrak{h}$. Let y in $\mathfrak{g}$ be a representative of the class Y. Then $[x, y]$ is in $\mathfrak{h}$ for all $x \in \mathfrak{h}$ and $\mathfrak{h}$ is an ideal of the $(m + 1)$-dimensional subalgebra of $\mathfrak{g}$ spanned by $\mathfrak{h}$ and y. If we iterate this process, we find a codimension 1 ideal of $\mathfrak{g}$. Let $\mathfrak{h}$ be this ideal. Choose $a \in \mathfrak{g} \setminus \mathfrak{h}$. The set $U = \{v \in V\colon\ x(v) = 0 \text{ for all } x \in \mathfrak{h}\}$ is not null. As $x(a(v)) = [x, a](v) = 0$ for all $v \in U$, the space U is a-invariant. By the induction hypothesis, there exists an element $u \in U$, $u \neq 0$ such that $a(u) = 0$. As $x(u) = 0$ for all $x \in U$ and $\mathfrak{g} = \mathfrak{h} \oplus Ka$, the lemma is proved. □

PROOF (OF THE THEOREM). The necessary condition is a consequence of Proposition 6. For prove the sufficient condition we use an induction on the dimension n of $\mathfrak{g}$. The theorem is clear for $n = 1$. Suppose the theorem is true for the Lie algebras of dimension less than n where $n \geqslant 2$. Let $\mathfrak{g}$ be a n-dimensional Lie algebra such that $\operatorname{ad} x$ is nilpotent for all x. From the Engel lemma, it exists $y \in \mathfrak{g}$, $y \neq 0$ such that $[x, y] = 0$ for all $x \in \mathfrak{g}$. This proves that the center $Z(\mathfrak{g})$ of $\mathfrak{g}$ is not trivial. Let $\mathfrak{g}'$ be the quotient algebra $\mathfrak{g}/Z(\mathfrak{g})$. By hypothesis $\mathfrak{g}'$ is nilpotent and from the Proposition 4, $\mathfrak{g}$ is nilpotent. □

COROLLARY 11. *Let $\mathfrak{g}$ be a Lie algebra, $\mathfrak{l}$ an ideal of $\mathfrak{g}$. If the quotient Lie algebra $\mathfrak{g}/\mathfrak{l}$ is nilpotent and if for all x in $\mathfrak{g}$, the restriction of* $\operatorname{ad} x$ *to $\mathfrak{l}$ is nilpotent, then the Lie algebra $\mathfrak{g}$ is also nilpotent.*

COROLLARY 12. *Let $\mathfrak{g}$ be a nilpotent Lie algebra. Then*
(i) $\operatorname{codim}[\mathfrak{g}, \mathfrak{g}] \geqslant 2$.
(ii) *If $\mathfrak{a}$ is a subspace of $\mathfrak{g}$, such that $\mathfrak{g} = \mathfrak{a} + [\mathfrak{g}, \mathfrak{g}]$, then $\mathfrak{a}$ generates $\mathfrak{g}$.*

1.5. *Lie's theorem*

The field K is algebraic closed and of characteristic 0.

Let $\mathfrak{g}$ be a Lie algebra and V a vector space over K. A linear representation of $\mathfrak{g}$ on V is a homomorphism of Lie algebras

$$\Phi : \mathfrak{g} \to \operatorname{gl}(V)$$

that is a linear map satisfying

$$\Phi[x, y]_{\mathfrak{g}} = \left[\Phi(x), \Phi(y)\right]_{\operatorname{gl}(V)} = \Phi(x) \circ \Phi(y) - \Phi(y) \circ \Phi(x)$$

with x and y in $\mathfrak{g}$ ($[\,,\,]_{\mathfrak{g}}$ is the bracket in $\mathfrak{g}$). Then, the vector space V is a $\mathfrak{g}$-module defined by

$$x.v = \Phi(x)(v), \quad x \in \mathfrak{g} \text{ and } v \in V.$$

We have naturally an equivalence between the notions of the linear representations of $\mathfrak{g}$ and the $\mathfrak{g}$-modules.

THEOREM 13 (Lie's theorem). *Let $\mathfrak{g}$ be a solvable Lie algebra over K and Φ a linear representation of $\mathfrak{g}$ in a vector space V. Then, there exists a flag $D = (V = V_0 \supset V_1 \supset \cdots \supset V_n = \{0\})$ of V such that $\Phi(\mathfrak{g}) \subset \mathfrak{b}(D)$.*

The Lie algebra $\mathfrak{b}(D)$ is described in the examples. We can also enunciate the following theorem:

THEOREM 14. *With the hypothesis of Lie's theorem, if $V \neq \{0\}$, there is a vector $v \in V$, $v \neq 0$ such that*

$$\Phi(x)(v) = \lambda(x)v \quad \textit{for all } x \in \mathfrak{g},$$

where λ is a linear form on $\mathfrak{g}$.

The equivalence between these theorems can be proved by using an induction on the dimension of V.

First we prove the following lemma.

LEMMA 15. *Let $\mathfrak{g}$ be a Lie algebra over K and $\mathfrak{i}$ an ideal of $\mathfrak{g}$. Let V be a $\mathfrak{g}$-module and $v \in V$ be a nonzero element such that $a \cdot v = \lambda(a)v$ for all $a \in \mathfrak{i}$, where λ is a given linear form on $\mathfrak{i}$. Then $\lambda([x, a]) = 0$ for all $x \in \mathfrak{g}$ and $a \in \mathfrak{i}$.*

PROOF. One chooses a vector $x \in \mathfrak{g}$. Let V_i be the subspace of V generated by the vectors

$$x \cdot v,\ x \cdot (x \cdot v),\ \ldots,\ x \cdot \bigl(x \cdots (x \cdot v)\bigr) = x^{i-1}v.$$

These subspaces define the ascending sequence

$$0 = V_0 \subset V_1 \subset \cdots \subset V_n = V.$$

Let p be the smallest integer satisfying $V_p = V_{p+1}$. The sequence (V_i) is stationary from index p and we have $\dim V_p = p$.

We will show that

$$a \cdot (x^i v) \equiv \lambda(a)x^i v (\bmod V_i)$$

for any $i \geqslant 0$ and any $a \in \mathfrak{i}$.

This relation is true for $i = 0$. Suppose that this relation is true for every i_0: $1 \leqslant i_0 < i$. Then

$$\begin{aligned} a \cdot x^i v &= a \cdot x \cdot x^{i-1} v = x \cdot a \cdot x^{i-1} v \\ &= x \cdot \lambda(a)x^{i-1}v + xv_{i-1} - \lambda[x, a]x^{i-1}v + w_{i-1} \end{aligned}$$

with v_{i-1} and w_{i-1} in V_{i-1}. As $x \cdot V_{i-1} \subset V_i$ and $\mathfrak{i} \cdot V_i \subset V_i$ we have $xv_{i-1} \in V_i$ and $\lambda[x, a]x^{i-1}v \in V_i$. Then

$$a \cdot x^i v \equiv x \cdot \lambda(a)x^{i-1}v \ (\bmod V_i) \equiv \lambda(a)x^i v \ (\bmod V_i).$$

This proves that $a \cdot x^i v \equiv \lambda(a)x^i v$ $(\bmod V_i)$. So, the matrix of endomorphism $\Phi(a)$ $(\Phi(a)(v) = a \cdot v)$ in the adapted to the flag of V basis, is triangular and all its diagonal elements are equal to $\lambda(a)$. Then $\operatorname{Tr}(\Phi(a)) = n\lambda(a)$, and as $[a, x] \in \mathfrak{i}$, we have $\operatorname{Tr}\Phi[a, x] = n\lambda[a, x]$. But

$$\operatorname{Tr}\Phi[a, x] = \operatorname{Tr}\bigl(\Phi(a) \circ \Phi(x) - \Phi(x) \circ \Phi(a)\bigr) = 0$$

and the linear form λ satisfies $\lambda[a, x] = 0$. □

Now let us return to the proof of the theorem.

One uses an induction on the dimension of $\mathfrak{g}$. If $\dim \mathfrak{g} = 0$, the theorem is true. Suppose $\dim \mathfrak{g} \neq 0$. As $\mathfrak{g}$ is solvable, we have:

$$D\mathfrak{g} = D^1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \subset \mathfrak{g} \quad \text{and} \quad D\mathfrak{g} \neq \mathfrak{g}.$$

Let I be a codimension 1 subspace of $\mathfrak{g}$, I containing $D\mathfrak{g}$. It is an ideal of $\mathfrak{g}$ because $I \supset D\mathfrak{g} \supset [I, \mathfrak{g}]$. From the hypothesis, there is a vector $v \in V$, $v \neq 0$, and a linear form λ on I such that

$$\Phi(a)(v) = \lambda(a)v \quad \text{for all } a \in I.$$

We put

$$W = \{w \in V\colon\ \Phi(a)(w) = \lambda(a)w \text{ for all } a \in I\}.$$

It is a not trivial subspace of V. By the previous lemma we have:

$$\lambda[x, a] = 0 \quad \text{for all } x \in \mathfrak{g} \text{ and } a \in I.$$

As

$$\begin{aligned} \Phi(a) \circ \Phi(x)(w) &= \Phi[a, x](w) + \Phi(x) \circ \Phi(a)(w) \\ &= \lambda[a, x](w) + \Phi(x)\big(\lambda(a)w\big) \\ &= \lambda(a)\Phi(x)(w) \end{aligned}$$

then $\Phi(x)(w) \in W$ and W is invariant for the action of $\mathfrak{g}$. Let $x \in \mathfrak{g}$, $x \notin I$. As $\Phi(x)(W) \subset W$, there is an eigenvector $v_0 \neq 0$, $v_0 \in W$. This vector is also an eigenvector for every linear operator $\Phi(y)$, $y \in Kx + I = \mathfrak{g}$. This proves the theorem.

COROLLARY 16. *Let $\mathfrak{g}$ be a solvable Lie algebra and let $\rho : \mathfrak{g} \to \operatorname{End} V$ be a linear representation, then there exists a basis of V such that the matrices of the endomorphisms $\rho(x)$, $x \in \mathfrak{g}$, in this basis are triangular.*

COROLLARY 17. *We consider the adjoint representation of $\mathfrak{g}$. Then, there is a basis of $\mathfrak{g}$ such that every adjoint operator* $\operatorname{ad} x$ *admits a matricial triangular representation with respect to this basis.*

COROLLARY 18. *The Lie algebra $\mathfrak{g}$ is solvable if and only if its derived algebra $D(\mathfrak{g})$ is nilpotent.*

PROOF. We suppose that $\mathfrak{g}$ is solvable. From the Lie theorem applied to the adjoint representation, we deduce the existence of a sequence of ideals:

$$\mathfrak{g} \supset I_1 \supset I_2 \supset \cdots \supset I_n = \{0\}.$$

If $x \in D\mathfrak{g}$, then $\operatorname{ad} x(I_j) \subset I_{j+1}$. The endomorphism $\operatorname{ad} x$ is then nilpotent. Its restriction to $D\mathfrak{g}$ is also nilpotent. From the Engel theorem, $D\mathfrak{g}$ is nilpotent. □

1.6. *Cartan criterion for a solvable Lie algebra*

In this section, our field K is a field of characteristic 0.

THEOREM 19. *Let $\mathfrak{g}$ be a subalgebra of $\mathrm{gl}(V)$ where V is a finite-dimensional space over K. Then $\mathfrak{g}$ is solvable if and only if $\mathrm{Tr}(xy) = 0$ for all $x \in \mathfrak{g}$ and for all $y \in D\mathfrak{g}$.*

The proof is based on the Jordan decomposition of an endomorphism.

COROLLARY 20. *Let $\mathfrak{g}$ be a Lie algebra such that $\mathrm{Tr}(\mathrm{ad}\, x \cdot \mathrm{ad}\, y) = 0$ for all $x \in [\mathfrak{g}, \mathfrak{g}]$ and $y \in \mathfrak{g}$. Then $\mathfrak{g}$ is solvable.*

References for this section: [7,8,21,27,46].

2. Invariants of nilpotent Lie algebras

2.1. *The dimension of characteristic ideals and the nilindex*

Let $\mathfrak{g}$ be a nilpotent Lie algebra. The dimension of $\mathfrak{g}$ is the simplest of the invariants and it is natural to try to classify the Lie algebras of a given dimension. Now, we know the classification of complex nilpotent Lie algebras of dimension $\leqslant 7$. It is clear that the elaboration of these classifications is made using a finer invariant than the dimension of $\mathfrak{g}$. Actually we have some correct classifications in the dimension $\leqslant 6$. Probably, all classifications that we know of 7-dimensional nilpotent Lie algebras contain some mistakes.

2.1.1. *The nilindex of $\mathfrak{g}$.* By definition, this is the smallest positive integer s such that $C^s\mathfrak{g} = \{0\}$. In particular, if s is the nilindex of $\mathfrak{g}$ we have $(\mathrm{ad}\, X)^s = 0$ for all X in $\mathfrak{g}$.

We can also consider the minimal integer k such that $(\mathrm{ad}\, X)^k = 0$ for all X in $\mathfrak{g}$.

EXAMPLES.
(1) If $s = 1$, then $\mathfrak{g}$ is Abelian. It is unique for a given dimension.
(2) If $s = 2$, then we say that $\mathfrak{g}$ is two-step nilpotent. The Heisenberg algebra is of this type. We note that the classification of two-step nilpotent Lie algebras is unknown for dimensions greater than 9.
(3) If $s = \dim \mathfrak{g} - 1$, we say that $\mathfrak{g}$ is filiform.

2.1.2. *The dimension of characteristic ideals.* Consider the derived series, the descending central series and the ascending central series:

$$\begin{aligned}
&\mathfrak{g} = D^0\mathfrak{g} \supset D^1\mathfrak{g} \supset \cdots \supset D^p\mathfrak{g} = \{0\},\\
&\mathfrak{g} = C^0\mathfrak{g} \supset C^1\mathfrak{g} \supset \cdots \supset C^s\mathfrak{g} = \{0\},\\
&\{0\} = C_0\mathfrak{g} \subset C_1\mathfrak{g} \subset \cdots \subset C_s\mathfrak{g} = \mathfrak{g}.
\end{aligned}$$

We put $d_i = \dim D^i\mathfrak{g}$, $c^i = \dim C^i\mathfrak{g}$, $c_i = \dim C_i\mathfrak{g}$.

The sequence of the dimensions $(n, d_1, \ldots, d_{p-1}, c^1, \ldots, c^{s-1}, c_1, \ldots, c_{s-1})$ is an invariant of $\mathfrak{g}$.

2.2. *The characteristic sequence* ($K = \mathbb{C}$)

Let $x \in \mathfrak{g}$. We denote by $c(X)$ the ordered sequence of similitude invariants of the nilpotent operator $\operatorname{ad} X$, that is the ordered sequence of the Jordan bloc dimensions of this operator. We can put the lexicographical order on the set of these sequences:

$$(c_1, c_2, \ldots, c_s) > (d_1, \ldots, d_p) \Leftrightarrow \text{there exists } i \text{ such that } c_j = d_j \text{ for } j < i \text{ and } c_i > d_i.$$

Let $c(\mathfrak{g}) = \operatorname{Max}_{x \in \mathfrak{g} \setminus D\mathfrak{g}}\{c(X)\}$.
The sequence $c(\mathfrak{g})$ is an invariant of $\mathfrak{g}$ called the *characteristic sequence*.

EXAMPLES.
(1) $c(\mathfrak{g}) = (n-1, 1)$. There is an element X in $\mathfrak{g}$ such that $(\operatorname{ad} X)^{n-2} \neq 0$, and the Jordan form of the matrix of $\operatorname{ad} X$ is:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & & \\ \vdots & & & & & 0 & 1 \\ 0 & & & & & 0 & 0 \end{pmatrix}$$

The nilindex of $\mathfrak{g}$ is equal to $n-1$ and $\mathfrak{g}$ is filiform.
(2) $c(\mathfrak{g}) = (1, 1, \ldots, 1)$. For an element X in $\mathfrak{g} \setminus D\mathfrak{g}$, the Jordan matrix of $\operatorname{ad} X$ is zero. In this case $\mathfrak{g}$ is Abelian.
(3) $c(\mathfrak{g}) = (2, 1, \ldots, 1)$. This case corresponds to the Heisenberg algebra.

2.3. *The rank of a nilpotent Lie algebra*

In the Lie algebra $\operatorname{Der} \mathfrak{g}$ of all derivations of a nilpotent Lie algebra $\mathfrak{g}$ we consider the maximal Abelian subalgebras consisting of semisimple elements. By a theorem of Mostow, these algebras are conjugate under the action of the group of inner automorphism of $\mathfrak{g}$. These Abelian algebras are sometimes called *maximal tori* of derivations.

So, the dimension of the maximal tori of $\operatorname{Der} \mathfrak{g}$ is an invariant of $\mathfrak{g}$. It is called *rank* of $\mathfrak{g}$. We note that if the rank of $\mathfrak{g}$ is maximal, it is equal to the dimension of $\mathfrak{g}/D\mathfrak{g}$.

2.4. *Other invariants*

(i) The dimensions of the cohomology spaces (with trivial coefficients, the cohomology with coefficients in $\mathfrak{g}$-module, in the adjoint module...).
(ii) The characteristic of the Lie algebra of derivations.

(iii) The dimension of the center.
(iv) The Dixmier invariant. It is the dimension of a maximal orbit for the coadjoint representation.

COMMENTS. It is very important to find the new invariants or to develop the existing ones for the different classification problems. Some invariants which are very well adapted for the low-dimensional case are not interesting for arbitrary dimensions. For example, the invariants concerning the maximal tori of derivations. We know now that the typical nilpotent Lie algebra in the case of great dimension is a characteristically nilpotent one, that is, a Lie algebra not admitting semisimple derivations. Probably the invariants concerning the characteristic sequence can be most useful for the classification problems in arbitrary dimensions.

References for this section: [7,8,16,21,28,27,50].

3. Some classes of nilpotent and solvable Lie algebras

In this section we suppose that the field K is a field of complex numbers $\mathbb{C}$.

3.1. *Filiform Lie algebras*

Let $\mathfrak{g}$ be a nilpotent Lie algebra of dimension n. Let

$$C^0\mathfrak{g} \supset C^1\mathfrak{g} \supset \cdots \supset C^{n-2}\mathfrak{g} \supset C^{n-1}\mathfrak{g} = 0$$

be the central descending series of $\mathfrak{g}$, where $C^0\mathfrak{g} = 0$, $C^i\mathfrak{g} = [\mathfrak{g}, C^{i-1}\mathfrak{g}]$, $1 \leqslant i \leqslant n-1$.

The Lie algebra $\mathfrak{g}$ is called *filiform* if $\dim C^k\mathfrak{g} = n-k-1$ for $1 \leqslant k \leqslant n-1$. We remark that the filiform Lie algebras have the maximal possible nilindex, viz. $n-1$. These algebras are the "least" nilpotent.

We know the complete classification of nilpotent Lie algebras (in particular filiform Lie algebras) of dimension $\leqslant 6$. For this reason and for convenience in this subsection we suppose that $n \geqslant 7$.

EXAMPLES.

(1) Let L_n be the $(n+1)$-dimensional Lie algebra defined by

$$[X_0, X_i] = X_{i+1}, \quad i = 1, \dots, n-1,$$

where $(X_0, X_1, \dots, X_n)$ is a basis of L_n (the undefined brackets being zero). This is, in a certain way, the simplest filiform Lie algebra.

(2) Let $n = 2k+1$ be an odd integer and let Q_n be the $(n+1)$-dimensional Lie algebra defined on the basis $(X_0, X_1, \dots, X_n)$ by

$$\begin{aligned} &[X_0, X_i] = X_{i+1}, && i = 1, \dots, n-1, \\ &[X_i, X_{n-i}] = (-1)^i X_n, && i = 1, \dots, k. \end{aligned}$$

This is a filiform Lie algebra. In the basis $(Z_0, Z_1, \dots, Z_n)$, where $Z_0 = X_0 + X_1$, $Z_i = X_i$, $i = 1, \dots, n$, this Lie algebra is defined by

$$\begin{aligned} &[Z_0, Z_i] = Z_{i+1}, && i = 1, \dots, n-2,\\ &[Z_i, Z_{n-i}] = (-1)^i Z_n, && i = 1, \dots, k. \end{aligned}$$

Consider the algebraic Lie algebra $\operatorname{Der}\mathfrak{g}$ of all derivations of $\mathfrak{g}$. Clearly, the central descending series of $\mathfrak{g}$ is an invariant flag under all derivations. As we have supposed that $\dim \mathfrak{g} \geqslant 7$, it follows that $\operatorname{Der}\mathfrak{g}$ is solvable.

Let Δ be the set of pairs of integers (k, r) such that $1 \leqslant k \leqslant n-1$, $2k+1 < r \leqslant n$, $r \geqslant 4$ (if n is odd we suppose that Δ contain also the pair $(\frac{n-1}{2}, n)$). For any element $(k, r) \in \Delta$, we can associate the 2-cocycle for the Chevalley cohomology of L_n with coefficients in the adjoint module. It is denoted $\Psi_{k,r}$ and defined by

$$\Psi_{k,r}(X_i, X_j) = -\Psi_{k,r}(X_j, X_i) = (-1)^{k-i} C_{j-k-1}^{k-i} X_{i+j+r-2k-1}$$

if $1 \leqslant i \leqslant k \leqslant j \leqslant n$, $i + j + r - 2k - 1 \leqslant n$ and $\Psi_{k,r}(X_i, X_j) = 0$ otherwise. We remark that this formula for $\Psi_{k,r}$ is uniquely determined from the conditions:

$$\begin{aligned} &\Psi_{k,r}(X_k, X_{k+1}) = X_r,\\ &\Psi_{k,r}(X_i, X_j) \in Z^2(L_n, L_n). \end{aligned}$$

Then, any $(n+1)$-dimensional filiform Lie algebra law $\mu \in F_m$ is isomorphic to $\mu_0 + \Psi$ where μ_0 is the law of L_n and Ψ is a 2-cocycle defined by

$$\Psi = \sum_{(k,r)\in\Delta} a_{k,r} \Psi_{k,r}$$

and verifying the relation $\Psi \circ \Psi = 0$ with

$$\Psi \circ \Psi(x, y, z) = \Psi\big(\Psi(x, y), z\big) + \Psi\big(\Psi(y, z), x\big) + \Psi\big(\Psi(z, x), y\big).$$

Let $\mathfrak{g}$ be a $(n+1)$-dimensional filiform Lie algebra with law μ. A basis $(X_0, X_1, \dots, X_n)$ of $\mathfrak{g}$ is called *adapted*, if

$$\mu(X_i, X_j) = \mu_0(X_i, X_j) + \Psi(X_i, X_j), \quad 0 \leqslant i, j \leqslant n.$$

3.1.1. *Filiform Lie algebras of rank 2.* We have only two types of filiform Lie algebras of rank 2: L_n and Q_n (this case is possible if $n = 2k+1$ is odd).

Case $\mathfrak{g} = L_n$.

Let $(X_0, X_1, \ldots, X_n)$ be the basis of L_n introduced in example above. We consider the endomorphisms $d_1, d_2, t, h_2, h_3, \ldots, h_{n-1}$ defined by

$$\begin{array}{lll}
d_1(X_0) = 0, & d_1(X_i) = X_i, & 1 \leqslant i \leqslant n, \\
d_2(X_0) = X_0, & d_2(X_i) = (i-1)X_i, & 1 \leqslant i \leqslant n, \\
t(X_0) = X_1, & t(X_i) = 0, & 1 \leqslant i \leqslant n, \\
h_j(X_0) = 0, & h_j(X_i) = X_{i+j}, & 2 \leqslant j \leqslant n-1,\ 1 \leqslant i \leqslant n-j.
\end{array}$$

LEMMA 21. *The endomorphisms* $(d_1, d_2, t, h_2, h_3, \ldots, h_{n-1}, \operatorname{ad} X_0, \operatorname{ad} X_1, \ldots, \operatorname{ad} X_{n-1})$ *form a basis for the Lie algebra of derivations* $\operatorname{Der} L_n$ *and a maximal torus of derivations of* L_n *is spanned by* d_1 *and* d_2.

Case $\mathfrak{g} = Q_n$, $n = 2k+1$.

Let $(Z_0, Z_1, \ldots, Z_n)$ be the basis of Q_n introduced in example above. We consider the endomorphisms $d_1, d_2, t, h_3, h_5, \ldots, h_{2k-1}$ defined by

$$\begin{array}{llll}
d_1(Z_0) = Z_0, & d_1(Z_i) = (i-1)Z_i, & 1 \leqslant i \leqslant n-1, & d_1(Z_n) = (n-2)Z_n, \\
d_2(Z_0) = 0, & d_2(Z_i) = Z_i, & 1 \leqslant i \leqslant n-1, & d_2(Z_n) = 2Z_n, \\
t(Z_0) = Z_n, & t(Z_i) = 0, & 1 \leqslant i \leqslant n, & \\
h_{2j+1}(Z_0) = 0, & h_{2j+1}(Z_i) = Z_{i+2j+1}, & 1 \leqslant j \leqslant k-1,\ 1 \leqslant i \leqslant n-2j-1. &
\end{array}$$

LEMMA 22. *The endomorphisms* $(d_1, d_2, t, h_3, h_5, \ldots, h_{2k-1}, \operatorname{ad} Z_0, \operatorname{ad} Z_1, \ldots, \operatorname{ad} Z_{n-1})$ *form a basis of the Lie algebra of derivations* $\operatorname{Der} Q_n$ *and a maximal torus of derivations of* Q_n *is spanned by* d_1 *and* d_2.

3.1.2. *Filiform Lie algebras of rank 1.* Let $\mathfrak{g}$ be a filiform Lie algebra of dimension $n+1$ and of rank 1. Then there is a basis $(Y_0, Y_1, \ldots, Y_n)$ of $\mathfrak{g}$ such that $\mathfrak{g}$ is one of the following families of Lie algebras:

$$\text{(i)}\quad \begin{array}{l}
\mathfrak{g} = A^r_{n+1}(\alpha_1, \ldots, \alpha_t), \quad 1 \leqslant r \leqslant n-3,\ t = \left[\dfrac{n-r-1}{2}\right], \\
{[Y_0, Y_i]} = Y_{i+1}, \quad 1 \leqslant i \leqslant n-1, \\
{[Y_i, Y_j]} = \left(\displaystyle\sum_{k=i}^{t} \alpha_k (-1)^{k-i} C^{k-i}_{j-k-1}\right) Y_{i+j+r}, \quad 1 \leqslant i \leqslant j \leqslant n,\ i+j+r \leqslant n.
\end{array}$$

$$\text{(ii)}\quad \begin{array}{l}
\mathfrak{g} = B^r_{n+1}(\alpha_1, \ldots, \alpha_t), \quad n = 2m+1, 1 \leqslant r \leqslant n-4,\ t = \left[\dfrac{n-r-2}{2}\right], \\
{[Y_0, Y_i]} = Y_{i+1}, \quad 1 \leqslant i \leqslant n-2, \\
{[Y_i, Y_{n-i}]} = (-1)^i Y_n, \quad 1 \leqslant i \leqslant m, \\
{[Y_i, Y_j]} = \left(\displaystyle\sum_{k=i}^{t} \alpha_k (-1)^{k-i} C^{k-i}_{j-k-1}\right) Y_{i+j+r}, \\
1 \leqslant i < j \leqslant n-1,\ i+j+r \leqslant n-1.
\end{array}$$

$$\text{(iii)}\quad \begin{array}{l} \mathfrak{g} = C_{n+1}(\alpha_1,\dots,\alpha_t), \quad n = 2m+1,\ t = m-1, \\ [Y_0, Y_i] = Y_{i+1}, \quad 1 \leqslant i \leqslant n-2, \\ [Y_i, Y_{n-i}] = (-1)^i Y_n, \quad 1 \leqslant i \leqslant m, \\ [Y_i, Y_{n-i-2k}] = (-1)^i \alpha_k Y_n, \quad 1 \leqslant k \leqslant m-1,\ 1 \leqslant i \leqslant n-2k-1, \end{array}$$

where the C_q^s are the binomial coefficients (we suppose that $C_q^s = 0$ if $q < 0$ or $q < s$), the $(\alpha_1,\dots,\alpha_t)$ are the parameters satisfying the polynomial relations emanating from the Jacobi identity and at least one parameter $\alpha_i \neq 0$. A maximal torus of derivations is spanned by d, where:

Case (i):

$$d(Y_0) = Y_0, \qquad d(Y_i) = (i+r)Y_i, \quad 1 \leqslant i \leqslant n.$$

Case (ii):

$$d(Y_0) = Y_0, \qquad d(Y_i) = (i+r)Y_i, \quad 1 \leqslant i \leqslant n-1, \qquad d(Y_n) = (n+2r)Y_n.$$

Case (iii):

$$d(Y_0) = 0, \qquad d(Y_i) = Y_i, \quad 1 \leqslant i \leqslant n-1, \qquad d(Y_n) = 2Y_n.$$

REMARK 1. Let $r = 1$. Then, up to isomorphism, there are only four Lie algebras of rank 1.

REMARK 2. The laws $C_{n+1}(\alpha_1,\dots,\alpha_t)$ satisfy the Jacobi identity for all values of parameters $(\alpha_1,\dots,\alpha_t)$.

REMARK 3. Let $\mathfrak{g}$ be a Lie algebra belonging to one of the families (i), (ii), (iii) and let at least one of parameters α_i be different to zero. Then we can transform one of these parameters to 1 using the automorphism ψ defined by $\psi(X_0) = aX_0$, $\psi(X_1) = bX_1$ (this is a unique type of automorphisms preserving the torus and the property of a basis to be adapted). Modulo this transformation we have a classification up to isomorphism of filiform Lie algebras of rank 1.

3.1.3. *Filiform Lie algebras of rank* 0. Lie algebras of this family are characteristically nilpotent filiform Lie algebras (a Lie algebra is called *characteristically nilpotent* if all of its derivations are nilpotent). Let $\mathfrak{g}$ be a filiform Lie algebra of dimension $n+1 \geqslant 6$ and of rank 0. We can associate to $\mathfrak{g}$ a graded Lie algebra, denoted $\operatorname{gr}\mathfrak{g}$, obtained from the natural filtration of nilpotent Lie algebra $\mathfrak{g}$ (filtration by the ideals of the descending central sequence). Then $\operatorname{gr}\mathfrak{g}$ is isomorphic to L_n or Q_n. If $\operatorname{gr}\mathfrak{g}$ is isomorphic to L_n then we can suppose that $\mathfrak{g}$ is defined by the relations:

$$\begin{aligned} [X_0, X_i] &= X_{i+1}, \quad 1 \leqslant i \leqslant n-1, \\ [X_i, X_j] &= \sum_{k=1}^{n-i-j} a_{ij}^k X_{i+j+k}, \quad 1 \leqslant i \leqslant j \leqslant n-2,\ i+j \leqslant n-1. \end{aligned}$$

If $\operatorname{gr}\mathfrak{g}$ is isomorphic to Q_n $(n = 2m+1)$ then we can suppose that $\mathfrak{g}$ is defined by the relations

$$\begin{aligned} [X_0, X_i] &= X_{i+1}, \quad 1 \leqslant i \leqslant n-1, \\ [X_i, X_{n-i}] &= (-1)^i X_n, \quad 1 \leqslant i \leqslant m, \\ [X_i, X_j] &= \sum_{k=1}^{n-i-j} a_{ij}^k X_{i+j+k}, \quad 1 \leqslant i \leqslant j \leqslant n-2,\ i+j \leqslant n-1, \end{aligned} \tag{1}$$

or by the relations

$$\begin{aligned} [Z_0, Z_i] &= Z_{i+1}, \quad i = 1, \ldots, n-2, \\ [Z_i, Z_{n-i}] &= (-1)^i Z_n, \quad i = 1, \ldots, m, \\ [Z_i, Z_j] &= \sum_{k=1}^{n-i-j} b_{ij}^k Z_{i+j+k}, \quad 1 \leqslant i \leqslant j \leqslant n-2,\ i+j \leqslant n-1. \end{aligned} \tag{2}$$

DEFINITION 3. Let $\mathfrak{g}$ be a filiform Lie algebra such that $\operatorname{gr}\mathfrak{g}$ is isomorphic to L_n but not isomorphic to $\mathfrak{g}$. The Lie algebra defined by

$$[X_0, X_i] = X_{i+1}, \quad 1 \leqslant i \leqslant n-1, \qquad [X_i, X_j] = a_{ij}^r X_{i+j+r},$$

where r is the smallest index, $r \geqslant 1$ such that $a_{ij}^r \neq 0$ for some (i, j), is called the *sill algebra* of $\mathfrak{g}$.

DEFINITION 4. Let $\mathfrak{g}$ be a filiform Lie algebra such that $\operatorname{gr}\mathfrak{g}$ is isomorphic to Q_n but not isomorphic to $\mathfrak{g}$ and suppose $\mathfrak{g}$ is defined by the relations (2). If $b_{ij}^{n-i-j} = 0$ for all $1 \leqslant i < j \leqslant n-2$, $i+j \leqslant n-1$, then the filiform Lie algebra defined by

$$\begin{aligned} &[Z_0, Z_i] = Z_{i+1}, && i = 1 \leqslant i \leqslant n-2, \\ &[Z_i, Z_{n-i}] = (-1)^i Z_n, && 1 \leqslant i \leqslant m, \\ &[Z_i, Z_j] = b_{ij}^r Z_{i+j+r}, && 1 \leqslant i < j \leqslant n-2, \end{aligned}$$

where r is smallest index, $r \geqslant 1$ with $b_{ij}^r \neq 0$ is called the *sill algebra of first kind* of $\mathfrak{g}$. If $b_{ij}^{n-i-j} \neq 0$ for some i and j, $1 \leqslant i < j \leqslant n-2$, then the filiform Lie algebra defined by

$$\begin{aligned} &[Z_0, Z_i] = Z_{i+1}, && i = 1 \leqslant i \leqslant n-2, \\ &[Z_i, Z_{n-i}] = (-1)^i Z_n, && 1 \leqslant i \leqslant m, \\ &[Z_i, Z_j] = b_{ij}^{n-i-j} Z_n, && 1 \leqslant i < j \leqslant n-2, \end{aligned}$$

is called *sill algebra of second kind* of $\mathfrak{g}$.

We note that if $\mathfrak{g}$ is a filiform Lie algebra non-isomorphic to Q_n but with $\operatorname{gr}\mathfrak{g} \cong Q_n$, there is only one sill algebra of $\mathfrak{g}$ (of first or second kind) non-isomorphic to Q_n.

THEOREM 23. *Let $\mathfrak{g}$ be a filiform Lie algebra of dimension $n+1 \geqslant 7$ non-isomorphic to L_n and Q_n. Then $\mathfrak{g}$ is characteristically nilpotent if and only if $\mathfrak{g}$ is not isomorphic to its sill algebra.*

COMMENTS. The role of filiform Lie algebras resides in the fact that the variety of their laws is a Zariski open set in the variety of nilpotent Lie algebra laws. Unfortunately we do not know the complete classification of the filiform Lie algebras. We can meet such Lie algebras in the study of different problems of differential geometry (for example, the problem of existence of an affine structure on a nilpotent Lie group) and progress in the study of this class of Lie algebras would be interesting.

References for this subsection: [21,22,29,30,53].

3.2. *Description of Lie algebras whose nilradical is filiform*

Let $\mathfrak{g}$ be a Lie algebra. Any maximal semisimple subalgebra $\mathfrak{s} \subset \mathfrak{g}$ is called a Levi subalgebra of $\mathfrak{g}$. The semidirect decomposition $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{r}$ holds, where $\mathfrak{r}$ is the radical in $\mathfrak{g}$ (the Levi decomposition), and all Levi subalgebras are mutually conjugate (Malcev theorem). These theorems suggest to consider the problem of the classification of Lie algebras with a fixed radical. In this subsection we study a similar problem: Description of Lie algebras with a given nilradical.

3.2.1. *Description of the decomposable Lie algebras whose nilradical is filiform.* A Lie algebra $\mathfrak{g}$ with nilradical $\mathfrak{n}$ is called *decomposable* if $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{n}$ (semidirect sum) where $\mathfrak{s}$ is a reductive Lie algebra, whose action on $\mathfrak{n}$ is completely reducible. Let $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{n}$ be a non-nilpotent decomposable Lie algebra whose nilradical $\mathfrak{n}$ is filiform. For any $a \in \mathfrak{s}$ the mapping $\operatorname{ad} a$ carries $\mathfrak{n}$ to itself, i.e. $\operatorname{ad} a|_{\mathfrak{n}}$ is a derivation of $\mathfrak{n}$. Thus, we have a Lie algebra homomorphism $\tau : \mathfrak{s} \to \operatorname{Der} \mathfrak{n}$. The structure of $\mathfrak{g}$ is uniquely defined by $\mathfrak{s}$, $\mathfrak{n}$ and τ, which can be specified arbitrarily. If the homomorphism τ is given, then we denote by $\mathfrak{g} = \mathfrak{s} \oplus_\tau \mathfrak{n}$ the corresponding semidirect sum. The brackets of this semidirect sum are defined by the following formula:

$$\bigl[(x, y), (z, t)\bigr] = \bigl([x, z], [y, t] + \tau(x)t - \tau(z)y\bigr), \quad x, z \in \mathfrak{s},\ y, t \in \mathfrak{n}.$$

We shall often identify $\mathfrak{s}$ and $\mathfrak{n}$ with the corresponding subalgebras in $\mathfrak{s} \oplus \mathfrak{n}$.

The semidirect sum $\mathfrak{g} = \mathfrak{s} \oplus_\tau \mathfrak{n}$ is called *faithful* if the homomorphism τ is a monomorphism.

As $\mathfrak{s}$ is reductive, then any semidirect sum $\mathfrak{g} = \mathfrak{s} \oplus_\tau \mathfrak{n}$ can be represented in the form of the direct sum of a faithful semidirect sum and some ideal of $\mathfrak{s}$. In what follows we suppose that $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{n}$ is a faithful semidirect sum. As the dimension of a maximal torus of $\mathfrak{n}$ is $\leqslant 2$, we have $\dim \mathfrak{s} \leqslant 2$. Case 1: $\dim \mathfrak{s} = 0$. This class of Lie algebras was described in Theorem 1. Case 2: $\dim \mathfrak{s} = 1$ or 2. From the classification of filiform Lie algebras admitting a torus of derivations (see 1.1 and 1.2) and from Malcev theorem, we have

THEOREM 24. *Let $\mathfrak{g}$ be a faithful decomposable Lie algebra with a filiform nilradical of codimension $\geqslant 1$. Then $\mathfrak{g}$ is isomorphic to the algebra $\mathfrak{g} = T \oplus_{\mathrm{id}} \mathfrak{n}$ where $\mathfrak{n}$ is one of the following: L_n, Q_n, $A^r_{n+1}(\alpha_1, \dots, \alpha_t)$, $B^r_{n+1}(\alpha_1, \dots, \alpha_t)$, $C_{n+1}(\alpha_1, \dots, \alpha_t)$ and T is a subalgebra of the maximal torus of derivations of $\mathfrak{n}$ described before. The Lie algebras $\mathfrak{g}_1 = T_1 \oplus_{\mathrm{id}} \mathfrak{n}$, $\mathfrak{g}_2 = T_2 \oplus_{\mathrm{id}} \mathfrak{n}$ are isomorphic if and only if $T_1 = T_2$.*

Let $\mathfrak{g}$ be a non-decomposable Lie algebra whose nilradical $\mathfrak{n}$ is filiform and let $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{m}$ be a Levi decomposition, where $\mathfrak{m}$ is the radical of $\mathfrak{g}$ and $\mathfrak{p}$ is a Levi subalgebra of $\mathfrak{g}$. We can also suppose that this is a faithful semidirect sum. In this case Lie algebra $\mathfrak{g}$ is also called *faithful*. We can write $\mathfrak{m} = \mathfrak{s} \oplus \mathfrak{n}$ (vectorial direct sum) where $\mathfrak{n}$ is the nilradical of $\mathfrak{g}$ and $\mathfrak{s}$ is a complementary subspace.

LEMMA 25. *Lie algebra $\mathfrak{g}$ is solvable.*

PROOF. Since $\mathfrak{p}$ is semi-simple, we may suppose that $[\mathfrak{p}, \mathfrak{s}] \in \mathfrak{s}$. As $[\mathfrak{p}, \mathfrak{s}] \subset \mathfrak{n}$, we have $[\mathfrak{p}, \mathfrak{s}] = 0$. It follows that, for a faithful $\mathfrak{g}$, the adjoint action of $\mathfrak{p}$ in $\mathfrak{n}$ is also faithful. Since $\operatorname{Der} \mathfrak{n}$ is solvable, $\mathfrak{p} = 0$. Thus, $\mathfrak{g}$ is solvable, whenever it is faithful. □

LEMMA 26. $\dim \mathfrak{s} \leqslant \operatorname{rank}(\mathfrak{n}) \leqslant 2$.

We consider the Chevalley decomposition of the derivations algebra of $\mathfrak{n}$:

$$\operatorname{Der} \mathfrak{n} = \mathfrak{t} \oplus \mathfrak{b},$$

where $\mathfrak{t}$ is a maximal torus and $\mathfrak{b}$ the ideal of all nilpotent elements. For any $s \in \mathfrak{s}$, consider the decomposition $\operatorname{ad} s = t + b$, where $t \in \mathfrak{t}$ and $b \in \mathfrak{b}$. Then the linear mapping $\mathfrak{s} \to \mathfrak{t}$ defined by $s \to t$ is injective. In fact, if $\mathfrak{t} = 0$, then $\operatorname{ad} s$ is nilpotent on $\mathfrak{n}$. But $[s, \mathfrak{s}] \subset \mathfrak{n}$. It follows that s is a nilpotent element of $\mathfrak{g}$. Hence $s = 0$. Thus $\dim \mathfrak{s} \leqslant \dim \mathfrak{t}$.

COROLLARY 27. *If the exists a non-decomposable faithful Lie algebra $\mathfrak{g}$ with $\dim \mathfrak{s} = 2$, then $\mathfrak{n} = L_n$ or Q_n.*

LEMMA 28. *Let $\mathfrak{g}$ be a faithful Lie algebra (decomposable or not) with a filiform nilradical $\mathfrak{n}$ and let $\dim \mathfrak{g} - \dim \mathfrak{n} = 2$. Then we may choose a complementary subspace $\mathfrak{s}$ to $\mathfrak{n}$ in $\mathfrak{g}$ such that $\operatorname{ad} \mathfrak{s}$ is a maximal torus of derivations of $\mathfrak{g}$.*

PROOF. We study the two following possible cases:

(1) $\mathfrak{g} = \mathfrak{s} \oplus L_n$,
(2) $\mathfrak{g} = \mathfrak{s} \oplus Q_n$.

Case (1): Let (s_1, s_2) be a basis of $\mathfrak{s}$. Consider the Jordan decomposition $\operatorname{ad} s_1 = c_1 + n_1$, where c_1 is an element of the maximal torus of derivations T of L_n described in Section 1, n_1 is a nilpotent derivation with $c_1 n_1 = n_1 c_1$. We consider also the Chevalley decomposition $\operatorname{ad} s_2 = c_2 + n_2$ with $c_2 \in T$, and n_2 is a nilpotent derivation of L_n. We may suppose that n_1 and n_2 are two exterior derivations of L_n which belong to the space spanned by $S - \{d_1, d_2\}$, where S is the basis of $\operatorname{Der} L_n$ described in the Section 1. Two cases are possible:

$$\text{(i)}\quad \begin{aligned} &c_1 = \mu d_1, \quad n_1 = \sum_{i=2}^{n-1} \beta_i h_i, \quad (\beta_2, \dots, \beta_{n-1}) \in \mathbb{C}^{n-2} - \{0\},\\ &c_2 = \alpha_1 d_1 + \alpha_2 d_2, \quad n_2 = \nu t + \sum_{i=2}^{n-1} \gamma_i h_i, \quad (\gamma_2, \dots, \gamma_{n-1}) \in \mathbb{C}^{n-2} - \{0\}, \end{aligned}$$

where $\mu \neq 0$, $\alpha_2 \neq 0$, t, h_i are the nilpotent derivations of L_n described in Section 1.

$$\text{(ii)}\quad \begin{aligned} &c_1 = \mu (d_1, d_2), \quad n_1 = \beta t,\\ &c_2 = \alpha_1 d_1 + \alpha_2 d_2, \quad n_2 = \nu t + \sum_{i=2}^{n-1} \gamma_i h_i, \quad (\gamma_2, \dots, \gamma_{n-1}) \in \mathbb{C}^{n-2} - \{0\}, \end{aligned}$$

where $\mu \neq 0$, $\alpha_1 \neq \alpha_2$. The endomorphism $f = [\operatorname{ad} s_1, \operatorname{ad} s_2] = \operatorname{ad}[s_1, s_2]$ in an interior derivation. In the case (i) we have

$$\begin{aligned} f(X_0) &= \mu\nu X_1 + \sum_{i \geqslant 2} \tau_i X_i,\\ f(X_p) &= \alpha_2 \sum_{2 \leqslant i \leqslant n-1} \beta_i X_{i+p}, \quad p = 1, \dots, n-2. \end{aligned}$$

These relations are only possible, if $\nu = \beta_2 = \cdots = \beta_{n-1} = 0$. Thus, $f = 0$ and the endomorphisms $\operatorname{ad} s_1, \operatorname{ad} s_2$ belong in a same torus. In the case (ii) we have also $\nu = \gamma_2 = \cdots = \gamma_{n-1} = 0$ and $[\operatorname{ad} s_1, \operatorname{ad} s_2] = \operatorname{ad}[s_1, s_2] = 0$.

Case (2): Analogously to case (1), by using the description of $\operatorname{Der} Q_n$ given in Section 1, we have $\operatorname{ad} s_1 \operatorname{ad} s_2 = \operatorname{ad} s_2 \operatorname{ad} s_1$. □

LEMMA 29. *Let* $\mathfrak{g}$ *be a faithful non-decomposable Lie algebra with a filiform nilradical* $\mathfrak{n}$. *Then* $\dim \mathfrak{g} - \dim \mathfrak{n} = 1$.

PROOF. Let $\mathfrak{s}$ be a subspace complementary to $\mathfrak{n}$ in $\mathfrak{g}$. We suppose $\dim \mathfrak{s} \geqslant 2$. Corollary 27 and Lemma 28 show that $\mathfrak{g}$ is isomorphic to one of the following Lie algebras:

$$\begin{aligned} \mathfrak{g}_1 &= (\mathbb{C} s_1 \oplus \mathbb{C} s_2) \oplus L_n,\\ \mathfrak{g}_2 &= (\mathbb{C} c_1 \oplus \mathbb{C} c_2) \oplus Q_n, \end{aligned}$$

where $[s_i, x] = d_i(x)$, $x \in L_n$, $i = 1, 2$, $[c_i, z] = d_i(z) = d_i(z)$, $z \in Q_n$, $i = 1, 2$, $[s_1, s_2] = \beta X_n$, $[c_1, c_2] = \gamma Z_n$, $(X_0, X_1, \dots, X_n)$ is the basis of L_n, $(Z_0, Z_1, \dots, Z_n)$ is the basis of Q_n, d_1 and d_2 are the elements of maximal torus T of derivation of $\mathfrak{n}$, described in the Section 1. The mappings

$$\psi : T \oplus_{\mathrm{id}} L_n \to \mathfrak{g}_1, \quad \eta : T \oplus_{\mathrm{id}} Q_n \to \mathfrak{g}_2$$

defined by $\psi(d_1) = s_1$, $\psi(d_2) = s_2 - \beta X_n$, $\psi(x) = x$, $x \in L_n$, and by $\eta(d_1) = c_1 + \gamma Z_n$, $\eta(d_2) = c_2$, $\psi(z) = z$, $z \in Q_n$, are the isomorphisms. Thus, $\mathfrak{g}$ is decomposable. □

Case $\dim \mathfrak{s} = 1$. Let f be a nonzero element of $\mathfrak{s}$ considered as a derivation of $\mathfrak{n}$ and let $f = f_s + f_n$ be the Jordan decomposition of f with f_s semisimple, f_n nilpotent, $[f_s, f_n] = 0$. We can suppose that $f_s \in T$, where T is the maximal torus of derivations of $\mathfrak{n}$ described before. As $f_n \neq 0$, f_s admits an eigenvalue with a multiplicity greater than 2. This implies that $\mathfrak{n} = L_n$ or Q_n or $\mathbf{C}_{n+1}(\alpha_1, \ldots, \alpha_t)$. Moreover, f_n should be an external derivation (that is $f_n \neq \operatorname{ad} X$ for any $X \in \mathfrak{n}$). We deduce the following description of non-decomposable faithful Lie algebras with a filiform nilradical.

THEOREM 30. *Let $\mathfrak{g}$ be a non-decomposable faithful Lie algebra with a filiform nilradical $\mathfrak{n}$. Then $\mathfrak{g} = \{\mathbb{C}.f\} \oplus \mathfrak{n}$, where $f = f_s + f_n$, $[f, X] = f(X)$ for all $X \in \mathfrak{n}$, and $\mathfrak{n}$, f_s, f_n are one of the following*:

(i) $\mathfrak{n} = L_n$, $f_s = d_1$, $f_n = \sum_{i=2}^{n-1} \beta_i h_i$, $(\beta_2, \ldots, \beta_{n-1}) \in \mathbb{C}^{n-2} - \{0\}$ *where h_i are the nilpotent derivations defined by* $h_i(X_0) = 0$, $h_i(X_j) = X_{i+j}$, $j = 1, \ldots, n-i$.

(ii) $\mathfrak{n} = L_n$, $f_s = d_1 + d_2$ *and f_n is defined by* $f_n(X_0) = X_1$ *and* $f_n(X_i) = 0$ *for* $i \neq 0$.

(iii) $\mathfrak{n} = Q_n (n = 2m+1)$, $f_s = d_2$, $f_n = \sum_{i=1}^{m-3} \beta_i h_{2i+1}$, $(\beta_1, \ldots, \beta_{m-1}) \in \mathbb{C}^{m-1} - \{0\}$, *where h_i is the nilpotent derivation of $\mathfrak{n}$ defined by* $h_{2i+1}(Z_0) = 0$, $h_{2i+1}(Z_j) = Z_{2i+j+1}$, $j = 1, \ldots, n-2i-2$.

(iv) $\mathfrak{n} = C_{n+1}(\alpha_1, \ldots, \alpha_t)$ $(n = 2m+1)$, $f_s = d$, $f_n = \sum_{i=1}^{m-1} \beta_i h_{2i+1}$, $(\beta_1, \ldots, \beta_{m-1}) \in \mathbb{C}^{m-1} - \{0\}$, *where* $h_{2i+1}(Y_0) = 0$, $h_{2i+1}(Y_j) = Z_{2i+j+1}$, $j = 1, \ldots, n-2i-2$.

We suppose here that d_1 and d_2 are the elements of a maximal torus of derivations of $\mathfrak{n}$, $(X_0, X_1, \ldots, X_n)$ is the basis of L_n, $(Z_0, Z_1, \ldots, Z_n)$ is the basis of Q_n, $(Y_0, Y_1, \ldots, Y_n)$ is the basis of $\mathbf{C}_{n+1}(\alpha_1, \ldots, \alpha_t)$, described in the Section 1.

REMARK 4. Analogously to Remark 3, we can transform one of the parameters β_i to 1 using the automorphism ψ. Modulo this transformation theorem 30 gives a classification up to isomorphism of non-decomposable faithful Lie algebras with a filiform nilradical.

References for this subsection: [22,23].

3.3. *Two-step nilpotent Lie algebras*

DEFINITION 5. A nilpotent Lie algebra $\mathfrak{g}$ is called two step nilpotent if $C^2(\mathfrak{g}) = \{0\}$.

EXAMPLES.

(1) The Abelian Lie algebra is two-step nilpotent.

(2) The *Heisenberg algebra* H_p of dimension $2p+1$ is defined by the brackets

$$[X_1, X_2] = [X_3, X_4] = \cdots = [X_{2p-1}, X_{2p}] = X_{2p} + 1.$$

The Heisenberg algebra is a model for truly two-step nilpotent Lie algebras. It is characterized by the following property:

PROPOSITION 31. *Every Lie algebra satisfying $Z(\mathfrak{g}) = C^1\mathfrak{g}$ and $\dim Z(\mathfrak{g}) = 1$ is isomorphic to a Heisenberg algebra.*

3.3.1. *On the structure of the two step nilpotent Lie algebras.* Let V be a vector space complementary of $C^1\mathfrak{g}$ in $\mathfrak{g}$: $\mathfrak{g} = C^1\mathfrak{g} \oplus V$.

Note that $s = \dim V = \dim \mathfrak{g}/C^1\mathfrak{g}$ is the minimal number of generators of $\mathfrak{g}$.

As $\mathfrak{g}$ is 2-step nilpotent, the derived algebra $C^1\mathfrak{g}$ is contained in the center $Z(\mathfrak{g})$. Denote by U the subspace $Z(\mathfrak{g}) \cap V$. It is an Abelian ideal of $\mathfrak{g}$. Consider a vector space W complementary to U in V. We have

$$\mathfrak{g} = C^1\mathfrak{g} \oplus U \oplus W \quad \text{and} \quad Z(\mathfrak{g}) = C^1\mathfrak{g} \oplus U.$$

We can deduce that $\mathfrak{k} = C^1\mathfrak{g} \oplus W$ is a two-step nilpotent subalgebra of $\mathfrak{g}$ satisfying $Z(\mathfrak{k}) = C^1\mathfrak{g}$ and $C^1\mathfrak{k}=C^1\mathfrak{g}$. Thus every 2-step nilpotent Lie algebra is a trivial extension of a 2-step nilpotent ones with the center equal to the derived algebra, by an Abelian ideal. So one can suppose that the algebra $\mathfrak{g}$ satisfies the condition $C^1\mathfrak{g} = Z(\mathfrak{g})$.

Consider $(X_1, \dots, X_p, Y_1, \dots, Y_{n-p})$ a basis of $\mathfrak{g}$ adapted to the decomposition:

$$\mathfrak{g} = C^1\mathfrak{g} \oplus V = Z(\mathfrak{g}) \oplus V \quad \text{with } \dim C^1\mathfrak{g} = p.$$

Obviously, we have $[X_i, X_j] = [X_i, Y_k] = 0$. Then the Lie algebra $\mathfrak{g}$ is entirely defined by the only brackets:

$$[Y_i, Y_j] = \sum_{k=1}^{p} a_{ij}^k X_k, \quad 1 \leqslant i < j \leqslant n - p.$$

The Jacobi identities are satisfied and the constants a_{ij}^k are free. Intrinsically this means that the structure of $\mathfrak{g}$ is defined by surjective linear map

$$\varphi : \Lambda^2 V \to Z(\mathfrak{g}).$$

The dual point of view permits a concrete analysis. The dual map

$$\varphi^* : Z(\mathfrak{g})^* \to \Lambda^2(U)^*$$

is injective. The data of this mapping (and then the data of the bracket of $\mathfrak{g}$) is the same that the data of a subspace of dimension $p = \dim Z(\mathfrak{g})$ in $\Lambda^2 V^*$. If $(\alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_{n-p})$ is the dual basis of the given basis $(X_1, \dots, X_p, Y_1, \dots, Y_{n-p})$, then the structural equations of $\mathfrak{g}$ (the dual Jacobi conditions) are:

$$d\alpha_k = \sum_{1 \leqslant i < j \leqslant n-p} a_{ij}^k \beta_i \wedge \beta_j.$$

The bilinear forms $(\theta_1 = d\alpha_1, \dots, \theta_k = d\alpha_k)$ are 2-forms on V. They generate the subspace $\varphi^*(Z(\mathfrak{g})^*)$. It appears clearly, in this notation, that the data of the 2-forms θ_i is equivalent to the data of the constants (a_{ij}^k).

3.3.2. *On the classification of two-step nilpotent Lie algebras.* At first sight, the problem of classification of two-step nilpotent Lie algebras appears easy to solve. Recall that the problem of classification of nilpotent Lie algebras, has been approached for the first time by Umlauf in his thesis in 1891.

Today it is still an open problem. If we introduce a new and strong hypothesis like two-step nilpotency, we can hope a simplification to this problem. However, this is not the case. From the previous results, we can claim:

PROPOSITION 32. *The classification up to isomorphism of* 2*-step nilpotent Lie algebras is equivalent with the classification of the orbits corresponding to the action of the linear group* $\mathrm{GL}(p, \mathbb{C})$ *in the Grassmannian of the* 2*-spaces in the vector space* $\Lambda^2(\mathbb{C}^p)$.

Now we consider the characteristic sequence of a 2-step nilpotent Lie algebra $\mathfrak{g}$. As the nilindex of $\mathfrak{g}$ is 2 (if $\mathfrak{g}$ is not Abelian) then the characteristic sequence of $\mathfrak{g}$ has the following form:

$$s(\mathfrak{g}) = (2, 2, 2, \ldots, 2, 1, 1, \ldots, 1).$$

Consider the class of 2-step nilpotent Lie algebras whose characteristic sequence is $(2, 2, \ldots, 2, 1)$. This class determines an open set in the algebraic set of 2-step nilpotent Lie algebras. Suppose that $\dim \mathfrak{g} = 2p + 1$. In this case $\dim Z(\mathfrak{g}) = p$ and the brackets of $\mathfrak{g}$ are

$$[Y_i, Y_j] = \sum_{k=1}^{p} a_{ij}^k X_k, \quad 1 \leqslant i < j \leqslant p + 1,$$

where $(X_1, \ldots, X_p, Y_1, \ldots, Y_{p-1})$ is a basis of $\mathfrak{g}$ and $(X_1, \ldots, X_p)$ a basis of $Z(\mathfrak{g})$. We can verify that the parameters a_{ij}^k are free (all the Jacobi conditions are satisfied). Then this class of algebras is parametrized by the α_{ij}^k and is isomorphic to the vector space $\mathrm{Alt}(\mathbb{C}^p)$ of the alternating bilinear form on $\mathbb{C}^p$ with values in $\mathbb{C}^p$. The classification of these 2-step nilpotent Lie algebras is equivalent to the classification of the bilinear form with values in $\mathbb{C}^P$ and contains the classification of all the Lie algebras of dimension p. For this, we claim that this classification is impossible to establish.

References for this subsection: [21].

3.4. *Characteristically nilpotent Lie algebras*

Here we give some basic properties of characteristically nilpotent Lie algebras. For the more profound results see the chapter "Varieties of Lie algebra laws" of this Handbook (this Volume 2) and for characteristically nilpotent filiform Lie algebras see also Section 3.1 above.

We have the following theorem by Jacobson:

THEOREM 33. *Every Lie algebra over a field of characteristic* 0 *having a nondegenerate derivation is nilpotent.*

COMMENTS. A derivation f on a Lie algebra $\mathfrak{g}$ is defined by:

$$\delta f(X,Y)=\bigl[f(X),Y\bigr]+\bigl[X,f(Y)\bigr]-f[X,Y]=0$$

(the operator δ is related to Chevalley cohomology). This derivation is nondegenerate if the linear map f is not degenerate. For example, if $\mathfrak{g}$ is semisimple, every derivation is inner, that is it can be written as $f=\operatorname{ad} X$ for some X in $\mathfrak{g}$. Such a derivation admits a non-trivial kernel and every derivation in a semi simple Lie algebra is degenerate.

In his work, Jacobson presents the following problem: does every nilpotent Lie algebra admit a nondegenerate derivation?

The answer to this problem is negative. The first example of a nilpotent algebra all of whose derivations are degenerate, was given by Dixmier and Lister in 1957, [13]. It is the following 8-dimensional Lie algebra whose brackets are

$$\begin{array}{ll}
[X_1,X_2]=X_5, & [X_2,X_3]=X_8,\\
[X_1,X_3]=X_6, & [X_2,X_4]=X_6,\\
[X_1,X_4]=X_7, & [X_2,X_6]=-X_7,\\
[X_1,X_5]=-X_8, & [X_3,X_4]=-X_5,\\
 & [X_3,X_5]=-X_7,\\
 & [X_4,X_6]=-X_8.
\end{array}$$

This example has been the starting point for studies of a new class of Lie algebras, named the *characteristically nilpotent* Lie algebras.

DEFINITION 6. Let $\mathfrak{g}$ be a nilpotent Lie algebra and $\operatorname{Der}\mathfrak{g}$ its derivations algebra. The Lie algebra $\mathfrak{g}$ is called *characteristically nilpotent* if every derivation f in $\operatorname{Der}\mathfrak{g}$ is a nilpotent endomorphism.

3.4.1. *Characterization of characteristically nilpotent Lie algebras.* Define

$$\mathfrak{g}^{[1]}=\operatorname{Der}\mathfrak{g}(\mathfrak{g})=\bigl\{Y\in\mathfrak{g}\colon\ \text{such that } Y=f(X),\ f\in\operatorname{Der}\mathfrak{g},\ X\in\mathfrak{g}\bigr\}$$

and

$$\mathfrak{g}^{[i]}=\operatorname{Der}\mathfrak{g}\bigl(\mathfrak{g}^{[i-1]}\bigr),\quad i>1.$$

The Lie algebra $\mathfrak{g}$ is characteristically nilpotent if and only if there is $k\in N$ such that $\mathfrak{g}^{[k]}=\{0\}$.

This sequence $\mathfrak{g}^{[i]}$ generalizes the central descending sequence; here we use the set of all derivations instead of just the set of inner derivations.

THEOREM 34. *Let* $\mathfrak{g}$ *be a nilpotent Lie algebra of dimension more than* 2. *Then* $\mathfrak{g}$ *is characteristically nilpotent if and only if the Lie algebra* $\operatorname{Der}\mathfrak{g}$ *is a nilpotent Lie algebra.*

Note that the nilpotency of $\operatorname{Der}\mathfrak{g}$ does not imply directly the nilpotency of its elements, but only the nilpotency of the adjoint operators. The proof of the theorem is based on the following lemma.

LEMMA 35. *Let L be a nilpotent Lie algebra such that $L = \mathfrak{l}_1 \oplus \mathfrak{l}_2$ where $\mathfrak{l}_1$ and $\mathfrak{l}_2$ are ideals, and $\mathfrak{l}_1$ is a central ideal. Then* $\operatorname{Der} L$ *is not nilpotent.*

PROOF. As l_1 is central, one has $[\mathfrak{l}_1, L] = \{0\}$. Let $x \neq 0$ be in $\mathfrak{l}_1$, and U such that $\mathfrak{l}_1 = U \oplus \mathbb{C}\{x\}$. As $\mathfrak{l}_2$ is an ideal of L, it is nilpotent. Then its center is not trivial. Let be $y \neq 0$ in $Z(\mathfrak{l}_2)$. Let f_1 and f_2 be the derivations of L defined by:

$$\begin{array}{lll} f_1(\mathfrak{l}_2) = 0, & f_1(x) = y, & f_1(U) = 0, \\ f_2(\mathfrak{l}_2) = 0, & f_2(x) = x, & f_2(U) = 0. \end{array}$$

Then, we have $[f_1, f_2] = f_1$ and $\operatorname{Der}\mathfrak{g}$ is not nilpotent. □

PROOF (OF THE THEOREM). It is obvious that if $\mathfrak{g}$ is characteristically nilpotent then $\operatorname{Der}\mathfrak{g}$ is nilpotent. Now we prove the reverse. Suppose $\operatorname{Der}\mathfrak{g}$ nilpotent and $\dim\mathfrak{g} > 1$. Let α be in the dual space of $\operatorname{Der}\mathfrak{g}$ and let be V_α the subspace of $X \in \mathfrak{g}$ such that there exists $m \in N$ with $(f - \alpha(f)\,\mathrm{Id})^m(X) = 0$ for all $f \in \operatorname{Der}\mathfrak{g}$. If $V_\alpha \neq \{0\}$, we call α a weight of $\operatorname{Der}\mathfrak{g}$. Denote by Ω the set of the weights of $\operatorname{Der}\mathfrak{g}$. One has:

$$\mathfrak{g} = \bigoplus_{\alpha \in \Omega} V_\alpha$$

and $[V_\alpha, V_\beta] \subset V_{\alpha+\beta}$ if $\alpha + \beta \in \Omega$ or $[V_\alpha, V_\beta] = \{0\}$ if not. As every $\operatorname{ad} X$ is a derivation of $\mathfrak{g}$, we deduce that each space V_α is an ideal of $\mathfrak{g}$. One deduces:

$$[V_\alpha, V_\beta] \subset V_\alpha \cap V_\beta \cap V_{\alpha+\beta}.$$

Then, $[V_\alpha, V_\beta] = \{0\}$ when $\alpha \neq 0$ or $\beta \neq 0$.

We consider the spaces

$$I_1 = \bigoplus_{\alpha \in \Omega,\ \alpha \neq 0} V_\alpha \quad \text{and} \quad I_2 = V_0.$$

Then $\mathfrak{g}$ is a direct vectorial sum of I_1 and I_2, I_1 and I_2 are ideals of $\mathfrak{g}$; I_1 being central, the previous lemma implies $I_1 = \{0\}$ or $I_2 = \{0\}$.

If $I_1 = \{0\}$, then $\mathfrak{g} = I_2 = V_0$ and every derivation of $\mathfrak{g}$ is nilpotent.

If $I_2 = \{0\}$, then $\mathfrak{g} = I_1$. This implies that $\mathfrak{g}$ is Abelian and $\operatorname{Der}\mathfrak{g} = \mathrm{gl}(n, \mathbb{C})$. As $\operatorname{Der}\mathfrak{g}$ is nilpotent, we deduce that $n = 1$ and we find a contradiction. □

EXAMPLES.

(i) The smallest dimension where we meet an example of a complex characteristically nilpotent Lie algebra is 7. This example has been given by Favre:

$$
\begin{array}{ll}
[X_1, X_2] = X_3, & [X_2, X_3] = X_6, \\
[X_1, X_3] = X_4, & [X_2, X_4] = [X_2, X_5] = -[X_3, X_4] = X_7, \\
[X_1, X_4] = X_5, & \\
[X_1, X_5] = X_6, & [X_1, X_6] = X_7.
\end{array}
$$

Note that, in dimension 7, there is a one parameter family of nonisomorphic characteristically nilpotent Lie algebras. In contrast to the Favre example, these algebras are not filiform.

(ii) In the first chapter of the Bourbaki book, [7], the following Lie algebra is given:

$$
\begin{array}{lll}
[X_1, X_2] = X_3, & [X_2, X_3] = X_5, & [X_3, X_4] = -X_7 + X_8, \\
[X_1, X_3] = X_4, & [X_2, X_4] = X_6, & [X_3, X_5] = X_8. \\
[X_1, X_4] = X_5, & [X_2, X_5] = X_7, & \\
[X_1, X_5] = X_6, & [X_2, X_6] = 2X_8, & \\
[X_1, X_6] = X_8, & & \\
[X_1, X_7] = X_8, & &
\end{array}
$$

This algebra is characteristically nilpotent and not filiform.

3.4.2. *Direct sum of characteristically nilpotent Lie algebras*

THEOREM 36. *Let $\mathfrak{g} = \bigoplus_{i=1}^{p} \mathfrak{g}_i$ be a Lie algebra, where the $\mathfrak{g}_i$ are ideals of $\mathfrak{g}$. Then $\mathfrak{g}$ is characteristically nilpotent if and only if each ideal $\mathfrak{g}_i$ is characteristically nilpotent.*

PROOF. We have seen, in a previous section that

$$
\operatorname{Der}\left(\sum_{i=1}^{p} \mathfrak{g}_i\right) = \left(\sum_{i=1}^{p} \operatorname{Der} \mathfrak{g}_i\right) \oplus \left(\sum_{i \neq j} D(\mathfrak{g}_i, \mathfrak{g}_j)\right),
$$

where

$$
D(\mathfrak{g}_i, \mathfrak{g}_j) = \left\{ f \in \operatorname{End} \mathfrak{g}\colon\ f(\mathfrak{g}_k) = 0 \text{ if } k \neq i,\ f(\mathfrak{g}_i) \subset Z(\mathfrak{g}_i) \text{ and } f[\mathfrak{g}_i, \mathfrak{g}_j] = 0 \right\}.
$$

We suppose that each $\mathfrak{g}_i$ is characteristically nilpotent. In this case, we have $Z(\mathfrak{g}_i) \subset [\mathfrak{g}_i, \mathfrak{g}_i]$. Indeed, if it is not the case, there is $X \neq 0$, $X \in Z(\mathfrak{g})$ with $X \notin [\mathfrak{g}_i, \mathfrak{g}_i]$. If U is a complementary subspace of the subspace generated by X in $\mathfrak{g}_i$, we define a derivation f by $f(X) = X$ and $f(U) = 0$. This derivation is not nilpotent. So, the center $Z(\mathfrak{g}_i)$ of $\mathfrak{g}_i$ is contained in $[\mathfrak{g}_i, \mathfrak{g}_i]$. One deduces:

(i) if $f_i \in \operatorname{Der}(\mathfrak{g}_i)$ and $f_i \in \operatorname{Der}(\mathfrak{g}_i)$ with $i \neq j$, then $f_i \circ f_j = f_j \circ f_i$.

(ii) $f_1 \circ f_2 = 0$ for all $f_1 \in D(\mathfrak{g}_i, \mathfrak{g}_j)$ and $f_2 \in D(\mathfrak{g}_k, \mathfrak{g}_l)$.

As we also have $f(Z(\mathfrak{g})) \subset Z(\mathfrak{g})$ for all $f \in \operatorname{Der}\mathfrak{g}$, we deduce that every derivation is nilpotent.

Conversely, assume that $\mathfrak{g} = \bigoplus_{i=1}^{p} \mathfrak{g}_i$ is characteristically nilpotent.

Any derivation of $\mathfrak{g}_i$ can be naturally extended in a derivation of $\mathfrak{g}$. Then this derivation is nilpotent. □

COMMENTS. The researches in this direction made recently have completely changed the point of view on nilpotent Lie algebras. We know now that the typical nilpotent Lie algebra is a characteristically nilpotent one. It is very interesting to study this class in the nonfiliform case.

References for this subsection: [9,11,13,14,21,27,29,30,32,35–37,56].

3.5. *k-Abelian filiform Lie algebras*

3.5.1. *Generalities.* A m-dimensional nilpotent Lie algebra $\mathfrak{g}$ (or corresponding law μ) is called *k-Abelian* ($0 \leqslant k \leqslant m-2$) if the ideal $C^k\mathfrak{g}$ of the central descending sequence of $\mathfrak{g}$ is Abelian. The minimal integer t, such that $\mathfrak{g}$ is t-Abelian, is called the *commutativity index* of $\mathfrak{g}$.

PROPOSITION 37. *Every $(n+1)$-dimensional k-Abelian filiform Lie algebra law $\mu \in \mathcal{F}^{n+1}(k)$ is isomorphic to*

$$\mu_0 + \sum_{(i,r)\in\Delta_k} a_{i,r}\psi_{i,r},$$

where μ_0 is the law of L_n and Δ_k is the subset of pairs (i,r) in Δ with $i \leqslant k$.

This proposition is a direct consequence from the results of 3.1 and the definition of a k-Abelian filiform Lie algebra.

COROLLARY 38. *All the brackets of an $(n+1)$-dimensional k-Abelian filiform Lie algebra relative to an adapted basis $(X_0, X_i, \dots, X_n)$ are uniquely determined by the brackets*

$$\begin{aligned}
&[X_1, X_2] = a_{1,0}X_4 + a_{1,1}X_5 + \cdots + a_{1,n-4}X_n,\\
&[X_2, X_3] = a_{2,0}X_6 + a_{2,1}X_7 + \cdots + a_{2,n-6}X_n,\\
&[X_k, X_{k+1}] = a_{k,0}X_{2k+2} + a_{k,1}X_{2k+3} + \cdots + a_{k,2k-2}X_n
\end{aligned}$$

if $2k+1 < n$, and by these brackets and a supplementary bracket

$$[X_{\frac{n-1}{2}}, X_{\frac{n+1}{2}}] = a_{\frac{n-1}{2},n}X_n$$

if $n = 2k+1$.

COROLLARY 39. *The commutativity index of an m-dimensional filiform Lie algebra is* $\leqslant [\frac{m-1}{2}]$.

In the following subsection we give the classification up to isomorphism of 2-Abelian filiform Lie algebras.

3.5.2. *Classification of the 2-Abelian filiform Lie algebras*

THEOREM 40. *Every $(n+1)$-dimensional $(n \geqslant 6)$ 2-Abelian filiform Lie algebra law is isomorphic to one of the following laws.*

$$\begin{aligned}(1)\ \mu_{1,s,t}(\Omega) &= \mu_0 + \Psi_{1,s} + \lambda_1\Psi_{1,s+1} + \cdots + \widehat{\lambda_{s-3}\Psi_{1,2s-3}} + \cdots \\ &\quad + \lambda_{t-s-1}\Psi_{1,t-1} + \Psi_{2,t} + \theta_1\Psi_{2,t+1} + \cdots + \theta_{n-t}\Psi_{2,n},\end{aligned}$$

where

$$4 \leqslant s < \frac{t+3}{2}, \quad t > \frac{n+4}{2}, \quad n > 7,$$
$$\Omega = (\lambda_1, \ldots, \lambda_{s-4}, \lambda_{s-2}, \ldots, \lambda_{t-s-1}, \theta_1, \ldots, \theta_{n-t}) \in \mathbb{C}^{n-s-2}.$$

$$\begin{aligned}(2)\ \mu_{2,s,t}(\Omega) &= \mu_0 + \Psi_{1,s} + \lambda_1\Psi_{1,s+1} + \cdots + \lambda_{t-s-1}\Psi_{1,t-1} + \Psi_{2,t} \\ &\quad + \theta_1\Psi_{2,t+1} + \theta_2\Psi_{2,t+2} + \cdots + \theta_{n-t}\Psi_{2,n},\end{aligned}$$

where

$$\frac{t+3}{2} \leqslant s < t, \quad t > \frac{n+4}{2}, \quad s+t-3 > n, \quad s < k, \quad t \neq s+2, \quad n \neq 2s-7,$$
$$\Omega = (\lambda_1, \lambda_2, \ldots, \lambda_{t-s-1}, \theta_1, \ldots, \theta_{n-t}) \in \mathbb{C}^{n-s-1}.$$

$$\begin{aligned}(3)\ \mu_{3,s,t}(\Omega) &= \mu_0 + \Psi_{1,s} + \lambda_1\Psi_{1,s+1} + \cdots + \lambda_{t-s-1}\Psi_{1,t} + \Psi_{2,t} + \theta_1\Psi_{2,t+1} \\ &\quad + \theta_2\Psi_{2,t+2} + \cdots + \widehat{\theta_{s-3}\Psi_{2,s+t-3}} + \cdots + \theta_{n-t}\Psi_{2,n},\end{aligned}$$

where

$$\frac{t+3}{2} \leqslant s < t, \quad t > \frac{n+4}{2}, \quad s+t-3 \leqslant n, \quad t \neq s+2,$$
$$\Omega = (\lambda_1, \ldots, \lambda_{t-s-1}, \theta_1, \ldots, \theta_{s-4}, \theta_{s-2}, \ldots, \theta_{n-t}) \in \mathbb{C}^{n-s-2}.$$

$$\begin{aligned}(4)\ \mu_{4,s,s+2}(\Omega) &= \mu_0 + \Psi_{1,s} + \Psi_{1,s+1} + \theta_0\Psi_{2,s+2} + \theta_1\Psi_{2,s+3} + \cdots \\ &\quad + \theta_{n-s-2}\Psi_{2,n},\end{aligned}$$

where

$$s > \frac{n+1}{2}, \quad \Omega = (\theta_0, \theta_1, \ldots, \theta_{n-s-2}) \in \left(\mathbb{C}^*, \mathbb{C}^{n-s-2}\right).$$

$$(5)\ \mu_{5,s,s+2}(\Omega) = \mu_0 + \Psi_{1,s} + \theta_0\Psi_{2,s+2} + \Psi_{2,k} + \theta_{k-s-1}\Psi_{2,k+1} + \cdots + \theta_{n-s-2}\Psi_{2,n},$$

where

$$\frac{n+1}{2} < s < k-2 \leqslant n-2, \quad \Omega = (\theta_0, \theta_{k-s-1}, \theta_{k-s}, \ldots, \theta_{n-s-2}) \in \left(\mathbb{C}^*, \mathbb{C}^{n-k}\right).$$

$$(6)\ \mu_{6,s,s+2}(\Omega) = \mu_0 + \Psi_{1,s} + \Psi_{1,s+1} + \theta_0\Psi_{2,s+2} + \theta_1\Psi_{2,s+3} + \cdots + \theta_{n-s-3}\Psi_{2,n-1},$$

where

$$n = 2s-1 \geqslant 9, \quad \Omega = (\theta_0, \theta_1, \ldots, \theta_{n-s-3}) \in \left(\mathbb{C}^*, \mathbb{C}^{n-s-3}\right).$$

$$(7)\ \mu_{7,s,s+2}(\Omega) = \mu_0 + \Psi_{1,s} + \theta_0\Psi_{2,s+2} + \Psi_{2,k} + \theta_{k-s-1}\Psi_{2,k+1} + \cdots + \theta_{n-s-3}\Psi_{2,n-1},$$

where

$$n = 2s-1, \quad s+2 < k \leqslant n-1,$$
$$\Omega = (\theta_0, \theta_{k-s-1}, \ldots, \theta_{n-s-3}) \in \left(\mathbb{C}^*, \mathbb{C}^{n-k-1}\right).$$

$$(8)\ \mu_{8,s,s+1}(\Omega) = \mu_0 + \Psi_{1,s} + \Psi_{2,s+1} + \theta_1\Psi_{2,s+2} + \cdots + \theta_{n-s-2}\Psi_{2,n+1},$$

where

$$n = 2s-3, \quad s \geqslant 4,$$
$$\Omega = (\theta_1, \theta_2, \ldots, \theta_{n-s-2}) \in \mathbb{C}^{n-s-2}.$$

$$(9)\ \mu_{9,s,s+2}(\theta_0) = \mu_0 + \Psi_{1,s} + \theta_0\Psi_{2,s+2}, \quad \textit{where } s > \frac{n}{2},\ \theta_0 \in \mathbb{C}^*.$$

$$(10)\ \mu_{10,4,6}(\theta_0) = \mu_0 + \Psi_{1,4} + \theta_0\Psi_{2,6} + \Psi_{2,7}, \quad \textit{where } n = 7,\ \theta_0 \in \mathbb{C}.$$

$$(11)\ \mu_{11,t}(\Omega) = \mu_0 + \Psi_{2,t} + \Psi_{2,k} + \theta_{k-t+1}\Psi_{2,k+1} + \cdots + \theta_{n-t-1}\Psi_{2,n-1},$$

where

$$n = 2t-5 > 11, \quad t < k, \quad \Omega = (\theta_{k-t+1}, \theta_{k-t+2}, \ldots, \theta_{n-t-1}) \in \mathbb{C}^{n-k-1}.$$

$$(12)\ \mu_{12,t}(\Omega) = \mu_0 + \Psi_{2,t} + \Psi_{2,k} + \theta_{k-t+1}\Psi_{2,k} + \cdots + \theta_{n-t}\Psi_{2,n},$$

where

$$\frac{n+5}{2} < t \leqslant n-1, \quad t < k, \quad \Omega = (\theta_{k-t+1}, \theta_{k-t+2}, \ldots, \theta_{n-t}) \in \mathbb{C}^{n-k}.$$

$$(13)\ \mu_{13,t} = \mu_0 + \Psi_{2,t}, \quad \textit{where } t > \frac{n+4}{2}.$$

$$(14)\ \mu_{14,s}(\Omega) = \mu_0 + \Psi_{1,s} + \Psi_{1,k} + \lambda_{k-s+1}\Psi_{1,k+1} + \cdots + \lambda_{s-3}\widehat{\Psi_{1,2s-3}} + \cdots + \lambda_{n-s}\Psi_{1,n},$$

where

$$4 \leqslant s < k < 2s-3 \leqslant n, \quad \Omega = (\lambda_{k-s+1}, \ldots, \lambda_{s-4}, \lambda_{s-2}, \ldots, \lambda_{n-s}) \in \mathbb{C}^{n-k-1}.$$

$$(15)\ \mu_{15,s}(\Omega) = \mu_0 + \Psi_{1,s} + \Psi_{1,k} + \lambda_{k-s+1}\Psi_{1,k+1} + \lambda_{k-s+2}\Psi_{1,k+2} + \cdots + \lambda_{n-s}\Psi_{1,n},$$

where

$$4 \leqslant s < k \leqslant n, \quad 2s-3 > n \text{ or } 2s-3 < k,$$
$$\Omega = (\lambda_{k-s+1}, \lambda_{k-s+2}, \ldots, \lambda_{n-s}) \in \mathbb{C}^{n-k}.$$

$(16)\ \mu_{16,s}(\Omega) = \mu_0 + \Psi_{1,s}$, *where* $4 \leqslant s \leqslant n$.

$(17)\ \mu_{17} = \mu_0$.

All the Lie algebra laws described above are pairwise non-isomorphic except in the following cases:

$\mu_{i,s,t}(\Omega) \cong \mu_{i,s,t}(\Omega')$, $i = 1,2,3$, *if and only if* $\lambda'_j = \lambda_j \varepsilon^j$, $j = 1, \ldots, t-s-1$, $\theta'_j = \theta_j \varepsilon^j$, $j = 1, \ldots, n-t$; *where* ε *is a* $(k-s-2)$*-th root of* 1.

$\mu_{i,s,t}(\Omega) \cong \mu_{i,s,t}(\Omega')$, $i = 5,7$, *if and only if* $\theta'_j = \theta_j \varepsilon^{j+t-k}$, $j = 0, 1, \ldots, n-s-2$ (*for* $i = 5$) *and* $j = 0, 1, \ldots, n-s-3$ (*for* $i = 7$), *where* ε *is a* $(k-s-2)$*-th root of* 1.

THEOREM 41. $\mu_{i,t}(\Omega) \cong \mu_{i,t}(\Omega')$, $i = 11, 12$, *if and only if* $\theta'_j = \theta_j \varepsilon^j$, $j = k-t+1, \ldots, n-t-1$ (*for* $i = 11$) *and* $j = k-t+1, \ldots, n-t$ (*for* $i = 12$), *where* ε *is a* $(k-t)$*-th root of* 1.

THEOREM 42. $\mu_{i,s}(\Omega) \cong \mu_{i,s}(\Omega')$, $i = 14, 15$, *if and only if* $\lambda'_j = \lambda_j \varepsilon^j$, $j = k-s+1, \ldots, n-s$, *where* ε *is a* $(k-s)$*-th root of* 1.

COMMENTS. We have a complete classification of the k-Abelian Lie algebras for $k \leqslant 2$. Progress in the case $k \geqslant 3$ can be interesting for a description of some irreducible components of the variety of nilpotent Lie algebra laws (see the chapter "Varieties of Lie algebra laws" of this Handbook).

References for this subsection: [10,19].

3.6. *Left-symmetric filiform Lie algebras*

Let ∇ be a linear connection on the Lie group G. This connection is called *locally flat* if

$$C(X,Y) = \nabla_{[X,Y]} - [\nabla_X, \nabla_Y] = 0,$$
$$T(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] = 0$$

for all elements X and Y in the space $\Gamma(TG)$ of sections of the tangent bundle of G. We say that the group G is locally flat if there exists a locally flat connection on G. Let G be a locally flat Lie group. The vector space $\Gamma(TG)$ is then an algebra with product $XY = \nabla_X Y$ and we have the following property for this product:

$$(XY)Z - X(YZ) = (YX)Z - Y(XZ).$$

An algebra with a product satisfying this property is called a *left-symmetric algebra*. If A is a left-symmetric algebra, A is a Lie algebra under the product $[X, Y] = XY - YX$. This Lie algebra is said to be sub-adjacent to the left-symmetric algebra A. The canonical Lie structure of $\Gamma(TG)$ coincides with the original Lie structure. G is locally flat if and only if $\Gamma(TG)$ has a left-symmetric product, which is compatible with its canonical Lie structure and satisfies the following conditions

$$(gX)Y = g(XY), \quad X(fY) = X(f)Y + f(XY)$$

for all appropriate elements. Let G be a finite dimensional Lie group and let $\mathfrak{g}$ be its Lie algebra. The existence of a left-symmetric product on the Lie algebra $\mathfrak{g}$ compatible with its Lie structure is equivalent to the existence of a left-invariant locally flat connection on G.

THEOREM 43. *Let $\mathfrak{g}$ be a filiform Lie algebra admitting a semisimple derivation (see Section 1 for its description) and let $\mathfrak{g} \neq C_{n+1}(\alpha_1, \ldots, \alpha_t)$. Then $\mathfrak{g}$ can be given a left-symmetric product compatible with its Lie structure.*

In fact, let d be a semisimple derivation of $\mathfrak{g}$. If $\mathfrak{g}$ is a Lie algebra of rank 1 (see 1.2), then all eigenvalues of d are different of zero. If $\mathfrak{g}$ is a Lie algebra of rank 2, then $\mathfrak{g} = L_n$ or Q_n and always there exist a derivation with eigenvalues different from zero and we can suppose that d is a such derivation. Finally, if $\mathfrak{g}$ is a filiform Lie algebra satisfying the conditions of the theorem then d is a nondegenerate derivation. In this case a left-symmetric product always exists.

COMMENTS. Recently Benoist has constructed an 11-dimensional filiform Lie algebra not provided with a affine structure (for a filiform Lie algebra this is equivalent to nonexistence of a left-symmetric product compatible with its Lie structure). It is very interesting to determine all low-dimensional filiform Lie algebras (for example of dimension $\leqslant 11$) not provided with such product.

References for this subsection: [5,6,11,18,33,34,39–42,51,54,55].

3.7. *Standard Lie algebras*

3.7.1. *Parabolic subalgebras of a semisimple algebra.* Let $\mathfrak{g}$ be a semisimple complex Lie algebra.

DEFINITION 7. A subalgebra $\mathfrak{g}_0$ of $\mathfrak{g}$ is called parabolic if it contains a Borel subalgebra of $\mathfrak{g}$.

Let's recall what a Borel subalgebra is. First we fix the following notations: $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, Δ the set of roots corresponding to $\mathfrak{h}$, S a basis of Δ (the simple roots), Δ_+ (respectively, Δ_-) the set of positive (respectively, negative) roots (recall that $\Delta = \Delta_+ \cup \Delta$) and

$$\mathfrak{g}_\alpha = \bigl\{X \in \mathfrak{g}\colon\ [X, H] = \alpha(H)X \text{ for all } H \in \mathfrak{h}\bigr\}.$$

This space $\mathfrak{g}_\alpha$ has dimension one. We choose a non-null vector X_α in $\mathfrak{g}_\alpha$. A Borel subalgebra $\mathfrak{b}$ of $\mathfrak{g}$ is a solvable maximal subalgebra of $\mathfrak{g}$. It is conjugate, by an inner automorphism, with a subalgebra of the following type

$$\mathfrak{b}' = \mathfrak{h} \oplus \sum_{\alpha \in \Delta} \mathfrak{g}_\alpha.$$

The parabolic subalgebras are determined, up to an inner automorphism, by a subset S_1 of S. Let Δ_1 be the set of roots whose decomposition on S contains only elements of $S \setminus S_1$. One writes $\Delta_2 = \Delta \setminus \Delta_1$, $\Delta_2^+ = \Delta_2 \cap \Delta_+$. The Lie algebra

$$\mathfrak{p} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta_2^+ \cup \Delta_1} \mathfrak{g}_\alpha$$

is parabolic and every parabolic subalgebra is conjugate to an algebra $\mathfrak{p}$.

We note that the nilradical of $\mathfrak{p}$ is

$$\mathfrak{n} = \sum_{\alpha \in \Delta_2^+} \mathfrak{g}_\alpha$$

and its reductive part is

$$\mathfrak{r} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta_1} \mathfrak{g}_\alpha.$$

3.7.2. *Standard subalgebra*

DEFINITION 8. A subalgebra of a semisimple Lie algebra is called standard if its normalizer is a parabolic subalgebra.

These algebras have been studied, for the first time, by G.B. Gurevich, but only in the case where the simple algebra $\mathfrak{g}$ is of type A_r. The motivation for this study can be founded in the theory of complex homogeneous spaces: let M be a compact homogeneous space $M = G/H$, G being a complex Lie group and M a closed subgroup. If $\mathfrak{g}$ and $\mathfrak{h}$ are the Lie algebras corresponding to G and H, the normalizer of $\mathfrak{h}$ in $\mathfrak{g}$ is parabolic. This result has been established by Tits. It permits to translate the study of homogeneous complex spaces into the study of standard subalgebras. In order to simplify terminology, we shall call *standard algebra* every standard subalgebra of a semisimple Lie algebra.

3.7.3. *Some properties of standard algebras.* We note that, for every standard algebra $\mathfrak{t}$ (not necessarily nilpotent) whose normalizer has the following form

$$\mathfrak{p} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta_1 \cup \Delta_2^+} \mathfrak{g}_\alpha$$

(which is the canonical form of parabolic algebras), one has:

$$\mathfrak{t} = \mathfrak{h}_1 \oplus \sum_{\alpha \in \Delta'} \mathfrak{g}_\alpha,$$

where $\Delta' \subset \Delta$ and $\mathfrak{h}_1 \subset \mathfrak{h}$.

We consider a partial order relation on the dual $\mathfrak{h}^*$ to the Cartan subalgebra $\mathfrak{h}$: $\omega_1 \geqslant \omega_2$ if and only if $\omega_1 - \omega_2$ is a linear combination of simple roots with non-negative coefficients.

PROPOSITION 44. *Let $\mathfrak{t}$ be a standard algebra such that its normalizer can be written as $\mathfrak{h} \oplus \sum_{\alpha \in \Delta_1 \cup \Delta_2^+} \mathfrak{g}_\alpha$. Suppose that α and β are positive roots with $\alpha \leqslant \beta$. If the subspace $\mathfrak{g}_\alpha$ is included in $\mathfrak{t}$, then $\mathfrak{g}_\beta$ also is in $\mathfrak{t}$.*

PROOF. We consider in each subspace $\mathfrak{g}_\alpha$ a nonzero vector X_α. If $\gamma \geqslant \alpha$, one can write $\gamma = \alpha + \gamma_1 + \cdots + \gamma_k$ where the γ_i are positive roots. One can order again the terms of this sum

$$\gamma = \gamma_{i_1} + \cdots + \gamma_{i_s} + \alpha + \gamma_{i_{s+1}} + \cdots + \gamma_{i_k}$$

such that every partial connected sum beginning which a γ_{i_1} is a positive root. Each vector X_{γ_i} is in the normalizer of $\mathfrak{t}$. Then the image of the vector X_α by the composition

$$\operatorname{ad} X_{\gamma_{i_k}} \circ \operatorname{ad} X_{\gamma_{i_{k-1}}} \circ \cdots \circ \operatorname{ad} X_{\gamma_{i_{s+1}}} \circ \operatorname{ad} X_{\gamma_{i_1} + \cdots + \gamma_{i_s}}$$

is in the space $\mathfrak{g}_\gamma$. This proves the proposition. □

For the description of standard algebras in the general case see the references. Here we study only nilpotent standard algebras.

3.7.4. *Nilpotent standard algebras.* Let R be a subset of Δ_+ whose elements are pairwise not comparable (for the ordering in $\mathfrak{h}^*$ given in the preceding section). One puts:

$$R_1 = \{\alpha \in \Delta_+ \colon \alpha \geqslant \beta \text{ for } \beta \in R\}.$$

The subspace $\mathfrak{n} = \sum_{\alpha \in R_1} \mathfrak{g}_\alpha$ is a nilpotent subalgebra of $\mathfrak{g}$.

The normalizer of $\mathfrak{n}$ contains a Borel subalgebra. Thus $\mathfrak{n}$ is a nilpotent standard subalgebra of $\mathfrak{g}$. We will say that $\mathfrak{n}$ is *the nilpotent standard subalgebra associated to* R. This process permits to construct more easily such subalgebras. The following shows that every standard algebra is of this type. So one obtains a complete description of these algebras.

THEOREM 45. *Let* $\mathfrak{n}$ *be a nilpotent standard subalgebra whose normalizer has the form* $\mathfrak{h}_1 \oplus \sum_{\alpha \in \Delta_1 \cup \Delta_2^+} \mathfrak{g}_\alpha$. *Then there is a subset* R *of* Δ_+ *consisting of pairwise not comparable elements of* Δ_+ *such that* $\mathfrak{n}$ *is the standard nilpotent subalgebra associated to* R.

PROOF. As $\mathfrak{n}$ is an ideal of its normalizer, it is contained in the radical of this normalizer. But $\mathfrak{n}$ can be written as $\mathfrak{n} = \sum_{\alpha \in \Delta'} \mathfrak{g}_\alpha$ for some $\Delta' \subset \Delta$. One deduces $\Delta' \subset \Delta_2^+ \subset \Delta_+$. Let R be the subset of Δ' formed by the minimal elements, still for the some partial ordering. These elements are pairwise noncomparable. Then R defines $\mathfrak{n}$. □

COROLLARY 46. *Every standard nilpotent algebra is conjugate to a nilpotent standard algebra associated to a set* R *of pairwise noncomparable roots.*

3.7.5. *Structure of normalizers of nilpotent standard subalgebras*

DEFINITION 9. A basic root $\alpha \in S$ is called extremal for $\beta \in \Delta_+$ if it satisfies $\alpha = \beta$ or $\beta - \alpha \in \Delta$.

If $\beta \in \Delta_+$, we note S^β the set of the extremal roots for β.

LEMMA 47. *Suppose that* $\alpha_1, \alpha_2, \alpha_3, \alpha_1 + \alpha_2, \alpha_1 + \alpha_2 + \alpha_3$ *are roots such that* $\alpha_1 + \alpha_3$ *and* $\alpha_2 + \alpha_3$ *are non-zero elements. Then one of these combinations* $\alpha_1 + \alpha_3$ *or* $\alpha_2 + \alpha_3$ *is a root.*

The proof is a direct consequence of the Jacobi identity concerning $(X_{\alpha_1}, X_{\alpha_2}, X_{\alpha_3})$.

LEMMA 48. *Let* α *and* β *be two positive roots such that* $\alpha + \beta \in \Delta$. *If* $\gamma \in S$ *is a non-extremal root for* α *and for* β, *then* γ *is non-extremal root for* $\alpha + \beta$.

PROOF. We suppose γ extremal for $\alpha + \beta$. As $\gamma \in S$, then $\gamma \neq \alpha + \beta$ and $\alpha + \beta - \gamma \in \Delta$. From the Lemma 1, $\alpha - \gamma$ or $\beta - \gamma$ is a root and γ is extremal for α or for β. This is impossible. □

LEMMA 49. *Let* α *and* β *be two roots such that the decomposition of* α *as sum of basic roots contains all the roots of* S^β *(the set of extremal roots of* β*) with a multiplicity greater than that of* β. *Then* $\alpha \geqslant \beta$.

PROOF. We put $\alpha = \beta + \gamma_1 + \cdots + \gamma_k - \xi_1 - \cdots - \xi_m$ with γ_i and ξ_j in S and $\gamma_i \neq \xi_j$ for every i and j. By hypothesis the roots $\xi_1, \ldots, \xi_m$ are not in S^β. Suppose now that α is not greater or equal to β (i.e. $m \geqslant 1$). We are going to show that $\nu = \beta + \gamma_1 + \cdots + \gamma_k$ is a root of Δ and that the $\xi_1, \ldots, \xi_m$ are not in S^ν. For $k = 0$, the results are obvious. So suppose $k \geqslant 1$ and we make an induction on k. One has:

$$(\beta, \beta) = (\beta, \alpha) + \sum_{i=1}^{k} (\beta, -\gamma_i) + \sum_{j=1}^{m} (\beta, \xi_j) > 0, \qquad (*)$$

where $(\,,)$ is the usual product on the set of the roots associated with the Killing–Cartan form.

As $(\beta, \alpha) \leqslant 0$ and $\sum_{j=1}^{m} (\beta, \xi_j) \leqslant 0$ (if not, $\beta - \alpha$ and $\beta - \sum \xi_j$ are roots, this is opposite to $m \geqslant 1$ or to $\xi_1, \ldots, \xi_m \notin S^\beta$) there exists (from $(*)$), an integer r, $1 \leqslant r \leqslant k$, such that $(\beta, -\gamma_r) > 0$. So $\beta + \gamma_r \in \Delta$. From the Lemma 47, $\xi_1, \ldots, \xi_m \notin S^{\beta+\gamma_r}$. We apply the induction hypothesis to prove the previous assertion.

So, one has

$$(\nu, \nu) = (\nu, \alpha) + \sum_{i=1}^{m} (\nu, \xi_i) > 0$$

and $\xi_1, \ldots, \xi_m \notin S^\nu$. One deduces $(\nu, \alpha) > 0$ and $\xi = \nu - \alpha = \xi_1 + \cdots + \xi_m$ is in Δ. We can suppose that each partial sum $\xi_1 + \cdots + \xi_r$ of ξ is in Δ. Let $\xi' = \xi_1 + \cdots + \xi_{m-1}$. Then $\alpha = -\xi' - \xi_m + \nu$. From Lemma 47, one has $\nu - \xi' = \nu - \xi_1 - \cdots - \xi_{m-1}$ is a root. By a simple induction we can affirm that $\nu - \xi_1 \in \Delta$. As $\xi_1 \notin S^\nu$, this is impossible. □

THEOREM 50. *Let $\mathfrak{n}$ be a standard nilpotent subalgebra associated to a set R of pairwise non-comparable roots. Then the normalizer $N(\mathfrak{n})$ of $\mathfrak{n}$ (it is parabolic because $\mathfrak{n}$ is standard) is defined by the subset*

$$S_1 = \bigcup_{\beta \in R} S^\beta \subset S.$$

PROOF. Let $\mathfrak{r}$ be a parabolic subalgebra associated to the system S_1 and let $\mathfrak{n} = \sum_{\alpha \in R_1} \mathfrak{g}_\alpha$ (R_1 is defined above). By Lemma 47, we can write $\mathfrak{r} \subset N(\mathfrak{n})$.

Suppose that $\mathfrak{r} \neq N(\mathfrak{n})$. As $\Delta = \Delta_1 \cup \Delta_2$, there is an $\alpha \in \Delta_2^+$ such that $X_{-\alpha} \in N(\mathfrak{n})$. This shows that the decomposition of α in simple roots contains one element ν of S_1 (by definition of S_1, ν is an extremal element for a $\beta \in R$). But $N(\mathfrak{n})$ is a parabolic subalgebra. Then the vector $X_{-\nu}$ is in the normalizer $N(\mathfrak{n})$. More, we have either $\beta - \nu \in \Delta$ or $\beta = \nu$. If $\beta - \nu \in \Delta$ then $[X_\beta, X_{-\nu}] = a.X_{\beta-\nu} \in \mathfrak{n}$ with $a \neq 0$. If $\beta = \nu$ then $[X_\beta, X_{-\nu}] = [X_\nu, X_{-\nu}] \in \mathfrak{n}$, but this vector is also in the Cartan subalgebra. These two jads form a contradiction. □

COROLLARY 51. *Every system $R \subset \Delta^+$ of pairwise noncomparable roots defines a nilpotent standard subalgebra:*

$$\mathfrak{n} = \sum_{\alpha \in R_1} \mathfrak{g}_\alpha.$$

Its normalizer is the parabolic subalgebra associated to the subsystem

$$S_1 = \bigcup_{\beta \in R} S^{\beta}.$$

Conversely, every nilpotent standard subalgebra is conjugate to a subalgebra $\mathfrak{n}$ *associated to a subsystem* $R \subset \Delta^+$ *of pairwise noncomparable roots.*

3.7.6. *On nilpotent algebras of maximal rank.* Let $\mathfrak{g}$ be a semi simple complex Lie algebra and let $\mathfrak{g}^+ = \sum_{\alpha \in \Delta^+} \mathfrak{g}_\alpha$ be its nilpotent part. If $\mathfrak{n}$ is a standard nilpotent Lie algebra then the quotient $\mathfrak{g}^+/\mathfrak{n}$ is also nilpotent. This process permits to construct a class of nilpotent Lie algebras. This class contains all nilpotent Lie algebras of maximal rank studied by G. Favre and L. Santharoubane.

3.7.7. *Complete Standard Lie algebras*

DEFINITION 10. A standard nilpotent algebra $\mathfrak{n}$ is called complete if it is the nilradical of its normalizer.

It is easy to see that such an algebra is conjugate to a subalgebra associated to a subset of simple roots.

THEOREM 52. *Let* $\mathfrak{n}$ *be a standard nilpotent Lie algebra associated to a system* $R \subset \Delta^+$ *of non-comparable roots. The following assertions are equivalent:*

(i) $\mathfrak{n}$ *is an intersection of complete standard nilpotent subalgebras associated to some subsystems of* S,

(ii) *every element* $\beta \in R$ *can be written under the form* $\beta = \alpha_1 + \cdots + \alpha_m$ *with* $\alpha_1, \ldots, \alpha_m \in S$ *and* $\alpha_i \neq \alpha_j$ *for* $i \neq j$.

PROOF. (ii) $\Rightarrow$ (i) We put $R = \{\beta_1, \ldots, \beta_k\}$ and $\phi_R = \{\bigcup_{i=1}^{k} \{\alpha_i\}\colon \alpha_i \in S^{\beta_i},\ i = 1, \ldots, k\}$.

Every element S' of ϕ_R is a subset of S and defines a standard nilpotent complete subalgebra $\mathfrak{n}_{S'}$. We easily show that $\mathfrak{n}_{S'} \supset \mathfrak{n}$ for all $S' \in \phi_R$. Then $\mathfrak{n}' = \bigcap_{S \in \phi_R} \mathfrak{n}_{S'}$ satisfies the inclusion $\mathfrak{n}' \supset \mathfrak{n}$. Suppose that $\alpha \in \Delta^+$, $X_\alpha \in \mathfrak{n}'$ and $X_\alpha \notin \mathfrak{n}$.

The decomposition of α in simple roots does not contain the roots of S^β when $\beta \in R$ (if not we will have $X_\alpha \in \mathfrak{n}$, by Lemma 47). If we consider the elements $S' = \bigcup_{i=1}^{k} \{\alpha_i\}$ of ϕ_R, where $\alpha \neq \alpha_i$ for $i = 1, \ldots, k$ (recall that $\alpha_i \in S^{\beta_i}$), then $X_\alpha \notin \mathfrak{n}_{S'}$; this is contrary to the choice of X_α.

(i) $\Rightarrow$ (ii) Let be $\mathfrak{n} = \bigcap_{i=1}^{n} \mathfrak{n}_i$ where $\mathfrak{n}_i$ is a complete nilpotent standard subalgebra associated to the system S_i of simple roots. Consider a root β in R. Suppose that its decomposition $\beta = k_i\alpha_1 + \cdots + k_m\alpha_m$ contains a coefficient k_i strictly greater than 1. As X_β is in $\mathfrak{n} = \bigcap \mathfrak{n}_i$, there is γ_i in S_i such that $\beta \geqslant \gamma_i$ for all i, $1 \leqslant i \leqslant n$. Then for the root $\beta' = \alpha_1 + \cdots + \alpha_m \in R^+$, we shall have also $\beta' \geqslant \gamma_i$ for all i. Then the vector $X_{\beta'}$ is in each $\mathfrak{n}_i$ for all i and thus $X_{\beta'}$ is in $\mathfrak{n}$. It is impossible because $\beta' < \beta$. □

REMARK 5. If the semisimple Lie algebra $\mathfrak{g}$ is of type A_n, condition (ii) is necessary satisfied. The decomposition of $\mathfrak{n}$ as intersection of complete subalgebra always is true. This result was already been proved by Gurevich.

COMMENTS. It would be very interesting to consider the standard subalgebras of the affine Kac–Moody algebras.

References for this subsection: [17,21,25,31].

4. On the Classification of low-dimensional nilpotent and solvable Lie algebras

4.1. *Classification of nilpotent Lie algebras of dimension less than 7*

4.1.1. *The classification in dimension less than 6.* By convention, all not defined brackets are null.

The classifications in dimensions $\leqslant 5$ are given for the nilpotent Lie algebras over a field of characteristic 0. In dimension 6 we give it in the cases $K = \mathbb{C}$ and $K = \mathbb{R}$.

Dimension 1

$\mathfrak{n}_1^1 = \mathfrak{a}_1 =$ the Abelian Lie algebra of dimension 1.

Dimension 2

$\mathfrak{n}_1^2 = \mathfrak{a}_1^2 =$ the Abelian algebra of dimension 2.

Dimension 3

$\mathfrak{n}_1^3 = \mathfrak{a}_1^3 =$ the Abelian algebra of dimension 3,
$\mathfrak{n}_2^3 = H_1 =$ the Heisenberg algebra defined by $[X_1, X_2] = X_3$.

From the dimension 4, we do not write down the decomposable Lie algebras, that is those Lie algebras which are direct sums of proper ideals.

Dimension 4

$\mathfrak{n}_1^4$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$. This is a filiform law.

Dimension 5

$\mathfrak{n}_1^5$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$,

$\mathfrak{n}_2^5$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$;
$[X_2, X_3] = X_5$.

These two laws are the 5-dimensional filiform laws.

$\mathfrak{n}_3^5$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_2, X_3] = X_5$,

$\mathfrak{n}_4^5$: $[X_1, X_3] = X_5$; $[X_2, X_4] = X_5$,

$\mathfrak{n}_5^5$: $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_2, X_3] = X_5$,

$\mathfrak{n}_6^5$: $[X_1, X_2] = X_3$; $[X_1, X_4] = X_5$.

Dimension 6 ($K = \mathbb{C}$)

$\mathfrak{n}_1^6$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_1, X_5] = X_6$,

$\mathfrak{n}_2^6$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_1, X_5] = X_6$; $[X_2, X_3] = X_6$,

$\mathfrak{n}_3^6$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_1, X_5] = X_6$; $[X_2, X_3] = X_5$; $[X_2, X_4] = X_6$,

$\mathfrak{n}_4^6$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_1, X_5] = X_6$; $[X_3, X_4] = X_6$; $[X_2, X_5] = -X_6$,

$\mathfrak{n}_5^6$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_1, X_5] = X_6$, $[X_3, X_4] = X_6$; $[X_2, X_3] = X_5$; $[X_2, X_4] = X_6$; $[X_2, X_5] = -X_6$.

These laws are the filiform laws.

$\mathfrak{n}_6^6$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_6$; $[X_2, X_3] = X_5$; $[X_2, X_5] = X_6$,

$\mathfrak{n}_6^7$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_6$; $[X_2, X_5] = X_6$,

$\mathfrak{n}_6^8$: $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_1, X_5] = X_6$; $[X_2, X_3] = X_5$; $[X_2, X_4] = X_6$,

$\mathfrak{n}_6^9$: $[X_1, X_2] = X_3 + X_5$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_6$; $[X_1, X_5] = X_6$,

$\mathfrak{n}_6^{10}$: $[X_1, X_3] = X_4$; $[X_1, X_4] = X_6$; $[X_2, X_3] = X_5$; $[X_2, X_5] = X_6$,

$\mathfrak{n}_6^{11}$: $[X_1, X_2] = X_5$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_6$; $[X_2, X_5] = X_6$,

$\mathfrak{n}_6^{12}$: $[X_1, X_3] = X_4$; $[X_1, X_4] = X_6$; $[X_2, X_5] = X_6$,

$\mathfrak{n}_6^{13}$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_2, X_3] = X_6$,

$\mathfrak{n}_6^{14}$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_5] = X_6$; $[X_2, X_3] = X_6$,

$\mathfrak{n}_6^{15}$: $[X_1, X_2] = X_3 + X_5$; $[X_1, X_3] = X_4$; $[X_2, X_5] = X_6$,

$\mathfrak{n}_6^{16}$: $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_2, X_3] = X_6$,

$\mathfrak{n}_6^{17}$: $[X_1, X_2] = X_6$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_5$; $[X_2, X_3] = X_5$,

$\mathfrak{n}_6^{18}$: $[X_1, X_2] = X_5$; $[X_1, X_3] = X_6$; $[X_2, X_4] = X_6$,

$\mathfrak{n}_6^{19}$: $[X_1, X_3] = X_6$; $[X_1, X_3] = X_4$; $[X_1, X_3] = X_5$,

$\mathfrak{n}_6^{20}$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_5] = X_6$.

If $K = \mathbb{R}$, we have $\mathfrak{n}_6^1, \mathfrak{n}_6^2, \ldots, \mathfrak{n}_6^{20}$ and

$\mathfrak{n}_6^{21}$: $[X_1, X_3] = X_5$; $[X_1, X_4] = X_6$; $[X_2, X_4] = X_5$; $[X_2, X_3] = -X_6$,

$\mathfrak{n}_6^{22}$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_5$; $[X_1, X_4] = X_6$; $[X_2, X_4] = X_5$; $[X_2, X_3] = -X_6$,

$\mathfrak{n}_6^{23}$: $[X_1, X_3] = X_4$; $[X_1, X_4] = X_6$; $[X_2, X_3] = X_5$; $[X_2, X_5] = -X_6$,

$\mathfrak{n}_6^{24}$: $[X_1, X_2] = X_3$; $[X_1, X_3] = X_4$; $[X_1, X_4] = X_6$; $[X_2, X_3] = X_5$; $[X_2, X_5] = -X_6$.

4.1.2. *On the classification in dimension* 7. The previous tables show that, for the dimension less than 6, there is a finite number of isomorphism classes of nilpotent Lie algebras. But, for the dimension 7 and more, we are going to find families with one or more parameters of nonisomorphic Lie algebras. It is the same to say that the variety of isomorphism classes has nonzero dimension. These one parameter families of isomorphic classes of 7-dimensional nilpotent algebra is not the only obstacle. There is, in the variety of isomorphism classes, a number of isolated points. So every table of dimension 7 is very difficult to establish. At this time we have 3 or 4 tables given by various authors. Unfortunately, it is very complicated to compare these lists because the techniques used and the invariants considered are very different (characteristic sequence, rank, extension). The links between the invariant used are not evident. Probably, all the classifications of 7-dimensional nilpotent Lie algebras contain some mistakes.

References for this subsection: [2,21,43,45,47,48].

4.2. *On the classification of the low-dimensional filiform Lie algebras*

We have a complete classification up to isomorphism of the complex filiform Lie algebras of dimension $\leqslant 11$ (see the references). Here we give the list for the dimensions $\leqslant 9$. We use the notations of Section 3.1.

Dimension 3

μ_3^1: μ_0.

Dimension 4

μ_4^1: μ_0.

Dimension 5

μ_5^1: μ_0,
μ_5^2: $\mu_0 + \Psi_{1,4}$.

Dimension 6

μ_6^1: μ_0,
μ_6^2: $\mu_0 + \Psi_{1,4}$,
μ_6^3: $\mu_0 + \Psi_{1,5}$,
μ_6^4: $\mu_0 + \Psi_{2,5}$,
μ_6^5: $\mu_0 + \Psi_{1,4} + \Psi_{2,5}$.

Dimension 7

μ_7^1: μ_0,
μ_7^2: $\mu_0 + \Psi_{1,5}$,
μ_7^3: $\mu_0 + \Psi_{1,6}$,
μ_7^4: $\mu_0 + \Psi_{1,5} + \Psi_{1,6}$,
μ_7^5: $\mu_0 + \Psi_{1,4}$,
μ_7^6: $\mu_0 + \Psi_{1,4} + \Psi_{1,6}$,
μ_7^7: $\mu_0 + \Psi_{1,5} + \Psi_{2,6}$,
μ_7^8: $\mu_0 + \alpha\Psi_{1,4} + \Psi_{2,6}$.

Dimension 8

μ_8^1: $\mu_0 + \Psi_{1,4} + \alpha\Psi_{1,5} - 2\Psi_{2,6} + \Psi_{3,7}$,
μ_8^2: $\mu_0 + \Psi_{1,5} + \alpha\Psi_{1,6} + \Psi_{3,7}$,
μ_8^3: $\mu_0 + \Psi_{1,6} + \Psi_{3,7}$,
μ_8^4: $\mu_0 + \Psi_{3,7}$,
μ_8^5: $\mu_0 + \alpha\Psi_{1,4} + \Psi_{2,6} + \Psi_{2,7}$, $\alpha \neq 0$,
μ_8^6: $\mu_0 + \alpha\Psi_{1,4} + \Psi_{2,6}$,
μ_8^7: $\mu_0 + \Psi_{1,5} + \Psi_{2,6}$,
μ_8^8: $\mu_0 + \Psi_{1,4} + \alpha\Psi_{1,6} + \Psi_{2,7}$,
μ_8^9: $\mu_0 + \alpha\Psi_{1,5} + \Psi_{1,6} + \Psi_{2,7}$,
μ_8^{10}: $\mu_0 + \alpha\Psi_{1,5} + \Psi_{2,7}$,
μ_8^{11}: $\mu_0 + \alpha\Psi_{1,4} + \Psi_{1,6} + \Psi_{1,7}$,
μ_8^{12}: $\mu_0 + \Psi_{1,4} + \Psi_{1,6}$,
μ_8^{13}: $\mu_0 + \Psi_{1,4} + \Psi_{1,7}$,
μ_8^{14}: $\mu_0 + \Psi_{1,4}$,
μ_8^{15}: $\mu_0 + \Psi_{1,5} + \Psi_{1,6}$,
μ_8^{16}: $\mu_0 + \Psi_{1,5}$,
μ_8^{17}: $\mu_0 + \Psi_{1,6}$,
μ_8^{18}: $\mu_0 + \Psi_{1,7}$,
μ_8^{19}: μ_0.

Dimension 9

μ_9^1: $\mu_0+\Psi_{1,4}+\alpha\Psi_{2,6}+\Psi_{2,7}+\beta\Psi_{2,8}+\frac{3\alpha^2}{\alpha+2}\Psi_{3,8},\ \alpha\neq 0,-2,$
μ_9^2: $\mu_0+\Psi_{1,4}+\alpha\Psi_{2,6}+\Psi_{2,8}+\frac{3\alpha^2}{\alpha+2}\Psi_{3,8},\ \alpha\neq 0,-2,$
μ_9^3: $\mu_0+\Psi_{1,4}+\alpha\Psi_{2,6}+\frac{3\alpha^2}{\alpha+2}\Psi_{3,8},\ \alpha\neq -2,$
μ_9^4: $\mu_0+\Psi_{1,4}+\Psi_{1,8}+4\Psi_{2,6}+\alpha\Psi_{2,7}+\beta\Psi_{2,8}+8\Psi_{3,8},$
μ_9^5: $\mu_0+\Psi_{1,4}+\alpha\Psi_{1,8}+4\Psi_{2,6}+\alpha\Psi_{2,8}+8\Psi_{3,8},$
μ_9^6: $\mu_0+\Psi_{1,4}+\Psi_{1,8}+4\Psi_{2,6}+8\Psi_{3,8},$
μ_9^7: $\mu_0+\Psi_{1,5}+\Psi_{2,6}+\alpha\Psi_{2,8}+3\Psi_{3,8},$
μ_9^8: $\mu_0+\Psi_{2,6}+\Psi_{2,8}+3\Psi_{3,8},$
μ_9^9: $\mu_0+\Psi_{2,6}+3\Psi_{3,8},$
μ_9^{10}: $\mu_0+\alpha\Psi_{1,5}+\beta\Psi_{1,6}+\Psi_{2,7}+\Psi_{3,8},\ \alpha\neq 1,-\frac{1}{3},$
μ_9^{11}: $\mu_0+\Psi_{1,5}+\alpha\Psi_{1,6}+\Psi_{2,7}+\beta\Psi_{2,8}+\Psi_{3,8},$
μ_9^{12}: $\mu_0+\frac{-1}{3}\Psi_{1,5}+\alpha\Psi_{1,6}+\Psi_{2,7}+\beta\Psi_{2,8}+\Psi_{3,8},$
μ_9^{13}: $\mu_0+\Psi_{1,5}+\alpha\Psi_{1,6}+\Psi_{3,8},$
μ_9^{14}: $\mu_0+\Psi_{1,6}+\alpha\Psi_{1,7}+\Psi_{3,8},$
μ_9^{15}: $\mu_0+\Psi_{1,7}+\Psi_{3,8},$
μ_9^{16}: $\mu_0+\Psi_{3,8},$
μ_9^{17}: $\mu_0+\Psi_{1,4}+\alpha\Psi_{1,6}+\Psi_{2,7}+\beta\Psi_{2,8},$
μ_9^{18}: $\mu_0+\alpha\Psi_{1,4}+\beta\Psi_{1,6}+\Psi_{1,7}+\Psi_{2,8},$
μ_9^{19}: $\mu_0+\Psi_{1,4}+\alpha\Psi_{1,6}+\Psi_{2,8},$
μ_9^{20}: $\mu_0+\alpha\Psi_{1,5}+\beta\Psi_{1,6}+\Psi_{2,7}+\Psi_{2,8},$
μ_9^{21}: $\mu_0+\alpha\Psi_{1,5}+\Psi_{1,6}+\Psi_{2,7},$
μ_9^{22}: $\mu_0+\alpha\Psi_{1,5}+\Psi_{2,7},$
μ_9^{23}: $\mu_0+\Psi_{1,5}+\alpha\Psi_{1,6}+\Psi_{2,8},$
μ_9^{24}: $\mu_0+\alpha\Psi_{1,6}+\Psi_{1,7}+\Psi_{2,8},$
μ_9^{25}: $\mu_0+\alpha\Psi_{1,6}+\Psi_{2,8},$
μ_9^{26}: $\mu_0+\alpha\Psi_{1,4}+\Psi_{1,6}+\Psi_{1,7}+\beta\Psi_{1,8},$
μ_9^{27}: $\mu_0+\Psi_{1,4}+\Psi_{1,6}+\alpha\Psi_{1,8},$
μ_9^{28}: $\mu_0+\alpha\Psi_{1,4}+\Psi_{1,7}+\Psi_{1,8},$
μ_9^{29}: $\mu_0+\Psi_{1,4}+\Psi_{1,7},$
μ_9^{30}: $\mu_0+\Psi_{1,4}+\Psi_{1,8},$
μ_9^{31}: $\mu_0+\Psi_{1,5}+\Psi_{1,6}+\alpha\Psi_{1,8},$
μ_9^{32}: $\mu_0+\Psi_{1,5}+\Psi_{1,8},$
μ_9^{33}: $\mu_0+\Psi_{1,5},$
μ_9^{34}: $\mu_0+\Psi_{1,6}+\Psi_{1,8},$
μ_9^{35}: $\mu_0+\Psi_{1,6},$

μ_9^{36}: $\mu_0 + \Psi_{1,7}$,
μ_9^{37}: $\mu_0 + \Psi_{1,8}$,
μ_9^{38}: μ_0.

References for this subsection: [1,20].

4.3. *On the rigid solvable Lie algebras of dimension* $\leqslant 8$

The following is the description up to isomorphism of the complex rigid solvable Lie algebras of dimension $\leqslant 8$. We give the nonzero brackets (the non-defined brackets being zero) relative to a basis. In this basis we denote by X a characteristic vector, by $Y_i, Y_i', Y_i'', \dots$ the eigenvectors of $\operatorname{ad} X$ corresponding to the eigenvalues $\lambda = i$ $(i \neq 0)$, by $X_0, X_0', X_0'', \dots$ the eigenvectors of $\operatorname{ad} X$ corresponding to the eigenvalues $\lambda = 0$. The rigid Lie algebras given by their laws μ_n^i and μ_n^j with $i \neq j$ are not isomorphic.

Dimension 2

$$\mu_2^i(X, X_1) = X_1.$$

Dimension 4

$$\begin{cases} \mu_4^1(X, X_1) = X_1, \\ \mu_4^1(X_0, X_2) = X_2. \end{cases}$$

Dimension 5

$$\begin{cases} \mu_5^1(X, Y_i) = iY_i, & i = 1, 2, 3, \\ \mu_5^1(X_0, Y_i) = Y_i, & i = 2, 3, \\ \mu_5^1(Y_1, Y_2) = Y_3. \end{cases}$$

Dimension 6

$$\begin{cases} \mu_6^1(X, Y_i) = iY_i, & i = 1, 2, 3, 4, 5, \\ \mu_6^1(Y_1, Y_i) = Y_{i+1}, & i = 2, 3, 4, \\ \mu_6^1(Y_2, Y_3) = Y_5, \end{cases}$$

$$\begin{cases} \mu_6^2(X, Y_i) = iY_i, & i = 1, 2, 3, 4, \\ \mu_6^2(X_0, Y_i) = Y_i, & i = 2, 3, 4, \\ \mu_6^2(Y_1, Y_i) = Y_{i+1}, & i = 2, 3, \end{cases}$$

$$\begin{cases} \mu_6^3(X, Y_i) = iY_i, & i = 1, 2, 3, \\ \mu_6^3(X_0, Y_2) = Y_2, \\ \mu_6^3(X_0', Y_3) = Y_3. \end{cases}$$

Dimension 7

$$\begin{cases} \mu_7^1(X, Y_i) = iY_i, & i = 1,2,3,4,5,6, \\ \mu_7^1(Y, Y_i) = Y_{i+1}, & i = 2,3,4,5, \\ \mu_7^1(Y_2, Y_i) = Y_{i+2}, & i = 3,4, \end{cases}$$

$$\begin{cases} \mu_7^2(X, Y_i) = iY_i, & i = 1,2,3,4,5,7, \\ \mu_7^2(Y_1, Y_i) = Y_{i+1}, & i = 2,3,4, \\ \mu_7^2(Y_2, Y_i) = Y_{i+2}, & i = 3,5, \\ \mu_7^2(Y_3, Y_4) = -Y_7, & \end{cases}$$

$$\begin{cases} \mu_7^3(X, Y_i) = iY_i, & i = 1,3,4,5,6,7, \\ \mu_7^3(Y_1, Y_i) = Y_{i+1}, & i = 3,4,5,6, \\ \mu_7^3(Y_3, Y_4) = Y_7, & \end{cases}$$

$$\begin{cases} \mu_7^4(X, Y_i) = iY_i, & i = 1,2,3,4,5, \\ \mu_7^4(X, Y_3') = 3Y_3', & \\ \mu_7^4(Y_1, Y_i) = Y_{i+1}, & i = 2,3,4, \\ \mu_7^4(Y_2, Y_3) = \mu_7^4(Y_2, Y_3') = Y_5, & \end{cases}$$

$$\begin{cases} \mu_7^5(X, Y_i) = iY_i, & i = 1,2,3,4,5, \\ \mu_7^5(X_0, Y_i) = Y_i, & i = 2,3,4,5, \\ \mu_7^5(Y_1, Y_i) = Y_{i+1}, & i = 2,3,4, \end{cases}$$

$$\begin{cases} \mu_7^6(X, Y_i) = iY_i, & i = 1,2,3,4,5, \\ \mu_7^6(X_0, Y_i) = Y_i, & i = 2,3,4, \\ \mu_7^6(X_0, Y_5) = 2Y_5, & \\ \mu_7^6(Y_1, Y_i) = Y_{i+1}, & i = 2,3, \\ \mu_7^6(Y_2, Y_3) = Y_5, & \end{cases}$$

$$\begin{cases} \mu_7^7(X, Y_i) = iY_i, & i = 1,2,3,4,5, \\ \mu_7^7(X_0, Y_i) = Y_i, & i = 3,4,5, \\ \mu_7^7(Y_1, Y_i) = Y_{i+1}, & i = 3,4, \\ \mu_7^7(Y_2, Y_3) = Y_5, & \end{cases}$$

$$\begin{cases} \mu_7^8(X, Y_i) = iY_i, & i = 1,2,3,4, \\ \mu_7^8(X_0, Y_i) = Y_i, & i = 2,3, \\ \mu_7^8(X_0', Y_4) = Y_4, & \\ \mu_7^8(Y_1, Y_2) = Y_3. & \end{cases}$$

Dimension 8

$$\begin{cases} \mu_8^1(X,Y_i)=iY_i, & i\in\{1,2,3,5,6,7,8\},\\ \mu_8^1(Y_1,Y_i)=Y_{i+1}, & i\in\{2,6,7\},\\ \mu_8^1(Y_2,Y_3)=Y_5, & \mu_8^1(Y_2,Y_5)=Y_7,\\ \mu_8^1(Y_2,Y_6)=Y_8, & \mu_8^1(Y_3,Y_5)=Y_8, \end{cases}$$

$$\begin{cases} \mu_8^2(X,Y_i)=iY_i, & i\in\{1,3,4,5,6,7,8\},\\ \mu_8^2(Y_1,Y_i)=Y_{i+1}, & i\in\{3,4,5,6,7\},\\ \mu_8^2(Y_3,Y_i)=Y_{i+3}, & i=4,5, \end{cases}$$

$$\begin{cases} \mu_8^3(X,Y_i)=iY_i, & i\in\{1,3,4,5,6,7,9\},\\ \mu_8^3(Y_1,Y_i)=Y_{i+1}, & i\in\{3,4,5,6\},\\ \mu_8^3(Y_3,Y_i)=Y_{i+3}, & i=4,6,\\ \mu_8^3(Y_4,Y_5)=-Y_9, & \end{cases}$$

$$\begin{cases} \mu_8^4(X,Y_i)=iY_i, & i\in\{1,4,5,6,7,8,9\},\\ \mu_8^4(Y_1,Y_i)=Y_{i+1}, & i\in\{4,5,6,7,8\},\\ \mu_8^4(Y_4,Y_5)=Y_9, & \end{cases}$$

$$\begin{cases} \mu_8^5(X,Y_i)=iY_i, & i\in\{2,3,4,5,6,7,8\},\\ \mu_8^5(Y_2,Y_i)=Y_{i+2}, & i\in\{3,4,5,6\},\\ \mu_8^5(Y_3,Y_i)=Y_{i+3}, & i=4,5, \end{cases}$$

$$\begin{cases} \mu_8^6(X,Y_i)=iY_i, & i\in\{2,3,4,6,7,8,10\},\\ \mu_8^6(Y_2,Y_i)=Y_{2+i}, & i=4,6,8,\\ \mu_8^6(Y_3,Y_i)=Y_{i+3}, & i=4,7,\\ \mu_8^6(Y_4,Y_6)=Y_{10}, & \end{cases}$$

$$\begin{cases} \mu_8^7(X,Y_i)=iY_i, & i\in\{2,3,5,6,7,8,9\},\\ \mu_8^7(Y_2,Y_i)=Y_{i+2}, & i=3,5,6,7,\\ \mu_8^7(Y_3,Y_i)=Y_{i+3}, & i=5,6, \end{cases}$$

$$\begin{cases} \mu_8^8(X,Y_i)=iY_i, & i\in\{2,3,5,7,8,9,11\},\\ \mu_8^8(Y_2,Y_i)=Y_{i+2}, & i\in\{3,5,7,9\},\\ \mu_8^8(Y_3,Y_i)=Y_{i+3}, & i=5,8, \end{cases}$$

$$\begin{cases} \mu_8^9(X,Y_i)=iY_i, & i\in\{1,2,3,4,5,6\},\\ \mu_8^9(X,Y_3')=3Y_3', & \\ \mu_8^9(Y_1,Y_i)=Y_{i+1}, & i=2,3,4,5,\\ \mu_8^9(Y_2,Y_i)=Y_{i+2}, & i=3,4,\\ \mu_8^9(Y_2,Y_3')=Y_5, & \mu_8^9(Y_3,Y_3')=Y_6, \end{cases}$$

$$\begin{cases} \mu_8^{10}(X,Y_i)=iY_i, & i=1,2,3,4,5,6,\\ \mu_8^{10}(X,Y_4')=4Y_4', & \\ \mu_8^{10}(Y_1,Y_i)=Y_{i+1}, & i=2,3,4,5,\\ \mu_8^{10}(Y_2,Y_i)=Y_{i+2}, & i=3,4,\\ \mu_8^{10}(Y_2,Y_4')=Y_6, & \end{cases}$$

$$\begin{cases} \mu_8^{11}(X,Y_i)=iY_i, & i=1,2,3,4,5,6,\\ \mu_8^{11}(X,Y_5')=5Y_5', & \\ \mu_8^{11}(Y_1,Y_i)=Y_{i+1}, & i=2,3,4,5,\\ \mu_8^{11}(Y_1,Y_5')=Y_6, & \\ \mu_8^{11}(Y_2,Y_3)=Y_5', & \mu_8^{11}(Y_2,Y_4)=Y_6, \end{cases}$$

$$\begin{cases} \mu_8^{12}(X,Y_i)=iY_i, & i=1,2,3,4,5,7,\\ \mu_8^{12}(X,Y_3')=3Y_3', & \\ \mu_8^{12}(Y_1,Y_2)=Y_3', & \mu_8^{12}(Y_1,Y_3')=Y_4,\\ \mu_8^{12}(Y_1,Y_4)=Y_5, & \\ \mu_8^{12}(Y_2,Y_3)=Y_5, & \mu_8^{12}(Y_2,Y_3')=Y_5,\\ \mu_8^{12}(Y_2,Y_5)=\mu_8^{12}(Y_3',Y_4)=Y_7, & \end{cases}$$

$$\begin{cases} \mu_8^{13}(X,Y_i)=iY_i, & i=1,2,3,4,5,7,\\ \mu_8^{13}(X,Y_5')=5Y_5', & \\ \mu_8^{13}(Y_1,Y_i)=Y_{i+1}, & i=2,3,4,\\ \mu_8^{13}(Y_2,Y_3)=Y_5', & \\ \mu_8^{13}(Y_2,Y_5)=\mu_8^{13}(Y_2,Y_5')=Y_7, & \mu_8^{13}(Y_3,Y_4)=-Y_7, \end{cases}$$

$$\begin{cases} \mu_8^{14}(X,Y_i)=iY_i, & i=1,3,4,5,6,7,\\ \mu_8^{14}(X,Y_4')=4Y_4', & \\ \mu_8^{14}(Y_1,Y_i)=Y_{i+1}, & i=3,4,5,6,\\ \mu_8^{14}(Y_3,Y_4)=\mu_8^{14}(Y_3,Y_4')=Y_7, & \end{cases}$$

$$\begin{cases} \mu_8^{15}(X,Y_i)=iY_i, & i=1,2,3,4,5, & \\ \mu_8^{15}(X,Y_3')=3Y_3', & & \\ \mu_8^{15}(X,Y_4')=4Y_4', & \mu_8^{15}(Y_1,Y_i)=Y_{i+1}, & i=2,3,4,\\ \mu_8^{15}(Y_1,Y_3')=Y_4', & \mu_8^{15}(Y_1,Y_4')=\mu_8^{15}(Y_2,Y_3)=Y_5, & \end{cases}$$

$$\begin{cases} \mu_8^{16}(X,Y_i)=iY_i, & i=1,2,3,4,5, & \\ \mu_8^{16}(X,Y_1')=Y_1', & & \\ \mu_8^{16}(X,Y_3')=3Y_3', & \mu_8^{16}(Y_1,Y_i)=Y_{i+1}, & i=2,4,\\ \mu_8^{16}(Y_1,Y_i')=Y_{i+1}, & i=1,3, & \\ \mu_8^{16}(Y_1',Y_i)=Y_{i+1}, & i=3,4, & \\ \mu_8^{16}(Y_1',Y_2)=Y_3', & \mu_8^{16}(Y_2,Y_3)=Y_5, & \\ \mu_8^{16}(Y_2,Y_3')=-Y_5, & & \end{cases}$$

$$\begin{cases} \mu_8^{17}(X, Y_i) = iY_i, & i = 1, 2, 3, 4, 5, \\ \mu_8^{17}(X, Y_i') = iY_i', & i = 3, 5, \\ \mu_8^{17}(Y_1, Y_i) = Y_{i+1}, & i = 2, 3, 4, \\ \mu_8^{17}(Y_1, Y_3') = Y_4, & \\ \mu_8^{17}(Y_2, Y_3) = Y_5', & \mu_8^{17}(Y_2, Y_3') = Y_5 + 2Y_5', \end{cases}$$

$$\begin{cases} \mu_8^{18}(X, Y_i) = iY_i, & \mu_8^{18}(X, Y_i') = iY_i', \quad i = 1, 2, 3, \\ \mu_8^{18}(X, Y_1'') = Y_1'', & \\ \mu_8^{18}(Y_1, Y_i') = Y_2, & \mu_8^{18}(Y_1, Y_1'') = Y_2', \\ \mu_8^{18}(Y_1'', Y_2') = Y_3', & \\ \mu_8^{18}(Y_1, Y_2') = Y_3', & \mu_8^{18}(Y_1', Y_2) = Y_3', \\ \mu_8^{18}(Y_1', Y_2') = \mu_8^{18}(Y_1'', Y_2) = Y_3, & \end{cases}$$

$$\begin{cases} \mu_8^{19}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 6, \\ \mu_8^{19}(Y_0, Y_i) = Y_i, & i = 2, 3, 4, 5, 6, \\ \mu_8^{19}(Y_1, Y_i) = Y_{i+1}, & i = 2, 3, 4, 5, \end{cases}$$

$$\begin{cases} \mu_8^{20}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 6, \\ \mu_8^{20}(Y_0, Y_i) = Y_i, & i = 3, 4, 5, 6, \\ \mu_8^{20}(Y_1, Y_i) = Y_{i+1}, & i = 3, 4, 5, \\ \mu_8^{20}(Y_2, Y_i) = Y_{i+2}, & i = 3, 4, \end{cases}$$

$$\begin{cases} \mu_8^{21}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 6, \\ \mu_8^{21}(Y_0, Y_i) = Y_i, & i = 4, 5, 6, \\ \mu_8^{21}(Y_1, Y_i) = Y_{i+1}, & i = 2, 4, 5, \\ \mu_8^{21}(Y_2, Y_4) = Y_6, & \end{cases}$$

$$\begin{cases} \mu_8^{22}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 6, \\ \mu_8^{22}(Y_0, Y_i) = Y_i, & i = 2, 3, 4, \\ \mu_8^{22}(Y_0, Y_i) = 2Y_i, & i = 5, 6, \\ \mu_8^{22}(Y_1, Y_i) = Y_{i+1}, & i = 2, 3, 5, \\ \mu_8^{22}(Y_2, Y_i) = Y_{i+2}, & i = 3, 4, \end{cases}$$

$$\begin{cases} \mu_8^{23}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 6, \\ \mu_8^{23}(Y_0, Y_6) = Y_6, & \\ \mu_8^{23}(Y_1, Y_i) = Y_{i+1}, & i = 2, 3, 4, \\ \mu_8^{23}(Y_2, Y_3) = Y_5, & \end{cases}$$

$$\begin{cases} \mu_8^{24}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 7, \\ \mu_8^{24}(Y_0, Y_i) = Y_i, & i = 3, 4, 5, \\ \mu_8^{24}(Y_0, Y_7) = 2Y_7, & \mu_8^{24}(Y_1, Y_i) = Y_{i+1}, \quad i = 3, 4, \\ \mu_8^{24}(Y_2, Y_3) = Y_5, & \mu_8^{24}(Y_3, Y_4) = Y_7, \end{cases}$$

$$\begin{cases} \mu_8^{25}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 7, \\ \mu_8^{25}(Y_0, Y_i) = Y_i, & i = 2, 3, 4, \\ \mu_8^{25}(Y_0, Y_5) = 2Y_5, & \mu_8^{25}(Y_0, Y_7) = 3Y_7, \\ \mu_8^{25}(Y_1, Y_i) = Y_{i+1} & i = 2, 3, \\ \mu_8^{25}(Y_2, Y_i) = Y_{i+2}, & i = 3, 5, \end{cases}$$

$$\begin{cases} \mu_8^{26}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 7, \\ \mu_8^{26}(Y_0, Y_i) = Y_i, & i = 2, 3, 4, 5, \\ \mu_8^{26}(Y_0, Y_7) = 2Y_7, & \mu_8^{26}(Y_1, Y_i) = Y_{i+1}, \quad i = 2, 3, 4, \\ \mu_8^{26}(Y_2, Y_5) = Y_7, & \mu_8^{26}(Y_3, Y_4) = -Y_7, \end{cases}$$

$$\begin{cases} \mu_8^{27}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 5, 6, 7, \\ \mu_8^{27}(Y_0, Y_i) = Y_i, & i = 2, 3, \\ \mu_8^{27}(Y_0, Y_5) = 2Y_5, & \mu_8^{27}(Y_0, Y_i) = 3Y_i, \quad i = 6, 7, \\ \mu_8^{27}(Y_1, Y_i) = Y_{i+1}, & i = 2, 6, \\ \mu_8^{27}(Y_2, Y_i) = Y_{i+2}, & i = 3, 5, \end{cases}$$

$$\begin{cases} \mu_8^{28}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, 7, \\ \mu_8^{28}(Y_0, Y_2) = Y_2, & \\ \mu_8^{28}(Y_0, Y_5) = 2Y_i, & i = 3, 4, \\ \mu_8^{28}(Y_0, Y_5) = 3Y_5, & \mu_8^{28}(Y_0, Y_7) = 4Y_7 \\ \mu_8^{28}(Y_1, Y_3) = Y_4, & \mu_8^{28}(Y_2, Y_i) = Y_{i+2}, \quad i = 3, 5, \\ \mu_8^{28}(Y_3, Y_4) = -Y_7, & \end{cases}$$

$$\begin{cases} \mu_8^{29}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, \\ \mu_8^{29}(X, Y_3') = 3Y_3', & \\ \mu_8^{29}(Y_0, Y_i) = Y_i, & i = 2, 3, 4, \\ \mu_8^{29}(Y_0, Y_3') = Y_3', & \\ \mu_8^{29}(Y_0, Y_5) = 2Y_5, & \mu_8^{29}(Y_1, Y_i) = Y_{i+1}, \quad i = 2, 3, \\ \mu_8^{29}(Y_1, Y_3') = Y_4, & \mu_8^{29}(Y_2, Y_3) = Y_5, \end{cases}$$

$$\begin{cases} \mu_8^{30}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, \\ \mu_8^{30}(X, Y_0') = 0, & \mu_8^{30}(Y_0, Y_5) = Y_5, \\ \mu_8^{30}(Y_0', Y_i) = Y_i, & i = 2, 3, 4, \\ \mu_8^{30}(Y_1, Y_i) = Y_{i+1}, & i = 2, 3, \end{cases}$$

$$\begin{cases} \mu_8^{31}(X, Y_i) = iY_i, & i = 0, 1, 2, 3, 4, 5, \\ \mu_8^{31}(X, Y_0') = 0, & \\ \mu_8^{31}(Y_0, Y_i) = Y_i, & i = 2, 3, \\ \mu_8^{31}(Y_0', Y_i) = Y_i, & i = 4, 5, \\ \mu_8^{31}(Y_1, Y_i) = Y_{i+1}, & i = 2, 4, \end{cases}$$

$$\begin{cases} \mu_8^{32}(X,Y_i)=iY_i, & i=0,1,2,3,4,5, \\ \mu_8^{32}(X,Y_0')=0, & \\ \mu_8^{32}(Y_0,Y_i)=Y_i, & i=2,4,5, \\ \mu_8^{32}(Y_0',Y_i)=Y_i, & i=3,4,5, \\ \mu_8^{32}(Y_1,Y_4)=Y_5, & \mu_8^{32}(Y_2,Y_3)=Y_5, \end{cases}$$

$$\begin{cases} \mu_8^{33}(X,Y_i)=iY_i, & i=0,1,2,3,4, \\ \mu_8^{33}(X,Y_0')=0, & \\ \mu_8^{33}(X,Y_0'')=0, & \mu_8^{33}(Y_0,Y_2)=Y_2, \\ \mu_8^{33}(Y_0',Y_3)=Y_3, & \mu_8^{33}(Y_0'',Y_4)=Y_4. \end{cases}$$

COMMENTS. The rigid Lie algebras play an important role in the description of irreducible components of the variety of Lie algebra laws (see the chapter "Varieties of Lie algebra laws" of this Handbook).

References for this subsection: [3,4,23,24,28].

References

[1] J.M. Ancochéa-Bermudez and M. Goze, *Classification des algèbres de Lie filiformes de dimension* 8, Arch. Math. **50** (1988), 511–525.

[2] J.M. Ancochéa-Bermudez and M. Goze, *Classification des algèbres de Lie nilpotentes de dimension* 7, Arch. Math. **52** (2) (1989), 157–185.

[3] J.M. Ancochéa-Bermudez and M. Goze, *Algèbres de Lie rigides dont le nilradical est filiforme*, C. R. Acad. Sci. Paris, Sér. I **312** (1991), 21–24.

[4] J.M. Ancochéa-Bermudez and M. Goze, *Le rang des systèmes linéaires d'une algèbre de Lie résoluble rigide*, Comm. Algebra **20** (3) (1992), 875–887.

[5] L. Auslander, *Simply transitive groups of affine motions*, Amer. J. Math. **99** (1977), 809–821.

[6] Y. Benoist, *Une nilvariété non affine*, J. Differential Geom. **41** (1) (1995), 21–52.

[7] N. Bourbaki, *Groupes et algèbres de Lie. Ch. I*, Hermann (1960).

[8] N. Bourbaki, *Groupes et algèbres de Lie. Ch. VII–VIII*, Hermann (1975).

[9] F. Bratzlavsky, *Une algèbre de Lie caractéristiquement nilpotente de dimension* 6, C. R. Acad. Sci. Paris, Sér. A **276** (1973), 1035–1037.

[10] F. Bratzlavsky, *Classification des algèbres de Lie nilpotentes de dimension n, de classe n − 1, dont l'idéal dérivé est commutatif*, Acad. Roy. Belg. Bull. Cl. Sci. **560** (1974), 858–865.

[11] R. Carles, *Sur les algèbres de Lie caractéristiquement nilpotentes*, Preprint No. 5, Univ. de Poitiers (1984).

[12] D. Burde and F. Grunewald, *Modules for certain Lie algebras of maximal class*, J. Pure Appl. Algebra **99** (1995), 239–254.

[13] J. Dixmier, *Sur les représentations unitaires de groupes de Lie nilpotentes III*, Canad. J. Math. **10** (1958), 321–348.

[14] J. Dixmier and W.G. Lister, *Derivations of nilpotent Lie algebras*, Proc. Amer. Math. Soc. **8** (1957), 155–158.

[15] J.L. Dyer, *A nilpotent Lie algebra with nilpotent automorphism group*, Bull. Amer. Math. Soc. **8** (1970), 155–158.

[16] G. Favre, *Une algèbre de Lie caractéristiquement nilpotente de dimension* 7, C. R. Acad. Sci. Paris, Sér. A **274** (1972), 1338–1339.

[17] G. Favre, *Système de poids sur une algèbre de Lie nilpotente*, Manuscripta Math. **9** (1973), 53–90.

[18] G. Favre and L.J. Santharoubane, *Nilpotent Lie algebras of maximal rank and of classical simple type*, Preprint.

[19] D. Fried, *Distality, completences and Affine structures*, J. Differential Geom. **24** (1986), 265–273.
[20] J.R. Gómez, M. Goze and Yu. Khakimdjanov, *On the k-Abelian filiform Lie algebras*, Comm. Algebra **25** (2) (1997), 431–450.
[21] J.R. Gómez, A. Jiménez-Merchan and Yu. Khakimdjanov, *Low-dimensional filiform Lie algebras* **30** (1998), 133–158.
[22] M. Goze and Yu. Khakimdjanov, *Nilpotent Lie Algebras*, MIA 361, Kluwer Acad. Publ., Dordrecht (1996).
[23] M. Goze and Yu. Khakimdjanov, *Sur les algèbres de Lie nilpotentes admettant un tore de dérivations*, Manuscripta Math. **84** (1994), 115–224.
[24] M. Goze and Yu. Khakimdjanov, *Some nilpotent Lie algebras and its applications*, Algebra and Operator Theory, Proceedings of the Colloquium in Tashkent, Kluwer Acad. Publ. (1998), 49–64.
[25] M. Goze and J.M. Ancochéa-Bermudez, *Algèbres de Lie rigides*, Algebra and Operator Theory, Proceedings of the Colloquium in Tashkent, Kluwer Acad. Publ. (1998), 65–91.
[26] G.B. Gurevich, *Standard Lie algebra*, Mat. Sb. **35** (77) (1954), 437–460.
[27] N. Jacobson, *Lie Algebras*, Intersc. Tracts 10, Wiley, New York (1962).
[28] N. Jacobson, *A note automorphisms and derivations of Lie algebras*, Proc. Amer. Math. Soc. **6** (1955), 281–283.
[29] Yu. Khakimdjanov, M. Goze and Sh. Ayupov (eds), Algebra and Operator Theory, Proceedings of the Colloquium in Tashkent, Kluwer Acad. Publ. (1998).
[30] Yu. Khakimdjanov, *Variété des lois d'algèbres de Lie nilpotentes*, Geom. Dedicata **40** (1991), 229–295.
[31] Yu. Khakimdjanov, *Characteristically nilpotent Lie algebras*, Mat. Sb. **181** (5) (1990), 642–655 (Russian). English transl. in Math. USSR Sbornik **70** (1) (1991).
[32] Yu. Khakimdjanov, *Standard subalgebras of reductive Lie algebras*, Vestnik Moskov. Univ. **4** (1974), 49–55. English transl. in Moscov Univ. Math. Bull. **29** (1–2) (1976).
[33] Yu. Khakimdjanov, *On the characteristically nilpotent Lie algebras*, Algebra i Logika **28** (6) (1989), 722–736.
[34] J. Kim, *Complete left-invariant affine structures on nilpotent Lie groups*, J. Differential Geom. **24** (1986), 373–394.
[35] J.L. Koszul, *Domaines bornés homogènes et orbites de groupes de transformations affines*, Bull. Soc. Math. France **89** (1961), 515–533.
[36] G. Leger, *A note on the derivations of Lie algebra*, Proc. Amer. Math. Soc. **4** (1953), 511–514.
[37] G. Leger and S. Tôgô, *Characteristically nilpotent Lie algebras*, Duke Math. J. **26** (4) (1959), 623–628.
[38] E. Luks, *What is the typical nilpotent Lie algebra*? Computers in Nonassociative Rings and Algebras, Academic Press (1977).
[39] A. Malcev, *On solvable Lie algebras*, Izv. Akad. Nauk SSSR Ser. Mat. **9** (1945), 329–356. English transl. in Amer. Math. Soc. Transl. No. 27 (1950).
[40] A. Medina, *Autour des connexions plates invariantes à gauche sur les groupes de Lie*, Thèse d'Etat, Montpellier, France (1979).
[41] A. Medina, *Flat left-invariant connections adapted to the automorphism structure of a Lie group*, J. Differential Geom. **16** (1981), 445–474.
[42] A. Medina and Ph. Revoy, *Groupes de Lie à structure symplectique invariante*, Séminaire Sud-Rhodanien, MSRI, Springer (1991), 247–266.
[43] J. Milnor, *On fundamental groups of complete affinely flat manifolds*, Adv. in Math. **25** (2) (1977), 178–187.
[44] V.V. Morosov, *Classification of nilpotent Lie algebras*, Izv. Vyssh. Uchebn. Zaved. **4** (1958), 161–171 (Russian).
[45] G.D. Mostow, *Fully reductible subgroups of algebraic groups*, Amer. J. Math. **78** (1956), 200–221.
[46] O.A. Nielsen, *Unitary representations and coadjoint orbits of low-dimensional nilpotent Lie groups*, Quenn's Papers in Pure Appl. Math. **63** (1983).
[47] A. Onishchik and Yu. Khakimdjanov, *Semidirect sums of Lie algebras*, Mat. Zametki **18** (1) (1975), 31–40 (Russian). English transl. in Math. Notes of Acad. Sci. USSR (1976), 600–604.
[48] M. Romdani, *Classification of real and complex nilpotent Lie algebras of dimension* 7, Linear and Multilinear Algebra **24** (1989), 167–189.
[49] E.N. Saffiulina, *On the Classification of Nilpotent Lie Algebras of Dimension* 7 (in Russian), Kazan (1976).
[50] C. Seeley, *Seven-dimensional nilpotent Lie algebras over the complex numbers*, PhD thesis, Chicago (1988).
[51] J.P. Serre, *Lie Algebras and Lie Groups*, W.A. Benjamin, New York (1965).

[52] J. Scheunemann, *Translations in certain groups of affine motions*, Proc. Amer. Math. Soc. **47** (1975), 223–228.

[53] K.A. Umlauf, *Über die Zusammensetzung der endlichen continuierlichen Transformationsgrouppen insbesondere der Grouppen vom Range null,* Thesis, Leipzig (1891).

[54] M. Vergne, *Cohomologie des algèbres de Lie nilpotentes. Application à l'étude de la variété des algèbres de Lie nilpotentes*, Bull. Soc. Math. France **98** (1970), 81–116.

[55] E.B. Vinberg, *Convex homogeneous cones*, Trans. Moscow Math. Soc. **12**, 340–403.

[56] S. Yamaguchi, *Derivations and affine structures at some nilpotent Lie algebras*, Mem. Fac. Sc. Kyushu Univ. **34** (1980), 151–170.

[57] S. Yamaguchi, *On some classes of nilpotent Lie algebras and their automorphism groups*, Mem. Fac. Sc. Kyushy Univ. **35** (1981), 341–351.

This Page Intentionally Left Blank

Section 5A
Groups and Semigroups

This Page Intentionally Left Blank

Infinite Abelian Groups: Methods and Results

A.V. Mikhalev and A.P. Mishina

Department of Mechanics and Mathematics, Moscow State University, Moscow, 119899 GSP Russia

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

0. Introduction

We say that a group is Abelian if the group operation (usually called addition) is commutative. The theory of Abelian groups can be considered as a part of the general theory of groups and also as a part of module theory over rings since every Abelian group is a module over the ring of integers. But at the same time the theory of Abelian groups now is an independent branch of algebra.

The beginning of the theory of Abelian groups dates back to the last century. In fact, we find some finite Abelian groups in the works of Gauss [46]. Frobenius and Stickelberger in 1878 ([7, Theorem 15.2]) proved that each finite Abelian group is a direct sum of cyclic groups of prime power orders. Thus, it was the first structure theorem in the history of Abelian group theory. Later on this theorem was generalized to all Abelian groups with a finite set of generators (Section 1, Point 1).

In succeeding years the theory of Abelian groups breaks down into three main parts (see [10]): torsion groups; torsion-free groups; mixed groups (Section 1, Point 3). But many problems considered now in the theory of Abelian groups are not related to exactly one of these parts. On the other hand, some of the methods developed first for one of these classes later were extended to larger classes of groups (and sometimes even to the class of all Abelian groups).

In the bibliography at the end of the paper the main monographs, survey papers and articles which have been cited in the text can be found.

In all what follows the word "group" means an Abelian group (with some exception for the automorphism groups of Abelian groups that can be non-Abelian). We shall not define the basic notions from general group theory (such as subgroups, quotient groups, homomorphisms, isomorphisms and so on) and also not the basic notions from set theory or other parts of algebra (for example, from ring theory).

We use the following standard notations:

Z and $\mathbb{Z}$, the group of integers (the infinite cyclic group) and the ring of integers, respectively;

Q and $\mathbb{Q}$, the group of rationals and the field of rationals, respectively;

Q_p (respectively, $\mathbb{Q}_p$), the group (the ring) of rationals with denominators prime to p;

J_p (respectively, $\mathbb{J}_p$), the group (the ring) of p-adic integers (see [7, Section 3]);

$\mathbb{K}_p$, the field of p-adic numbers;

tG (respectively G_p) the torsion part (the p-component) of G (see Section 1, Point 3);

$G[n]$ the set of all $g \in G$ with $ng = 0$

(in particular, $G[p] = \{g \in G \mid pg = 0\}$ is the *p-socle* of G);

nG the set of all ng, $g \in G$;

$|M|$ the cardinality of the set M;

$\langle M \rangle_*$ the pure subgroup (see Section 1, Point 5) of a torsion-free group G generated by subset $M \subseteq G$ (i.e. the intersection of all pure subgroups of G containing M), which is called the *pure hull* of M.

If A is a torsion-free group, then the dimension of the quotient group A/pA for a prime p regarded as a vector space over the prime field $\mathbb{F}_p$ is called the *p-rank* of A (notation: $r_p(A)$).

1. Some basic definitions

1. The cyclic groups, the groups of type p^∞ (p is a prime number), the groups isomorphic to Q or to its subgroups give us Abelian groups with the simplest structure.

We say that a group is *cyclic* if it consists of all multiples of one of its elements, say a (the notation for such a group is $\langle a\rangle$). If the order $o(a)$ of the element a is finite (i.e. there exists a least natural number $o(a) = n$ with $na = 0$), then all the distinct elements of $\langle a\rangle$ are $0, a, 2a, \dots, (n-1)a$. In this case we use the notation: $\langle a\rangle = Z(n)$. If $o(a) = \infty$ (i.e. $na \neq 0$ for all $n \neq 0$), then $0, \pm a, \dots, \pm na, \dots$ are all the distinct elements of $\langle a\rangle$.

Any infinite cyclic group is isomorphic to the group Z. If $o(a) = n$, then the group $\langle a\rangle$ is isomorphic to the group of all residue classes modulo n, i.e. it is isomorphic to the quotient group Z/nZ. Any subgroup, as well as any quotient group, of a cyclic group is also cyclic.

The *group of type* p^∞ (called also *quasicyclic*, notation: $Z(p^\infty)$) is the union of the ascending chain of its nonzero cyclic subgroups

$$\langle c_0\rangle \subset \cdots \subset \langle c_n\rangle \subset \cdots,$$

where $pc_0 = 0$, $pc_{i+1} = c_i$, $i = 0, 1, 2, \dots$. The group $Z(p^\infty)$ is isomorphic to the quotient group of the *group of p-fractions* (i.e. of all rational numbers with denominators powers of p) by the subgroup of integers. All *proper* subgroups of the group $Z(p^\infty)$ (i.e. subgroups different from zero and the whole group itself) are exhausted by the subgroups $\langle c_i\rangle$, $i = 0, 1, 2, \dots$.

If M is a subset of the group G such that every element $g \in G$ can be represented as $g = \sum k_i a_i$, where $a_i \in M$ and $k_i \in \mathbb{Z}$ (the number of summands in the sum for every $g \in G$ is finite), then we say that M is a *set of generators* for the group G or that G is *generated* by the subset M (notation: $G = \langle M\rangle$). Any group has a set of generators (we can take, for example, $M = G$). The set $M = \{c_0, c_1, \dots, c_i, \dots\}$ is a set of generators for the group $Z(p^\infty)$. One of the sets of generators for the cyclic group $\langle a\rangle$ is the element a.

2. A group A is called *sum* of its subgroups A_i (i running over an index set I, finite or infinite) if every element $a \in A$ is a (finite) sum of elements taken from distinct subgroups A_i:

$$a = \sum_{s=1}^{k} a_{i_s}, \quad a_{i_s} \in A_{i_s}.$$

We say that the group A is *direct sum* of its subgroups A_i if every element $a \in A$ has a unique representation of this kind. We use the following notation for the direct sum: $A = \bigoplus_{i\in I} A_i$ or $A = A_1 \oplus \cdots \oplus A_k$ if the number of summands is finite.

Every group with a finite number of generators (in other words, a *finitely generated group*, see above Point 1) can be decomposed into a direct sum of cyclic groups ([7, Theorem 15.5]).

If $A = \bigoplus_{i\in I} A_i$, then A_i are called *direct summands* of the group A. We call the group A *indecomposable* if it is not a direct sum of proper subgroups.

Examples of indecomposable groups: Z; Q; J_p; all subgroups of Q; pure (Point 5) subgroups of J_p ([11, Section 32]); any group $Z(p^k)$, where $k \leqslant \infty$ and p is prime number.

The direct decompositions $A = \bigoplus A_i = \bigoplus B_j$ of the group A are called *isomorphic* if there exists a bijective correspondence between the summands A_i and B_j such that corresponding groups are isomorphic. An example of isomorphic direct decompositions is given by different decompositions of a group as a direct sum of cyclic groups into direct sums of cyclic groups of infinite or prime power order ([7, Theorem 17.4]).

If $A = \bigoplus_{i \in I} A_i$ and $A_i = \bigoplus_j A_{ij}$ $(\forall i \in I)$ then $A = \bigoplus_{i,j} A_{ij}$; this direct decomposition is called a *refinement* of the original decomposition. If A is a countable reduced p-group (Points 3, 4) then any two of its direct decompositions have isomorphic refinements if and only if A is a direct sum of cyclic groups [11, Section 28] and Section 2, Point 1 below.

For arbitrary groups B_i $(i \in I)$ one can construct a group $A = \bigoplus_{i \in I} A_i$ with $A_i \cong B_i$ for all $i \in I$. We define the group A as the set of sequences $(\dots, b_i, \dots)$ with $b_i \in B_i$, with only a finite number of nonzero b_i in each sequence and componentwise addition. Thus, we obtain a group $A = \bigoplus_{i \in I} A_i$ where A_i consists of all sequences with elements from B_i at the i-th position, and zeros elsewhere. Evidently $A_i \cong B_i$. We call the group A the *outer direct sum* of the groups B_i with usual notation: $A = \bigoplus_{i \in I} B_i$. If for all $i \in I$ we have $B_i \cong B$, where B is a fixed group, then we write $\bigoplus_{i \in I} B_i = \bigoplus_{|I|} B$, where $|I|$ is the cardinality of the set I.

If we consider all possible sequences $(\dots, b_i, \dots)$ where $b_i \in B_i$ with componentwise addition, then we obtain the group B called the *direct product* of groups B_i (or the *complete direct sum*, or the *Cartesian sum*; notation: $B = \prod_{i \in I} B_i$). If $B_i \cong C$ for all $i \in I$ one can also write $B = \prod_{|I|} C = C^{|I|}$.

The subgroup of $\prod_{i \in I} B_i$ consisting of all sequences $(\dots, b_i, \dots)$ such that for any n almost all b_i satisfy $b_i \in nB_i$ is called a *regular direct sum* of B_i (by *almost all* we mean all with the possible exception of a finite number of elements).

3. If the orders of all elements of the group G are finite then G is called a *torsion group*. If the orders of all elements of G are bounded by some fixed natural number then G is called a *bounded group*. If the order of any element of the group G is a power of some fixed prime p then G is called a *primary group* (or more precisely *p-primary* or a *p-group*). A p-group G is called *elementary* if it is a direct sum of an arbitrary set of cyclic groups of order p.

All finite order elements of a group G form a subgroup tG called the *torsion part* of G. The set of all elements whose orders are powers of a given prime p forms a subgroup called the *p-component* G_p of G. Every torsion group is the direct sum of its p-components ($G = \bigoplus G_p$, see [7, Theorem 8.4]).

In case when all nonzero elements of a group G have infinite order G is called *torsion-free*. A direct sum $F = \bigoplus_{\alpha \in \mathfrak{A}} \langle a_\alpha \rangle$ of infinite cyclic groups $\langle a_\alpha \rangle$ is said to be a *free Abelian group*. In this case we say that the set of elements $\{a_\alpha\}$ is a *free set of generators* of the free group F.

Each subgroup of a free Abelian group is free ([7, Theorem 14.5]). The direct product $G = \prod Z$ of an infinite set of infinite cyclic groups is not a free Abelian group, although all subgroups of cardinality less than $\aleph_1$ are free ([7, Theorem 19.2]). Groups with this property are called $\aleph_1$*-free*. There exist indecomposable (Point 2) $\aleph_1$-free groups of cardinality $\aleph_1$ (R. Göbel and S. Shelah, 1994).

Every Abelian group G is a homomorphic image of a certain free Abelian group F. If $M = \{g_\alpha\}_{\alpha \in \mathfrak{A}}$ is a set of generators of G and $F = \bigoplus_{\alpha \in \mathfrak{A}} \langle a_\alpha \rangle$ is the corresponding free group then one of the possible homomorphisms is $\eta : F \to G$ where

$$\eta\left(\sum_{i=1}^{s} k_{\alpha_i} a_{\alpha_i}\right) = \sum_{i=1}^{s} k_{\alpha_i} g_{\alpha_i}, \quad k_{\alpha_i} \in Z.$$

The relation $k_1 g_{\alpha_1} + \cdots + k_s g_{\alpha_s} = 0$ holds in G if and only if $k_1 a_{\alpha_1} + \cdots + k_s a_{\alpha_s} \in \ker \eta = H$ (the kernel of the homomorphism η). If H is generated by a subset L then the system of relations corresponding to the elements of L is the *set of defining relations* of the group G. Each relation for elements of M results from the defining relations. The group G is determined by its set of generators and defining relations (in fact $G \cong F/H$). If there exists a set of generators and defining relations such that each defining relation contains at most two generators, we call G a *simply presented group* (see [22], [38, Section 7].

The group G is said to be *mixed* if it contains nonzero elements of both finite and infinite order.

4. If for an element a of a group A there exists $b \in A$ such that $a = p^n b$ for some prime p and some integer $n > 0$, then we say that a is *divisible* by p^n in A. In this case, if $a \neq p^{n+1} c$ for any $c \in A$, we say that n is the *p-height* of a in A: $h_p(a) = n$. If for all $n > 0$ the equation $a = p^n x$ has a solution in A, we say that a is an *element of infinite p-height*: $h_p(a) = \infty$. If a has infinite p-height for every prime p then the element a is divisible in A by any $k > 0$. If this property takes place for all elements of A the group A is called *divisible*.

Every divisible group is a direct sum of groups isomorphic to Q and/or groups of type p^∞ for some p ([7, Theorem 23.1]).

Every group G contains a greatest divisible subgroup (the sum of all divisible subgroups of G, see Point 2). We call this subgroup the *divisible part* dG of G. If $dG = 0$ the group G is called *reduced*. Each divisible group is a direct summand in every group, containing it as a subgroup. Thus for any group G we have $G = dG \oplus R$, where R is a reduced group, which is uniquely determined up to isomorphism ($R \cong G/dG$) (see [7, Section 21]).

A subgroup H of G is called an *absolute direct summand* if $G = H \oplus K$ for any subgroup K maximal with respect to $H \cap K = 0$. Let us note that dG is an absolute direct summand of G.

Each group G can be embedded as a subgroup in a divisible group D. Moreover, there exists a minimal divisible subgroup of D containing G. Any minimal divisible subgroup containing G, is uniquely determined up to *isomorphism over* G (this means that the isomorphism leaves invariant all the elements of G). Every such subgroup is called the *divisible hull* of G ([7, Section 24]).

If G is a p-group then each of its elements has infinite q-height for all prime $q \neq p$. Note that a reduced p-group can also contain elements with infinite p-height. For example, we may take the so called *Prüfer group* determined by generators $a_0, a_1, \ldots, a_n, \ldots$ and defining relations $pa_0 = 0$, $pa_1 = a_0$, $p^2 a_2 = a_0, \ldots, p^n a_n = a_0, \ldots$ ([7, Section 35]).

For any group G let

$$pG=\{g\in G\mid h_p(g)\geqslant 1\},\ p^2G=p(pG),\ \ldots,\ p^\omega G=\bigcap_{n=1}^{\infty}p^nG$$

and in general $p^{\alpha+1}G=p(p^\alpha G)$ for all ordinal numbers α and

$$p^\beta G=\bigcap_{\gamma<\beta}p^\gamma G$$

if β is a limit ordinal. Then $p^\omega G$ is the set of all elements of infinite p-height in G. It is evident that $G\supseteq pG\supseteq\cdots\supseteq p^\alpha G\supseteq\cdots$. If $g\in p^\alpha G\setminus p^{\alpha+1}G$ we call α the *generalised p-height* of the element g (notation: $h_p^*(g)$). There exists a first ordinal λ such that $p^\lambda G=p^{\lambda+1}G$. If G is a p-group then we say that this λ is the *length* of G, and in this case $p^\lambda G=dG$. If $g\in p^\lambda G$ then we set $h_p^*(g)=\infty$.

Let G be a p-group,

$$G^1=\bigcap_{n=1}^{\infty}p^nG=p^\omega G,\qquad G^{\sigma+1}=\left(G^\sigma\right)^1,\qquad G^\rho=\bigcap_{\sigma<\rho}G^\sigma$$

if ρ is a limit ordinal, and let τ be the least ordinal such that $G^\tau=G^{\tau+1}$. We say that the sequence $G=G^0\supset G^1\supset\cdots\supset G^\sigma\supset\cdots\supset G^\tau$ is the *Ulm sequence* for G, τ is its *Ulm type*, G^σ is its σ*-th Ulm subgroup*, $G_\sigma=G^\sigma/G^{\sigma+1}$ is its σ*-th Ulm factor*. All Ulm factors, except may be $G_{\tau-1}$ (if $\tau-1$ exists), are unbounded (Point 3) groups ([7, Lemma 37.3], [11, Section 27]).

5. We call a subgroup A of the group G *pure*, if $a=ng$ for $a\in A$, $g\in G$, n a natural number, implies $a=na_1$ for some $a_1\in A$ (i.e. $nA=A\cap nG$). The notion of purity was introduced by Prüfer in 1923 (see [7, Ch. V]) and since that time it is one of the most important notions in the theory of Abelian groups.

Examples of pure subgroups: any direct summand; any subgroup with torsion-free factor group (in particular, the torsion part of the group); the subgroup $\bigoplus_{i\in I}G_i$ of the group $\prod_{i\in I}G_i$ for arbitrary groups G_i, $i\in I$.

A subgroup A is pure in the group G if and only if every coset $g+A$, $g\in G$, contains an element of the group G with order equal to the order of the element $g+A$ in G/A ([7, Theorem 28.1]).

If A is a pure subgroup in G and A is bounded (Point 3) or G/A is a direct sum of cyclic groups then A is a direct summand of G ([7, Theorems 27.5 and 28.2]). At the same time every unbounded p-group G decomposable into a direct sum of cyclic groups contains a pure subgroup which is not a direct summand of G [11, Section 25].

If a group G is a direct summand of every group which contains it as a pure subgroup, then G is called an *algebraically compact* group. For different characterizations of algebraically compact groups and their structure see [7, Sections 38–40], [11, Section D.29, Point 5], [28, Section 5] and also Yu.L. Ershov [42]. One can describe the algebraically compact groups by means of invariants which are countable systems of cardinals [7, Section 40].

For a torsion-free reduced algebraically compact group we may take for such a system of invariants the set of p-ranks of the group (Section 0) for all primes p (these p-ranks can be arbitrary cardinals) [11, Section D.33, Point 4]. A torsion group is algebraically compact if and only if it is a direct sum of a divisible group and a bounded (Point 3) group ([7, Corollaries 38.3, 40.3]).

Every group A can be embedded as a pure subgroup into a minimal algebraically compact group ([7, Theorem 41.7]); this group is uniquely determined up to isomorphism over A (see Point 4).

If a group G is a direct summand of any group containing it as a subgroup with torsion-free quotient group then G is called *cotorsion*.

More information on cotorsion groups can be found in [7, Sections 54–56, 58], [11, Section D.33, Point 4] and also [27, Section 4], [28, Section 9], [29, Section 10], [30, Section 5], [31, Section 7], [26, Section 1, Point 5].

If there are no nonzero cotorsion subgroups in a group G then G is called *cotorsion-free*. Any such group is reduced and torsion-free.

6. A sequence of groups A_i and homomorphisms α_i

$$A_0 \xrightarrow{\alpha_0} A_1 \xrightarrow{\alpha_1} \cdots \xrightarrow{\alpha_{k-1}} A_k \quad (k \geqslant 1)$$

is called *exact* if $\operatorname{Im} \alpha_i = \ker \alpha_{i+1}$, $i = 0, \dots, k-2$. An exact sequence of the form

$$0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0 \tag{$*$}$$

is called a *short exact sequence*. The group G is called *injective* (*projective*) with respect to the given exact sequence $(*)$ if for any homomorphism $\varphi : A \to G$ (for any homomorphism $\psi : G \to C$, respectively) there is a homomorphism $\nu : B \to G$ (homomorphism $\mu : G \to B$, respectively), such that $\varphi = \nu\alpha$ (respectively, $\psi = \beta\mu$).

2. Primary groups

1. L.Ya. Kulikov ([7, Theorem 17.1]) obtained a criterion for decomposability of a p-group G into direct sum of cyclic groups, that gives as important corollaries (*the first and the second Prüfer theorems*, see [7, Theorems 17.2, 17.3] and [11, Section 24]):

(1) Any bounded group (Section 1, Point 3) is a direct sum of cyclic groups;

(2) A countable p-group is a direct sum of cyclic groups if and only if it does not contain nonzero elements of infinite height (i.e. p-height, see Section 1, Point 4).

In the uncountable case a p-group without (nonzero) elements of infinite height can be not a direct sum of cyclic groups (see [11, Section 26]). But p-groups without elements of infinite height (and among p-groups only they) have the property that every finite system of elements of the group is contained in some direct summand of it which is a direct sum of cyclic groups (such p-groups are called *separable*, see [7, Section 65]).

2. One of the most useful tools for studying p-groups is the notion of a basic subgroup ([7, Section 33], [11, Section 26]) introduced by L.Ya. Kulikov. A subgroup B of a p-group G is called *basic* if it is pure in G (Section 1, Point 5), is a direct sum of cyclic groups and is such that G/B is a divisible group (Section 1, Point 4). Any p-group contains a basic subgroup, and all basic subgroups of any p-group are isomorphic to each other. Let a p-group G without elements of infinite height have a basic subgroup $B = \bigoplus B_i$, where $B_i = \bigoplus Z(p^i)$ (i is fixed). Then the group G is isomorphic to a pure subgroup containing B of the torsion part $\bar{B}$ of the group $\prod B_i$; the converse is also true ([55], [11, Section 26]).

But so far there are no criteria for two pure subgroups of the group $\bar{B}$, containing $\bigoplus B_i$ to be isomorphic to each other. H. Leptin (1960) has shown that if $B_i = Z(p^i)$ $(i = 1, 2, \ldots)$ then $\bar{B}$ contains $2^{\aleph}$ non-isomorphic pure subgroups which contain $\bigoplus B_i$.

The groups of type $\bar{B} = t(\prod B_i)$ $(B_i = \bigoplus Z(p^i))$ are called *torsion complete* or *closed* p-groups ([7, Section 68], [11, Section D.30, Point 2]. Two torsion complete p-groups are isomorphic if and only if their basic subgroups are isomorphic. In such groups (and in the case of separable p-groups only those) every isomorphism between two basic subgroups can be extended to a (uniquely determined) automorphism of the whole group (see H. Leptin, 1960). A reduced p-group is a direct summand of every p-group which contains it as a pure subgroup if and only if it is a torsion complete p-group [7, Theorem 68.4].

3. For countable p-groups containing elements of infinite height there is the *Ulm theorem* [11, Section 28]: if two countable reduced p-groups G and H have the same Ulm type τ (Section 1, Point 4) and for all $\alpha < \tau$ the Ulm factors G_α and H_α of the groups G and H (see ibid.) are isomorphic, then $G \cong H$. Since in this case all Ulm factors are countable p-groups without elements of infinite height, i.e. direct sums of cyclic groups (Point 1) that are completely characterized by the sets of orders of their cyclic summands, we see that countable p-groups are characterized by numerical invariants.

L. Zippin proved the following *existence theorem* [11, Section 27]: if τ is a given ordinal of at most countable cardinality and for each α, $0 \leqslant \alpha < \tau$, a countable p-group A_α without infinite height elements is given, and, moreover, for all α, except, perhaps, $\alpha = \tau - 1$ (for nonlimit τ), the group A_α is unbounded, then there exists a countable reduced p-group of type τ, for which the sequence $A_0, A_1, \ldots, A_\alpha, \ldots$ $(\alpha < \tau)$ serves as the sequence of Ulm factors.

L.Ya. Kulikov and L. Fuchs independently generalised the Zippin theorem to the case of p-groups of arbitrary cardinality ([7, Theorem 105.3], [11, Section D.30, Point 3]). The analogue of the Ulm theorem for arbitrary uncountable p-groups does not hold.

I. Kaplansky and G.W. Mackey (see [10]) gave a new proof and a partly new formulation of the Ulm theorem. For a reduced p-group G they considered the socles $P_\alpha = (p^\alpha G)[p]$ of the subgroups $p^\alpha G$. By the first Prüfer theorem the group $P_\alpha / P_{\alpha+1}$ is a direct sum of cyclic groups of order p. It can be considered as a vector space over the p-element field $\mathbb{F}_p$. The dimension $\dim_{\mathbb{F}_p} (P_\alpha / P_{\alpha+1}) = f_G(\alpha)$ is called the *α-th Ulm–Kaplansky invariant* of the group G (or its *α-th Ulm invariant*).

The *Ulm–Kaplansky theorem* says that two countable reduced p-groups are isomorphic if and only if they have the same Ulm invariants.

This theorem can be applied also to groups without elements of infinite height. If G is a direct sum of cyclic groups, then $f_G(n)$ $(n = 0, 1, 2, \ldots)$ is the number of cyclic direct summands of order p^{n+1} in the decomposition of G.

As to the existence of countable reduced Abelian p-groups with fixed Ulm invariants see [10, Section 11, Exercises 36, 42].

For uncountable groups the Ulm Kaplansky theorem is not true in the general case (see [10, Section 11, Exercise 33]). But it can be extended to some classes of groups. G. Kolettis (1960) has proved such a theorem for direct sums of countable reduced p-groups. G. Kolettis also found necessary and sufficient conditions for the existence of direct sums of countable reduced p-groups with given Ulm invariants.

The largest additive (i.e. closed under direct sums) class of reduced Abelian p-groups in which groups are uniquely determined by their Ulm–Kaplansky invariants, is the class of totally projective groups. This class contains the class of all direct sums of countable reduced p-groups, see [7, Theorem 82.4]). The original definition of totally projective groups was given by R.J. Nunke ([7, Section 82]): a reduced p-group A is called *totally projective* if $p^\sigma \operatorname{Ext}(A/p^\sigma A, C) = 0$ (see Section 8, Part II) for all ordinals σ and any group C. Some equivalent definitions can be found in [7, Section 83]. In particular, a reduced p-group is totally projective if and only if it is simply presented (Section 1, Point 3) (see [7, Theorem 83.5]), and if and only if it satisfies *the third axiom of countability with respect to nice subgroups* (R. Göbel and S. Shelah, 1994). This means that A contains a collection $\mathcal{C}$ of *nice subgroups* C (i.e. such that $p^\alpha(A/C) = (p^\alpha A + C)/C\ \forall\alpha$) with the properties: (1) $0 \in \mathcal{C}$; (2) if $\{C_i\}_{i\in I} \subseteq \mathcal{C}$ then $\sum_{i\in I} C_i \in \mathcal{C}$; (3) for every countable subgroup $B \subseteq A$ there exists a countable $C \in \mathcal{C}$ such that $B \subseteq C$ ([7, Section 81,Theorem 82.3]; see also [9] and P. Hill and W. Ullery, 1996). The third axiom of countability with respect to nice subgroups is called *Hill's condition* by R.B. Warfield, Jr. [37].

The existence of totally projective p-groups with fixed Ulm–Kaplansky invariants is discussed in [7, Theorem 83.6]). Totally projective groups are also considered in Section 1 of [26–31].

R. Göbel [20] has given a review of results, connected with the Ulm theorem and its generalizations.

4. P. Hill (1985) considered a class of p-groups G, stricly containing the class of totally projective groups. He called them *A-groups* and found a system of cardinal invariants containing the Ulm–Kaplansky invariants and determining A-groups up to isomorphism. The corresponding existence theorem was also proved.

Each A-group H is an *isotype subgroup* of some totally projective group G (i.e. $H \cap p^\alpha G = p^\alpha H$ for all α), see P. Hill and Ch. Megibben (1985). Thus, Warfield's question about the existence of a reasonable structure theory describing a sufficiently large class of isotype subgroups of totally projective groups has a positive answer.

Some other information on A-groups can be found in [21].

5. We say that a p-group G from a given class $\mathfrak{C}$ is determined in $\mathfrak{C}$ by its socle $G[p]$ or *p^n-socle* $G[p^n] = \{g \in G \mid p^n g = 0\}$ if for all $H \in \mathfrak{C}$ the existence of an isomorphism $\varphi : G[p] \to H[p]$ (an isomorphism $\psi : G[p^n] \to H[p^n]$) that preserves height (Section 2, Point 1) in G and H respectively, implies $G \cong H$.

One class of p-groups in which the groups are determined by their socles is the smallest class containing all groups $Z(p^n)$ $(n = 1, 2, \ldots)$ and closed under arbitrary direct sums, direct summands and *t-products* $\prod_{i \in I}^{t} G_i$ of group families $\{G_i\}_{i \in I}$ where I is a nonmeasurable [7, Section 94]) set, $\prod_{i \in I}^{t} G_i$ is the torsion part of the group $\prod_{i \in I} G_i$ (see P. Keef, 1992). Moreover in the above mentioned class each height preserving isomorphism of $G[p]$ and $H[p]$ can be extended to an isomorphism $G \cong H$.

But if G is a pure subgroup of a torsion complete group $\bar{B}$ (Section 2, Point 2) and G is neither a direct sum of cyclic groups, nor a torsion complete group, and the generalized continuum hypothesis GCH holds, then there exists a pure subgroup $H \subseteq \bar{B}$ such that $H[p] = G[p]$ but H and G are not isomorphic (S. Shelah, 1987).

Let G be a dense in the p-adic topology (with base $\{p^k G\}_{k=1,2,\ldots}$) pure subgroup of a separable p-group (Section 2, Point 1). Then the question if there exists a non-isomorphic to G pure subgroup with the same p^n-socle is undecidable under ZFC (Zermelo–Frenkel axioms and the axiom of choice), see D.O. Cutler (1989).

For any $n > 0$, D.O. Cutler (1987) has constructed an example of a separable p-group determined by its p^{n+1}-socle in the class of all p-groups but not determined by its p^n-socle.

6. For other results concerning primary Abelian groups see [7, Ch. XI, XII] and Section 1 of the reviews [26–31].

3. Torsion-free groups

1. In the theory of torsion-free groups the notion of rank is essential. We define the *rank*, $r(G)$, of a torsion-free group G as cardinality of a maximal linearly independent (over the ring of integers) system of elements of G.

Torsion-free groups of rank 1 have the simplest structure theory. Each such group is isomorphic to some subgroup of the group Q (Section 0), and vice versa.

Let R be a torsion-free group of rank 1. For $0 \neq a \in R$ let us consider the sequence $\chi(a) = (k_1, \ldots, k_i, \ldots)$ of p_i-heights k_i of the element a in the group R (Section 1, Point 4), where p_i runs over all primes enumerated in a fixed order. We say that the sequence $\chi(a)$ is the *characteristic* of a. Two characteristics are said to be *equivalent*, if they have symbol ∞ at the same places and, moreover, these characteristics can differ at most at a finite number of positions (i.e. they coincide *almost everywhere*). The set of all characteristics equivalent to one of them is called a *type*. The characteristics of all nonzero elements of a torsion-free group R of rank 1 constitute a type called the *type of* R (notation: $\tau(R)$). Two torsion-free groups of rank 1 are isomorphic if and only if they have the same type. Let G be a torsion-free group and $0 \neq a \in G$. Then the set of all elements $b \in G$ *depending* on a (i.e. there exist nonzero integers k and l such that $ka = lb$) coincides with the set of all elements of the pure subgroup $\langle a \rangle_*$ of G generated by a. If $A = \langle a \rangle_*$ then $r(A) = 1$, and the type $\tau(A)$ is called the *type of the element* a of G.

One sais that $\tau_1 \geqslant \tau_2$ for the types τ_1, τ_2 (or that τ_2 divides τ_1) if there exist characteristics $\chi_1 = (k_1, \ldots, k_i, \ldots) \in \tau_1$ and $\chi_2 = (l_1, \ldots, l_i, \ldots) \in \tau_2$ such that $k_i \geqslant l_i$, $i = 1, 2, \ldots$. If G is a finite rank torsion-free group then the g.c.d. of the types of all its pure subgroups of rank 1 is called the *inner type* of G (notation: $IT(G)$) and l.c.m. of the types of all rank 1 torsion-free quotients of G is called its *outer type* (notation: $OT(G)$). If $IT(G) = OT(G)$ then G is a direct sum of rank 1 groups of type $OT(G)$ ([63]). If G is a torsion-free group of infinite rank then always $OT(G) = \infty$ (i.e. the type containing the characteristic $(\infty, \infty, \ldots)$) and $IT(G)$ can be not defined.

2. If G is a direct sum of torsion-free rank 1 groups then G is called *completely decomposable*. Every two decompositions of such a group into direct sum of rank 1 groups are isomorphic; any direct summand of a completely decomposable group is also completely decomposable ([7, Proposition 86.1 and Theorem 86.7]). A.A. Kravchenko (1982) has found necessary and sufficient conditions for a completely decomposable group

$$G = \bigoplus R_\alpha, \quad r(R_\alpha) = 1,$$

to have all its pure subgroups completely decomposable (the conditions are on the types of the groups R_α). This result completes a series of investigations on this topic (see L. Bican (1974), S.F. Kozhukhov (1974), M.I. Kushnir (1974)). For groups of finite rank see also L.G. Nongxa and C. Vinsonhaler (1995) (the conditions are on the types of all the elements of G).

3. Pure subgroups of completely decomposable finite rank torsion-free groups are called *Butler groups*. Equivalently: a torsion-free group is called a Butler group if it is an epimorphic image of a completely decomposable finite rank torsion-free group. These groups were introduced by M.C. Butler (1965). There exist several other definitions for them equivalent to the one given above, see [30, Sections 2, 5], [31, Section 6].

Some classes of Butler groups were described by invariants up to isomorphism (D. Arnold and C. Vinsonhaler (1992)) or up to quasi-isomorphism (Section 6), see D. Arnold and C. Vinsonhaler (1989, 1993), L. Fuchs and C. Metelli (1991).

Additional information on Butler groups can be found in [30, Section 2], [31, Section 6], [15,17], A.V. Yakovlev (1995).

A torsion-free group G (of arbitrary rank) is called B_1*-group* (and sometimes Butler group or B-group) if $\operatorname{Bext}(G, T) = 0$ (Section 8, Part II, Point 7) for any torsion group T.

A countable torsion-free group is a B_1-group if and only if all its pure subgroups of finite rank are Butler groups (L. Bican and L. Salce (1983)); groups with the last property are called *finitely Butler*. In particular, the class of all B_1-groups of finite rank coincides with the class of all Butler groups of finite rank.

A *homogeneous* torsion-free group (i.e. such that all its nonzero elements have the same type) is a B_1-group if and only if it is completely decomposable (L. Bican (1984)). A pure subgroup of a B_1-group G may be not a B_1-group (ibid), but it does have to be a B_1-group if the set of all different types of elements of G is countable (see [17]).

For a survey of the results obtained for B_1-groups see [14,17,18] and [31, Section 6].

A torsion-free group G is called a B_2*-group* (in some articles: a *Butler group of infinite rank*) if there exists a well ordered increasing sequence of pure subgroups

$$0 = G_0 \subset G_1 \subset \cdots \subset G_\alpha \subset \cdots \subset G_\tau = G = \bigcup G_\alpha,$$

where $G_\alpha = \bigcup_{\beta<\alpha} G_\beta$ if α is a limit ordinal and $G_{\alpha+1} = G_\alpha + B_\alpha$ for some Butler group B_α of finite rank ($\forall\, \alpha < \tau$).

Every B_2-group is a B_1-group. For groups of cardinality $\leqslant \aleph_1$ the converse also holds (see [17]). On the relationships between B_1 and B_2-groups see also L. Bican (1994) and K.M. Rangaswary (1994).

Every almost completely decomposable torsion-free group (Point 4 below) is a B_2-group, see U. Albrecht and P. Hill (1987).

For more about B_2-groups see [17,18], [30, Section 5], [31, Section 6].

4. A torsion-free group G is called *almost completely decomposable* if there exists an $n > 0$ such that $nG \subseteq \bigoplus A_i \subseteq G$, where the A_i are subgroups of rank 1. In this case if G has a finite rank then $|G/nG| < \aleph_0$, and the subgroup $\bigoplus A_i$ has finite index in G. For more on almost completely decomposable groups see [24], Section 4 of [29], [30,31], E.A. Blagoveshchenskaya and A. Mader (1994), A.V. Yakovlev (1995).

For torsion-free groups of finite rank S.F. Kozhukhov (1980) obtained necessary and sufficient conditions for a group to be almost completely decomposable.

5. A torsion-free group is called *separable* if every finite system of elements of it is contained in a completely decomposable direct summand of the group. For information on such groups see [7, Sections 87, 96] and [31, Section 5]. Note that Theorem 96.6 from [7] does not hold (see A.P. Mishina (1962)).

Many generalizations of the notion of a separable torsion-free group were given in Section 4 of [28], [29,30] and [31, Section 5]. Note that some authors call a torsion-free group separable if every finite system of elements is contained in a free direct summand.

6. Let P_0 be a direct product of a countable number of infinite cyclic groups:

$$P_0 = \prod_{i=1}^{\infty} \langle e_i \rangle, \qquad o(e_i) = \infty.$$

A torsion-free group G is called *slender* if for every homomorphism $\eta : P_0 \to G$ we have $\eta(e_i) = 0$ for almost all i (Section 3, Point 1). A torsion-free group is slender if and only if it does not contain subgroups isomorphic to the groups Q, P_0, or J_p, for an arbitrary prime p ([7, Theorem 95.3]).

If $P = \prod_{i \in I} A_i$ where all the A_i are arbitrary torsion-free groups, and G is a slender group then for every homomorphism $\eta : P \to G$ for almost all $i \in I$ we have $\eta(A_i) = 0$, moreover, if the cardinality $|I|$ is non-measurable [7, Section 94] and $\eta(A_i) = 0$ for all $i \in I$ then $\eta = 0$ ([7, Theorem 94.4], a result of J. Łoś)).

Direct sums of slender groups ([7, Theorem 94.3]) are slender groups.

For more about slender groups see also [30, Section 4], [31, Section 4], [7, Sections 94, 95]. In the literature one also finds the dual notions of *coslender* groups and *dual slender* groups ([31, Section 4]).

7. For any infinite cardinal $\mathfrak{m}$ there exists $2^{\mathfrak{m}}$ non-isomorphic torsion-free indecomposable groups of cardinality $\mathfrak{M}$ (see S. Shelah (1974) and also [19, Section 8]). Note that the indecomposable torsion groups are only the groups $Z(p^k)$ ($k \leqslant \infty$, p prime) and there are no mixed indecomposable groups at all ([7, Corollary 27.3]).

Every torsion-free group G of finite rank is either indecomposable or the direct sum of a finite number of indecomposable groups. But different decompositions of a group G into direct sum of indecomposable groups may be non-isomorphic (Section 1, Point 2). Moreover, for arbitrary integers $n > k > 1$ one can find a torsion-free group A of rank n such that for any partition of n into k parts, $n = r_1 + \cdots + r_k$, $r_i \geqslant 1$, $i = 1, \ldots, k$, there is a direct decomposition $A = A_1 \oplus \cdots \oplus A_k$ where every A_i is an indecomposable group of rank r_i ([7, Theorem 90.2]). But every torsion-free group of finite rank may have only a finite number of non-isomorphic direct summands (E.L. Lady, 1974). This gives a solution of Problem 69 from [7, Ch. XIII].

Other examples of "bad decompositions" of a torsion-free group into direct sum of indecomposable subgroups can be found in: [7, Sections 90, 91], [11, Section D.31, Point 3]; V.A. Kakashkin (1984), E.A. Blagoveshchenskaya (1992) (two solutions to Problem 67 from [7, Ch. XIII]); E.A. Blagoveshchenskaya and A.V. Yakovlev (1989) (a solution to Problem 68 from [7, Ch. XIII]).

But for torsion-free groups of finite rank, a theorem that replaces the theorem on isomorphism of direct decompositions can be obtained if instead of isomorphism of groups one considers on quasi-isomorphisms (see Section 6, Point 1).

8. For a description of torsion-free groups from some classes of groups by means of invariants, see Section 5, Points 2, 3, 4 of the present article.

Some additional information on the topics mentioned above and on other questions from the theory of torsion-free groups can be found in [1,7, Ch. XIII], [11, Ch. 8 and Section D.31] and also [27, Sections 2, 6], [28, Sections 2, 4], [29, Sections 2, 4], [30, Sections 2, 4], [31, Sections 2, 4, 5, 6].

4. Mixed groups

1. If a mixed group G *splits*, i.e. $G = tG \oplus H$ for some subgroup H, then the description of the structure of G can be in some sense reduced to the description of the torsion group tG and the torsion-free group $H \cong G/tG$. But, for example, the group $G = \prod \langle a_i \rangle$ for a_i of order p_i and p_i running over all distinct primes, has torsion part $tG = \bigoplus \langle a_i \rangle$ and does not split (see [7, Section 100]).

All groups G with the torsion part isomorphic to a given torsion group T split if and only if T is a direct sum of a bounded (Section 1, Point 3) group and a divisible one (Section 1, Point 4) group, see [7, Theorem 100.1], [11, Section 29].

All groups G with $G/tG \cong H$ for a given torsion-free group H split if and only if H is a free Abelian group (see [7, Section 101]).

Different criteria for the splitting of mixed groups can be found in [11, Section D.32] and in Section 3 of the surveys [26–31].

2. Some classes of nonsplittig mixed groups can be described by means of invariants. For example, let G be a group such that $r(G/tG) = 1$ ($r(G/tG)$ is called the *torsion-free rank* of G). For an element $g \in G$ of infinite order one can introduce the notion of the *height matrix* $\mathbb{H}(g)$, the rows of this matrix are indexed by the primes p and the row with the index p consists of the generalized p-heights of the elements $g, pg, \dots, p^n g, \dots$ (see [7, Section 103]). A notion of equivalence of such matrices was determined and it was shown that two countable groups of torsion-free rank 1, are isomorphic if and only if their torsion parts are isomorphic and their height matrices are equivalent ([7, Theorem 104.3]). In the uncountable case this statement does not hold (Ch. Megibben, 1967).

A theorem about the existence of a mixed group with a fixed torsion part T and $r(G/T) = 1$, and containing an infinite order element with a given height matrix also holds (the height matrices have to satisfy suitable necessary and sufficient conditions), see [7, Theorem 103.3] with a correction in R. Hunter, E. Walker (1981); see also T. Koyama (1978).

For mixed groups of torsion-free rank 1 I. Kaplansky and G.W. Mackey obtained a result analogous to the Ulm theorem for p-groups (see [20]).

A group G is called *p-local* (p a prime) if the multiplication by any prime $q \neq p$ is an automorphism of G, i.e. if G can be considered in a natural way as a $\mathbb{Q}_p$-module (Section 0). Any p-group is p-local.

L.Ya. Kulikov [56] investigated mixed p-local groups that he called *generalized primary groups*. He considered cases in which Ulm theorem holds for such groups.

3. R.B. Warfield, Jr. suggested a new approach to investigation of mixed groups: instead of considering such a group as an extension (Section 8, Part II) of a torsion group by a torsion-free group he regarded it as an extension of a torsion-free group by a torsion group (see [36]).

Namely, R.B. Warfield, Jr. (1972) called a p-local group G (Point 2) a *λ-elementary KT-module* (λ is a limit ordinal) if $p^\lambda G \cong Q_p$ and $G/p^\lambda G$ is a totally projective p-group (see Section 2, Point 3). The direct sum of a totally projective group and a set of λ-elementary KT-modules (for different λ) is called a *KT-module*. For a KT-module M and an arbitrary limit ordinal μ let $h(\mu, M) = \dim_{\mathbb{F}_p}(p^\mu M/(p^{\mu+1}M + t(p^\mu M)))$ (R.B. Warfield, 1975). Two KT-modules M_1 and M_2 are isomorphic if and only if their Ulm–Kaplansky invariants are the same (i.e. $f(\alpha, M_1) = f(\alpha, M_2)$, where $f(\alpha, M_i) = \dim_{\mathbb{F}_p}((p^\alpha M_i)[p]/p^{\alpha+1}M_i[p]), \alpha \geqslant 0)$ and $h(\mu, M_1) = h(\mu, M_2)$ for any limit ordinal μ. The class of all KT-modules is a maximal, closed under finite direct sums, class of $\mathbb{Q}_p$-modules in which the above mentioned statement holds. There are also conditions for the functions f and h for which there exists a KT-module M such that $f(\alpha, M) = f(\alpha)$ and $h(\mu, M) = h(\mu)$ for arbitrary α and limit ordinal μ.

KT-modules are exactly the reduced p-local h-pure projective groups (see R.B. Warfield, Jr., 1972) (a group is called *h-pure projective* if it is projective (Section 1, Point 6) with respect to any exact sequence of $\mathbb{Q}_p$-modules $0 \to A \to B \to C \to 0$ such that the sequence $0 \to p^\alpha A \to p^\alpha B \to p^\alpha C \to 0$ is exact for every ordinal α).

A survey of new methods, as well as of basic results on classification of mixed groups, has been published by R.B. Warfield, Jr. [37]. There he has formulated also 27 problems; some of them have been already solved, see [38, Section 7] for problems 11, 13, 16, 17, 18, 19, 23, and A.A. Kravchenko (1984) for problems 6, 7.

4. For global (i.e. not necessary p-local) groups the next notion was introduced by J. Rotman (1960) (see also R.O. Stanton, 1977): a set $X = \{x_i\}_{i \in I}$ of elements of an Abelian group A is called a *decomposition basis* for A if $\{x_i\}_{i \in I}$ is a set of free generators of the subgroup $\langle X \rangle$, $A/\langle X \rangle$ is torsion and for arbitrary $a = \sum r_i x_i \in \langle X \rangle$ and any p, $h_p^*(a) = \min\{h_p^*(r_i x_i)\}$ for generalized p-heights h_p^* (Section 1, Point 4) taken in A (compare the definition of direct sum of valuated groups in Section 10). The group A is called a *Warfield group* (see R. Hunter and F. Richman, 1981) if it contains a decomposition basis X which generates a *nice* subgroup $\langle X \rangle$ (a subgroup B of a mixed group A is called nice in A if $(p^\alpha(A/B)/((p^\alpha A + B)/B))[p] = 0$ for any ordinal α and prime p, see [22]), and, moreover, $A/\langle X \rangle$ is a simply presented group. Warfield groups are exactly the direct summands of simply presented groups.

One can give a number of equivalent definitions of a Warfield group A (see [22]). In particular: A has a decomposition basis and satisfies the third axiom of countability (Section 2, Point 3) with respect to nice subgroups. These definitions generalize different characterizations of totally projective groups [7, Section 83].

R. Hunter and F. Richman (1981) introduced the notion of a global Warfield invariant. Two Warfield groups G and H are isomorphic if and only if their Warfield invariants are the same and their Ulm invariants are equal for any prime p.

For an existence theorem for p-local Warfield groups see [29, Section 10].

5. Ch. Megibben (1980) (see also K.M. Rangaswamy, 1984) calls an Abelian group *completely decomposable* if it is a direct sum of torsion-free groups of rank 1 and groups $Z(p^k)$, $k \leqslant \infty$ (compare Section 3, Point 2). A group is said to be *separable* if every finite set of its elements is contained in some completely decomposable (in the previous sense) direct summand of this group. One can prove (Ch. Megibben, 1980) that a mixed group G is separable if and only if tG and G/tG are separable (Section 2, Point 1 and Section 3, Point 5) and tG is a balanced subgroup in G (Section 8, Part II, Point 7).

V.M. Misyakov (1991) gave a criterion for separability of direct products of arbitrary Abelian groups.

5. Classification theorems

The Ulm theorem (Section 2, Point 3) gives us a full description of all countable (and even all totally projective) p-groups by means of numerical invariants. Some other group classes also can be described in terms of invariants. However, a description of a class of groups by means of invariants is reasonable only if it helps to solve some problems concerning groups of this class. For example, the Ulm theorem for totally projective p-groups helps us to get a positive solution for the following two Kaplansky test problems (see [7, Section 83, ex. 11, 12]):

(i) Do conditions $G = G_1 \oplus G_2$, $H = H_1 \oplus H_2$, $G \cong H_1$, $H \cong G_1$ imply $G \cong H$?
(ii) Is the condition $G \cong H$ necessary for $G \oplus G \cong H \oplus H$?

1. A.V. Ivanov [48] considered groups $G = \bigoplus_{i=1}^{\infty} G_i$, where the G_i are torsion-complete (Section 2, Point 2) p-groups. Let us associate to a group G the matrix (a_{ik}), where the

$a_{ik} = f_k(G_i)$ are the Ulm invariants of the group G_i (which determine this group, see [7, Section 68, Point b] and [10, Section 11]). This matrix is called the *decomposition matrix* of the group G. In the class of such matrices one can introduce two equivalence relations: $(a_{ik}) \sim (b_{ik})$ and $(a_{ik}) \approx (b_{ik})$. We can prove that if $G = \bigoplus_{i=1}^{\infty} G_i = \bigoplus_{i=1}^{\infty} H_i$ are different decompositions of the group G into direct sums of torsion-complete p-groups, and (a_{ik}), (b_{ik}) are the associated decomposition matrices then $(a_{ik}) \sim (b_{ik})$ and $(a_{ik}) \approx (b_{ik})$. We say that the equivalence class with respect to $\sim$ (to $\approx$) of matrices of decompositions of G into countable direct sums of torsion-complete p-groups is the *first* (the *second*) *invariant* of the group G. Two countable direct sums of torsion-complete p-groups are isomorphic if and only if their 1st and 2nd invariants and their Ulm invariants coincide. By means of this criterion we get a positive solution of the Kaplansky test problems for countable direct sums of torsion complete p-groups. One can show that therefore these problems have positive solutions also for arbitrary direct sums of torsion-complete groups (see A.V. Ivanov, [49]).

The same theory (with the same results) can be developed for countable direct sums $G = \bigoplus_{i=1}^{\infty} G_i$ of reduced algebraically compact groups, see [26, Section 5, Point 1].

Let us consider the class $\mathcal{A}$ of all countable direct sums of torsion-complete p-groups (respectively, of reduced algebraically compact groups) and let $G, H \in \mathcal{A}$. In terms of the invariants mentioned above we can formulate conditions for the group G to be isomorphic to a pure subgroup of the group H and also describe *pure-correct* groups $G \in \mathcal{A}$, i.e. such that if G and $H \in \mathcal{A}$ are isomorphic to pure subgroups of each other then $G \cong H$. Moreover, we can obtain necessary and sufficient conditions on the decomposition matrix for the group G to be *cancellable* in $\mathcal{A}$ (this means that for all $A, B \in \mathcal{A}$ from $G \oplus A \cong G \oplus B$ it follows $A \cong B$), see [48].

2. A.G. Kurosh (1937), D. Derry (1937) and A.I. Maltsev (1938) have described in terms of invariants all finite rank torsion-free groups. L. Fuchs has extended this theory to arbitrary countable torsion-free groups, see [7, Section 93].

The description was obtained in the following way. Let us fix a prime p. A countable torsion-free group A can be embedded into a group

$$A_p^* = J_p \otimes A = \bigoplus_n \mathbb{K}_p v_n \oplus \bigoplus_m \mathbb{J}_p w_m,$$

where $\otimes$ means tensor product (see Section 9) and n, m run over index sets of cardinalities k_p and l_p respectively, $k_p + l_p = r(A)$. Let $\{a_i\}$ be a maximal independent system of elements of A. Then

$$a_i = \sum_n \alpha_{in} v_n + \sum_m \beta_{im} w_m \quad (\alpha_{im} \in \mathbb{K}_p,\ \beta_{im} \in \mathbb{J}_p).$$

Let M_p be the matrix with elements that are coefficients from the last equalities. We construct such a matrix M_p for each prime p. Let now the matrix sequence $M_{p_1}, M_{p_2}, \ldots$ corresponds to A (its terms depend on the choice of $\{a_i\}$ and $\{v_n, w_m\}$). Now determine the connection between matrices M_p and M'_p, obtained by different systems $\{a_i\}$ and $\{v_n, w_m\}$.

If the elements of matrix sequences $(M_{p_1}, M_{p_2}, \ldots)$ and $(M'_{p_1}, M'_{p_2}, \ldots)$ are related in the above mentioned way, then these sequences are called equivalent. The cardinalities k_p, l_p (for all p) and the class of equivalent matrix sequences are a *complete set of invariants* for countable torsion-free groups. This means that two groups of this type are isomorphic if and only if the corresponding invariants coincide ([7, Theorem 93.4]). Also the corresponding existence theorem holds ([7, Theorem 93.5]).

This theory enabled A.G. Kurosh ([11, Section 32 c]) to prove the existence of indecomposable torsion-free groups of any finite rank.

If A is a finite rank torsion-free group then the number k_p for A coincides with the number of summands of type p^∞ in a decomposition into indecomposable summands of the divisible part of A/C, where C is a subgroup generated by a maximal independent set of elements of A (L. Procházka, 1962).

M.N. O'Campbell (1960) obtained a description of countable torsion-free groups by means of invariants in which he used equivalence classes of sequences of special type integer valued square matrices (the order of matrices is equal to the rank of the group). Using this language he formulated necessary and sufficient conditions for a countable torsion-free group to be free. In the same terms M.I. Kushnir (1972) gave necessary and sufficient conditions for a torsion-free group of finite rank to be quotient divisible (a torsion-free group G is called *quotient divisible* if it contains a free subgroup F such that G/F is a divisible torsion group).

However, the above methods to describe countable torsion-free groups have not yet found a sufficiently large field of applications. A.V. Yakovlev (1976) has pointed out the complexity of the classification problem even for finite rank torsion-free groups.

The I. Kaplansky test problems in the class of countable torsion-free groups have a negative answer (see L.S. Corner, 1964).

3. A description of one class of torsion-free groups was given by A.A. Fomin.

Let $\prod_p \mathbb{K}_p$ be the direct product of the fields of p-adic numbers $\mathbb{K}_p$. The set of all sequences (α_p), $\alpha_p \in \mathbb{K}_p$, with almost all α_p p-adic integers form a subring $\mathbb{K}_p$ of $\prod_p \mathbb{K}_p$. The subring $\mathbb{K}$ can be considered as a linear space over the field $\mathbb{Q}$ of rationals. A.A. Fomin [45] calls $\mathbb{K}_p$ *the ring of universal numbers* and $\prod_p \mathbb{J}_p$ ($\mathbb{J}_p$ is the ring of p-adic integers) *the ring of universal integers*. If (α_p) is an universal integer then we denote by h_p the greatest power of p dividing α_p in $\mathbb{J}_p$ (if $\alpha_p = 0$, then $h_p = \infty$). The type determined by the characteristic (h_p) is called *the type of the universal integer* (α_p).

For an arbitrary universal number α there exists a non-zero integer m such that $m\alpha$ is a universal integer. Let $\operatorname{type}(\alpha) = \operatorname{type}(m\alpha)$. It is easy to see that the type of α does not depend on the choice of m.

The set of all universal numbers which types are greater or equal to a fixed type τ, forms an ideal $\mathbb{I}_\tau$ of the ring $\mathbb{K}_p$; $\mathbb{Q}(\tau) = \mathbb{K}/\mathbb{I}_\tau$ is called the *ring of τ-adic numbers*. If the type τ is defined by the characteristic (m_p), then

$$\mathbb{Q}(\tau) = \mathbb{Q} \otimes \prod_p \mathbb{Z}_{p^{m_p}},$$

where $\mathbb{Z}_{p^{m_p}}$ is the residue class ring modulo p^{m_p} if $m_p \neq \infty$, and $\mathbb{Z}_{p^{m_p}} = \mathbb{J}_p$, if $m_p = \infty$; here $\otimes$ denotes the tensor product (Section 9). If α is an universal integer then we call $\alpha + \mathbb{I}_\tau$ an *integer τ-adic number*.

Let $\tau_1, \ldots, \tau_m$ be some types. A.A. Fomin considers certain finite dimensional subspaces of the linear space $M = \mathbb{Q}(\tau_1) \oplus \cdots \oplus \mathbb{Q}(\tau_m)$ over $\mathbb{Q}$ which he calls *$(\tau_1, \ldots, \tau_m)$-spaces*.

For a torsion-free group A of finite rank n we define its *Richman type* as the class of quasi-isomorphisms (Section 6) of the quotient group A/F, where F is a free subgroup of rank n in A. Since

$$A/F \cong \bigoplus_p \left(Z(p^{i_{1p}}) \oplus \cdots \oplus Z(p^{i_{np}})\right), \quad 0 \leqslant i_{1p} \leqslant \cdots \leqslant i_{np} \leqslant \infty,$$

the Richman type is determined by the sequence of types, containing the characteristics $(i_{1p})_{p\in\pi}, \ldots, (i_{np})_{p\in\pi}$ (with π the set of all primes). If some of these types are zero (i.e. we have $0, \ldots, 0, \tau_1, \ldots, \tau_m$, where $0 < \tau_1 \leqslant \cdots \leqslant \tau_m$, and 0 is the type of the group Z), then A.A. Fomin calls the sequence of types $\tau_1, \ldots, \tau_m$ the *reduced Richman type* of the group A ([45]). Groups of reduced Richman type τ are called *groups with one τ-adic relation* [44].

A.A. Fomin [45] described torsion-free finite rank groups of given Richman type that do not contain free direct summands (those groups are called *coreduced*). For this purpose he exploited $(\tau_1, \ldots, \tau_m)$-spaces as invariants. This classification was obtained up to quasi-isomorphisms (Section 6). It was shown that the category of $(\tau_1, \ldots, \tau_m)$-spaces is equivalent to the category of coreduced torsion-free finite rank groups with given reduced Richman type $\tau_1 \leqslant \cdots \leqslant \tau_m$, where instead of homomorphisms we take quasi-homomorphisms from the group A into the group B (*quasi-homomorphisms* are defined as elements of the divisible hull of $\operatorname{Hom}(A, B)$ (Section 8, Part I).

A.A. Fomin (1986) described also (up to quasi-isomorphisms) in terms of τ-adic invariants all pure-free groups (we say that a group is *pure-free* if each of its nontrivial pure subgroup is free). Every pure-free torsion-free group which is not free has finite rank and is strongly indecomposable (Section 6, Point 1). Applying the above mentioned invariants A.A. Fomin described endomorphism groups and quasi-endomorphism rings for pure-free groups (Section 7, Part I, Points 4, 5).

4. Another class of finite rank torsion-free groups (A, T-groups) was described by S.F. Kozhukhov in his doctoral dissertation, see also Section 4 in [30] and [31], where particular cases of (A, T)-groups ((T, m)-groups, (A, P)-groups, (A, p, r)-groups) are discussed.

A finite rank torsion-free group G is said to be (A, T) group, if there exists an uniquely determined subgroup $A = \bigoplus A_i$ of G where the A_i are strongly indecomposable (Section 6, Point 1) pure subgroups of G and $\operatorname{Hom}(A_i, A_j) = 0$ if $i \neq j$ (for Hom see Section 8, Part I) and $G/A = T$ is finite. It is made clear, when the class of (A, T)-groups is not empty.

With each (A, T)-group it is possible to associate a well defined set of rectangular matrices with integer coefficients which is called the (A, T)-type of G; one can establish the relations between two (A, T)-types corresponding to isomorphic groups (one gets a complete set of invariants (Point 2) for the groups under consideration). In terms of (A, T)-types one can obtain necessary and sufficient condition for an (A, T)-group to be decomposable into a direct sum of nontrivial subgroups. In one particular case of (A, T)-groups S.F. Kozhukhov gives a description of their structure and also considers their automorphism groups (Section 7, Part II, Point 1).

5. For classification theorems for some classes of mixed groups see Section 4, Points 2, 3, 4.

6. Quasi-isomorphisms

1. We say that two Abelian groups A and B are *quasi-isomorphic* (notation $A \sim B$) if there exist subgroups $A' \subseteq A$ and $B' \subseteq B$ such that $A' \cong B'$ and A/A', B/B' are bounded groups. In case of torsion-free groups A and B this is equivalent to the following property: one of the groups, for example A, is isomorphic to a subgroup $B' \subseteq B$ with B/B' bounded. Thus, almost completely decomposable groups (Section 3, Point 4) are exactly the torsion-free groups quasi-isomorphic to completely decomposable torsion-free groups. Two finite rank torsion-free groups A and B are quasi-isomorphic if and only if each of them is isomorphic to a subgroup of the other ([7, Section 92, Exercise 5], [51]). Groups with this property are called *equivalent* (or in [11, Section D.35, Point 3] nearly isomorphic). One can find in [12, p. 67] an example of completely decomposable torsion-free groups A and B, non-isomorphic but *purely equivalent* (i.e. $A \cong B' \subseteq B$, $B \cong A' \subseteq A$, where A' and B' are pure subgroups of the corresponding groups). This gives a negative solution to problem 28 from [6, Ch. VII] and as well shows that the corresponding statements from [6, Ch. VII, Exercise 26] and [6, Ch. VIII, Exercise 10] are wrong. Moreover, this example and results by L.I. Vlasova (see Point 3 below) show that infinite rank torsion-free groups can be equivalent but not quasi-isomorphic.

2. A torsion-free group G is said to be *strongly indecomposable* if there is no torsion-free group quasi isomorphic to G and decomposable into a direct sum of two nontrivial subgroups. B. Jonsson [51] proved that if the groups $A_1, \dots, A_m, B_1, \dots, B_n$ are strongly indecomposable torsion-free groups of finite rank and the direct sums $A_1 \oplus \cdots \oplus A_m$ and $B_1 \oplus \cdots \oplus B_n$ are quasi-isomorphic, then $m = n$ and (under a corresponding numeration) the summands A_i and B_i are quasi-isomorphic $(i = 1, \dots, n)$.

We say that a subgroup A of the group G is a *quasi-summand* of G if there exists a subgroup $B \subseteq G$ such that $G/(A \oplus B)$ is bounded. If all pure subgroups of a finite rank torsion-free group G are quasi-summands then G is a completely decomposable group (Section 3, Point 2) with a linearly ordered set of types of its elements, and vice versa (see E. Muller and O. Mutzbauer, 1992).

3. R.A. Beaumont and R.S. Pierce (1964) constructed an example of two p-groups with isomorphic basic subgroups (Section 2, Point 2), which are quasi-isomorphic, but non-isomorphic. But in some special classes of groups quasi-isomorphic groups are necessarily isomorphic, for example, in the class of complete decomposable torsion-free groups (see L.I. Vlasova, 1998).

But a torsion-free completely decomposable group can be quasi-isomorphic to a torsion-free group which is not completely decomposable (L. Procházka, 1965).

A splitting mixed group can be quasi-isomorphic to a non-splitting group (E.A. Walker, 1964). This gives an answer to a question formulated by L. Fuchs.

4. Two torsion groups G and H are quasi-isomorphic if and only if their p-components satisfy the following condition: G_p and H_p are quasi-isomorphic for all prime p and, moreover, $G_p \cong H_p$ for almost all p (R.A. Beaumont and R.S. Pierce, 1964).

If two p-groups G and H are quasi-isomorphic then necessarily $p^{\omega}G \cong p^{\omega}H$, and the groups $G/p^{\omega}G$ and $H/p^{\omega}H$ are quasi-isomorphic (the converse does not hold, see D.O. Cutler, 1966). If G is a direct sum of cyclic p-groups then any p-group H quasi-isomorphic to G is also a direct sum of cyclic groups (L. Fuchs, 1952). R.A. Beaumont and R.S. Pierce (1965) obtained necessary and sufficient conditions for two direct sums of cyclic p-groups to be quasi-isomorphic (these conditions are formulated in terms of the Ulm–Kaplansky invariants); these conditions are also necessary for two p-groups to be quasi-isomorphic, and they are sufficient for countable p-groups (see R.A. Beaumont and R.S. Pierce, 1964).

5. E.L. Lady (1975), see also [1], calls two finite rank torsion-free groups A, B *near-isomorphic* ($A \cong_n B$) if for any integer $k \neq 0$ there exists an integer m with $(k, m) = 1$ and homomorphisms $f: A \to B$ and $g: B \to A$ such that $gf = m \cdot 1_A$ (1_A is the identity isomorphism on A), $fg = m \cdot 1_B$. If $A \cong_n B$ then A and B can be not isomorphic but are quasi-isomorphic ([1, Corollary 6.2 and Section 7]). If $A \cong_n X \oplus Y$ then $A = X' \oplus Y'$ with $X' \cong_n X$, $Y' \cong_n Y$ (Arnold theorem: see [1, Corollary 12.9]). For nore on near-isomorphism see also [24, Sections 6, 7], D.M. Arnold (1995) and A. Mader and C. Vinsonhaler (1997).

6. Some more information on quasi-isomorphisms can be found in [27, Section 8], [28, Section 6], [29, Section 7], [30, Section 7], [31, Section 8], [7, Section 92], [11, Section D.35, Point 3].

7. Rings and groups of endomorphisms. Automorphism groups

I. Let G be an Abelian group. If we add endomorphisms of G according to the rule of addition of homomorphisms (Section 8, Part I) and multiply two endomorphisms by composition, then we obtain an associative ring with identity, which is called the *endomorphism ring* $E(G)$ of G. Examples of endomorphism rings can be found in [7, Section 106], [11, Section 21].

1. If the groups A and C are isomorphic then the group isomorphism $\varphi: A \to C$ induces a ring isomorphism $\varphi^{\sharp}: E(A) \to E(C)$ determined by the formula $\varphi^{\sharp}: \eta \to \varphi\eta\varphi^{-1}$.

The *Baer–Kaplansky theorem* holds [7, Section 108]: if A and C are torsion groups with isomorphic rings of endomorphisms then the groups A and C are isomorphic; moreover, any ring isomorphism $\psi: E(A) \to E(C)$ is induced by a certain group isomorphism $\varphi: A \to C$.

The Baer–Kaplansky theorem does not hold for the class of all Abelian groups (see [7, Section 106, Examples 3 and 5]). But it is valid in some classes of groups, see [61].

We say that the group A is determined by its endomorphism ring in the class L of Abelian groups if from $E(A) \cong E(B)$, where $B \in L$, it follows that $A \cong B$. Concerning the determination of groups by their endomorphism rings see [29, Section 7], [30, Section 7], [31, Section 9] and S.T. Files (1994).

May and Toubassi (1977, 1981) found conditions under which a mixed group with torsion-free rank one (Section 4, Point 2) is determined by its endomorphism ring in the class of these groups. This gives a negative solution to the Fuchs problem 87 ([7, Ch. XV]). For a survey of results on isomorphisms of endomorphism rings of mixed groups, see May [25].

2. Any associative ring A with identity such that its additive group A^+ is (1) countable, (2) reduced, (3) torsion-free, is isomorphic to the ring of endomorphisms of a countable reduced torsion-free group. Moreover, if $r(A^+) = n$ then A is isomorphic to the endomorphism ring of some reduced torsion-free group of rank $2n$ (*Corner theorem*, see [7, Theorems 110.1, 110.2 and Exercise 7 from Section 110].

Note that each ring A of the class under consideration can be realized as an endomorphism ring by $2^{\aleph_0}$ non-isomorphic groups G of the corresponding class ([41]). However, there exist rings A that satisfy any two, but not the third, of the three conditions for A^+ of being (1) countable, (2) reduced, and (3) torsion-free, and that are not realized as endomorphism rings of any groups (see [41] and the examples $\mathbb{J}_p \oplus \mathbb{J}_p$, $\mathbb{Q} \oplus \mathbb{Q}$, $\mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z}$).

W. Liebert (see [7, Theorem 109.1]), gave necessary and sufficient conditions for an associative ring with identity to be isomorphic to the ring of endomorphisms of a separable p-group (Section 2, Point 1). Unlike the case of torsion-free groups, a rather special class of rings is obtained in the last case.

3. For relations between properties of the group G and properties of its endomorphism ring see [7, Ch. XV, Sections 107, 111, 112], Section 7 in [28], [29,30] and [31, Section 9]).

In particular U. Albrecht (1986) established when the endomorphism ring of the group A is a principal ideal domain.

Necessary and sufficient conditions for commutativity of the endomorphism ring of a mixed group have been found by Ph. Schultz (1973).

P.A. Krylov (1979) gives necessary and sufficient condition on the group G under which the ring $E(G)$ is isomorphic to a subring of the field of rationals. A.V. Ivanov ([50]) finds a full description of all groups with left (right) selfinjective rings of endomorphisms.

4. The same questions as for the ring of endomorphisms $E(G)$ of the group G can be posed for its additive group $\operatorname{End} G$ (called the *endomorphism group* of G), see [29, Section 7], [30, Section 7], [31, Section 9].

If A and C are p-groups with isomorphic groups of endomorphisms then A and C can be non-isomorphic (see [11, Section D.34, Point 2]). J.W. Armstrong (1963) found conditions under which isomorphism of the endomorphism groups $\operatorname{End} A$ and $\operatorname{End} C$ implies $A \cong C$, but only in the case when A and C are direct sums of cyclic p-groups. S.Ya. Grinshpon (1973) assuming the general continuum hypothesis GCH obtained necessary and sufficient condition on a p-group A under which $\operatorname{End} A \cong \operatorname{End} C$ implies $A \cong C$ for an arbitrary p-group C (see Problem 41 in [6, Ch. IX]).

There are torsion-free groups with isomorphic groups of endomorphisms but with non-isomorphic rings of endomorphisms, see E. Sąsiada (1959) and [41] (Problem 42 in [6, Ch. IX]).

Results concerning the determination of a group in the given class of groups by its *multiplicative semigroup of endomorphisms* can be found in [30, Section 7] and A.M. Sebel'din (1991, 1994, 1995).

5. Let G be a torsion-free group. Then it can be embedded as a subgroup into a linear space V over the field $\mathbb{Q}$ of rationals (the divisible hull of G) in such a way that the quotient V/G is torsion. We say that a linear transformation φ of the linear space V is a *quasi-endomorphism* of the group G if for some natural number n we have $n\varphi(G) \subseteq G$ (R.A. Beaumont and R.S. Pierce, [2]). The set of all quasi-endomorphisms of the group G forms a ring with natural addition and multiplication. This ring is isomorphic to $\mathbb{Q} \otimes E(G)$ (tensor product, see Section 9).

Any algebra of finite dimension n over the field of rationals is isomorphic to the ring of quasi-endomorphisms of some torsion-free group of rank $2n$ ([41]).

P.A. Krylov (1979) considered relations between properties of a torsion-free group G of finite rank, its ring of endomorphisms $E(G)$ and its ring of quasi-endomorphisms.

In a series of papers the rings of quasi-endomorphisms for some finite rank torsion-free groups were described (see D.M. Arnold and C. Vinsonhaler (1989, 1991), A.A. Fomin (1986), [44]).

II. The set of all automorphisms of a (not necessary Abelian) group G forms a group with respect to natural multiplication operation as composition of automorphisms. This group is denoted by $\operatorname{Aut} G$ and is called the *automorphism group* of G. In the case of an Abelian group G it is the group of invertible elements of the endomorphism ring $E(G)$.

1. The automorphism group of an Abelian group is not necessarily Abelian. For example, if G is an Abelian p-group then $\operatorname{Aut} G$ is an Abelian group if and only if either $G \cong Z(p^k)$, $0 \leqslant k \leqslant \infty$, p prime, or $G \cong Z(2^\infty) \oplus Z(2)$ ([7, Corollary 115.2]).

If G is an infinite torsion Abelian group, then $|G| \leqslant |\operatorname{Aut} G|$, see E.A. Walker (1960). But there are some infinite Abelian groups with finite automorphism groups (Example: $\operatorname{Aut}(Z) \cong Z(2)$). If for a torsion-free group G the group $\operatorname{Aut} G$ is finite then for each infinite cardinal $\mathfrak{m}$ there are $2^{\mathfrak{m}}$ non-isomorphic Abelian torsion-free groups G' of cardinality $\mathfrak{M}$ for which $\operatorname{Aut}(G') \cong \operatorname{Aut}(G)$ ([3, Theorem 6.12]).

2. A torsion group and a torsion-free group can have isomorphic groups of automorphisms (Example: Z and $Z(3)$). But for Abelian p-groups ($p \geqslant 5$) H. Leptin [58] has proved that if $\operatorname{Aut}(G) \cong \operatorname{Aut}(H)$, then $G \cong H$. The same conclusion for any $p \geqslant 3$ has been obtained later by W. Liebert [59]. But not every isomorphism between automorphism groups is induced by an isomorphism of the groups; example: $G \cong Z(p)$, $H \cong Z(p)$ (for any $p \geqslant 5$) due to A.V. Ivanov, see [26, Section 7, Part II, Point 2]. This provides a negative solution to the problem 89 from Fuchs [7, Ch. XVI]. For reduced p-groups G, H ($p \geqslant 3$) with isomorphic groups of automorphisms W. Liebert (1989) gave a description of all isomorphisms between $\operatorname{Aut}(G)$ and $\operatorname{Aut}(H)$ which is analogous to the description of isomorphisms of the classical linear groups.

3. There are infinitely many non-isomorphic countable Abelian torsion-free groups G for which $\operatorname{Aut}(G)$ is a countable divisible Abelian group. But there is no such countable divisible Abelian group of finite torsion-free rank (Section 4, Point 2) that can be realized as an automorphism group for any group (not necessarily Abelian), see M.R. Dixon, M.J. Evans (1990).

A survey of results on automorphism groups of Abelian torsion-free groups obtained in the last decades was given (with proofs) in the book by I.Kh. Bekker and S.F. Kozhukhov [3].

For automorphism groups of Abelian groups see also [7, Ch. XVI], Section 9 in [27], [31] and Section 7 in [28], [29,30].

8. Homomorphism groups. Groups of extensions

I. Let us define the *group* $\operatorname{Hom}(A, B)$, *of homomorphisms* from an Abelian group A to an Abelian group B as the set of all homomorphisms $\alpha : A \to B$, with addition according to the rule $(\alpha_1 + \alpha_2)a = \alpha_1 a + \alpha_2 a$, for any $a \in A$. In particular, $\operatorname{Hom}(A, A) = \operatorname{End} A$ is the group of endomorphisms (Section 7, Part I, Point 4) of the group A.

1. Some introductory information on the groups $\operatorname{Hom}(A, B)$ can be found in [7, Section 43].

In particular, for any Abelian groups we have

$$\operatorname{Hom}\Bigl(A, \prod_{i\in I} B_i\Bigr) \cong \prod_{i\in I} \operatorname{Hom}(A, B_i), \quad \operatorname{Hom}\Bigl(\bigoplus_{i\in I} A_i, B\Bigr) \cong \prod_{i\in I} \operatorname{Hom}(A_i, B).$$

A group A is called *selfsmall* if

$$\operatorname{Hom}\Bigl(A, \bigoplus_{\mathfrak{m}} A\Bigr) \cong \bigoplus_{\mathfrak{m}} \operatorname{Hom}(A, A)$$

for any cardinal number $\mathfrak{m}$. Concerning these groups see D.M. Arnold and C.E. Murley (1975). In particular, any selfsmall torsion group is finite.

2. If A is a divisible group or B is a torsion-free group, then $\operatorname{Hom}(A, B)$ is torsion-free ([7, Section 43, F, C]). If A (or B) is a torsion-free divisible group then $\operatorname{Hom}(A, B)$ is also torsion-free divisible [7, Section 43, G, D].

If A is a torsion group then $\operatorname{Hom}(A, B)$ is a reduced algebraically compact group (Section 1, Point 5) for any group B ([7, Theorem 46.1]). If B is an algebraically compact group, then for any group A the group $\operatorname{Hom}(A, B)$ is also algebraically compact ([7, Theorem 47.7]).

The structure of the group $\operatorname{Hom}(A, B)$ in some particular cases is considered in [7, Section 47], Section 8 in [28], [29,30] and [31, Section 10].

The group $\operatorname{Hom}(A, D/Z)$ where D is the additive group of all real numbers is called the *character group* of the group A [7, Section 47].

3. For the question when a given group is isomorphic to a group Hom(A,B), as well as when $A \cong \mathrm{Hom}(A, B)$ or $A \cong \mathrm{Hom}(B, A)$, see respectively R.B. Warfield, Jr. [63] and A.M. Sebel'din (1973, 1974), P. Grosse (1965). In particular, from the results of P. Grosse it follows that $\mathrm{Hom}(A, Q/Z) \cong A$ if and only if A is finite (see also F.R. Beyl and A. Hanna, 1976).

4. Note the following important result: there exist two non-isomorphic groups A and A' such that $\mathrm{Hom}(A, X) \cong \mathrm{Hom}(A', X)$ for any group X; examples:

(i) $A = \bigoplus_{2^{\aleph_0}} B$, $A' = \bigoplus_{2^{\aleph_0}} \bar{B}$ with $B = \bigoplus_{i<\omega} \bigoplus_{\aleph_0} Z(p^i)$ (p prime) and $\bar{B}$ beeing the corresponding torsion complete group (Section 2, Point 2), see P. Hill (1971), and

(ii) $A = (\bigoplus_{\aleph_0} Z) \oplus Q$, $A' = A \oplus Q$, see A.M. Sebel'din (1974).

This gives a negative solution to Fuchs problem 34 ([7, Ch. VIII]). But if A and B are completely decomposable torsion-free groups, A is of finite rank and $\mathrm{Hom}(A, C) \cong \mathrm{Hom}(B, C)$ for any torsion-free group C of rank 1, then $A \cong B$ (see L.I. Vlasova, 1979).

If A and B are arbitrary reduced torsion-free groups of finite rank then there exist torsion-free groups $C_1, \ldots, C_n$ of finite rank such that quasi-isomorphism (Section 6) of the groups $\mathrm{Hom}(A, C_i)$ and $\mathrm{Hom}(B, C_i)$, $i = 1, \ldots, n$, implies quasi-isomorphism of the groups A and B ([52, Corollary 1.12]).

5. Let G be an arbitrary group, $G^* = \mathrm{Hom}(G, Z)$, $G^{**} = (G^*)^*$. The group G is called *reflexive* if the *natural mapping* $\sigma : G \to G^{**}$ (under which $\sigma(g)(g^*) = g^*(g)$ for all $g \in G$, $g^* \in G^*$) is an isomorphism. If the natural mapping $G \to \mathrm{Hom}(\mathrm{Hom}(G, A), A)$ for groups G and A is an isomorphism then the group G is called *A-reflexive*.

R.B.Warfield, Jr. [63] found the class $LF(\tau)$ of all torsion-free groups G of finite rank which are R-reflexive for a given torsion-free group R of rank 1 and type τ; see also P.A. Krylov (1979). Thus in the category $LF(\tau)$ the functor $\mathrm{Hom}(-, R)$ is a duality. This duality (called *Warfield duality*) is a generalization of the duality for linear spaces over a field.

D.M. Arnold (1972) has constructed a duality in the category QD of quotient divisible (Section 5, Point 2) torsion-free groups of finite rank with quasi-homomorphisms as morphisms (Section 5, Point 3) (the so-called *Arnold duality*). For this purpose he used the results of R.A. Beamont and R.S. Pierce (1961) describing groups in QD by means of invariants (that are subspaces of p-adic linear spaces of dimension equal to the rank of the group).

A.A. Fomin ([43]) constructed (for every type τ) a duality in the category D of all torsion-free groups A of finite rank in each of which there is a free subgroup F such that A/F is a torsion group any p-component of which is either a divisible group or a reduced group (morphisms in this category are quasi-isomorphisms). The class D of groups contains the class QD and all classes $LF(\tau)$. This *Fomin duality* restricted to $LF(\tau)$ coincides with Warfield duality, and for the type $\tau \ni (\infty, \infty, \ldots)$ coincides with Arnold duality on the class QD. C. Vinsonhaler and W. Wickless (1990) constructed a duality on classes of torsion-free groups of finite rank which also generalises both Warfield and Arnold duality; a special case of this duality is Fomin duality.

6. We say that a non-empty class of groups is *torsion* if it is closed under extensions (see Part II below), homomorphic images and direct sums, and is *torsion-free* if it is closed under extensions, subgroups and direct products. A pair $(\mathfrak{T}, \mathfrak{F})$ of classes of groups is called a *torsion theory* if $\mathfrak{T}$ is a torsion class and $\mathfrak{F}$ is the greatest torsion-free class with $\mathrm{Hom}(T, F) = 0$ for arbitrary $T \in \mathfrak{T}$ and $F \in \mathfrak{F}$ (see M. Dugas and G. Herden, 1983).

Let $\mathfrak{X}$ be an arbitrary class of groups. Then

$$\mathfrak{X}^{\perp} = \{A \mid \mathrm{Hom}(X, A) = 0,\ \forall X \in \mathfrak{X}\}$$

is called the *torsion-free class generated by* $\mathfrak{X}$, and

$$^{\perp}\mathfrak{X} = \{A \mid \mathrm{Hom}(A, X) = 0,\ \forall X \in \mathfrak{X}\}$$

the *torsion class cogenerated by* $\mathfrak{X}$. The pair $(^{\perp}(\mathfrak{X}^{\perp}), \mathfrak{X}^{\perp})$ is called the *torsion theory generated by the class* $\mathfrak{X}$, and the pair $(^{\perp}\mathfrak{X}, (^{\perp}\mathfrak{X})^{\perp})$ the *torsion theory cogenerated by the class* $\mathfrak{X}$. On the classes $\mathfrak{X}^{\perp}$ and $^{\perp}\mathfrak{X}$ see [30, Sections 8, 11], [31, Section 10]. P. Schultz, C. Vinsonhaler and W.J. Wickless (1992) studied the equivalence relation: $A \approx B \Leftrightarrow {}^{\perp}A = {}^{\perp}B$ and $A^{\perp} = B^{\perp}$. W.J. Wickless (1992) considered the relation: $G \approx H$ if and only if $(^{\perp}G)^{\perp} = (^{\perp}H)^{\perp}$.

II. For two given groups A and B there exists a group G such that $G \supseteq A$ and $G/A \cong B$ (for example $G = A \oplus B$). We say that such a group G is an *extension* of the group A by the group B. Two extensions G_1 and G_2 of A by B are called *equivalent* if there is an isomorphism φ between the groups G_1 and G_2 such that its restriction on A is equal to the identity mapping of A and the induced mapping $B \cong G_1/A \xrightarrow{\bar{\varphi}} G_2/A \cong B$, where $\bar{\varphi}(g_1 + A) = \varphi(g_1) + A$ for $g_1 \in G_1$, is equal to the identity mapping on B.

The set of all classes of equivalent extensions of A by B with the natural addition operation is an Abelian group (see [7, Sections 49, 50], [6, Section 61]). This group $\mathrm{Ext}(B, A)$ is called the *group of extensions* of A by B. The class of extensions equivalent to the extension $A \oplus B$ is the zero element in $\mathrm{Ext}(B, A)$.

If two extensions of a torsion group T by a torsion-free group K are isomorphic (but not necessarily equivalent), then they determine elements of the group $\mathrm{Ext}(K, T)$ of the same order (see A. Mader, 1968). For the same groups T and K an extension G of T by K determines an element of $\mathrm{Ext}(K,T)$ of finite order if and only if the group G is quasi-isomorphic (Section 6) to the group $T \oplus K$ (C.P. Walker, 1964).

1. Some elementary properties of the group $\mathrm{Ext}(B, A)$ (depending on the properties of the groups A and B) are listed in [7, Section 52], in particular

$$\mathrm{Ext}\left(\bigoplus_{i \in I} A_i, G\right) \cong \prod_{i \in I} \mathrm{Ext}(A_i, G),$$
$$\mathrm{Ext}\left(A, \prod_{i \in I} G_i\right) \cong \prod_{i \in I} \mathrm{Ext}(A, G_i)$$

for all A, A_i, G and G_i [7, Theorem 52.2].

For any group G and exact sequence

$$E\colon\ 0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$$

the following sequences are exact:

$$\begin{aligned} &0 \to \mathrm{Hom}(G, A) \xrightarrow{\alpha_*} \mathrm{Hom}(G, B) \xrightarrow{\beta_*} \mathrm{Hom}(G, C) \\ &\xrightarrow{E_*} \mathrm{Ext}(G, A) \xrightarrow{\bar{\alpha}_*} \mathrm{Ext}(G, B) \xrightarrow{\bar{\beta}_*} \mathrm{Ext}(G, C) \to 0 \end{aligned}$$

and

$$\begin{aligned} &0 \to \mathrm{Hom}(C, G) \xrightarrow{\beta_*} \mathrm{Hom}(B, G) \xrightarrow{\alpha_*} \mathrm{Hom}(A, G) \\ &\xrightarrow{E_*} \mathrm{Ext}(C, G) \xrightarrow{\bar{\beta}^*} \mathrm{Ext}(A, G) \xrightarrow{\bar{\alpha}^*} \mathrm{Ext}(A, G) \to 0 \end{aligned}$$

([7, Theorem 51.3]), where α_*, β_*, $\bar{\alpha}_*$, $\bar{\beta}_*$ and β^*, α^*, $\bar{\beta}^*$, $\bar{\alpha}^*$ are homomorphisms induced by homomorphisms α and β respectively and E_* and E^* are the so called *connecting homomorphisms* (see [7, Section 51]). These exact sequences play an important role in the proofs of many statements in Abelian group theory.

2. For two arbitrary groups A and B the group $\mathrm{Ext}(B, A)$ is cotorsion (Section 1, Point 5; see [7, Theorem 54.6]). Moreover, for an arbitrary reduced group H the group $\mathrm{Ext}(Q/Z, H)$ is a *cotorsion hull*, i.e. a minimal cotorsion group which contains H ([7, Proof of the Theorem 58.1], K.M. Rangaswamy, 1964).

If A is a torsion-free group or any algebraically compact group, then $\mathrm{Ext}(B, A)$ is an algebraically compact group for every group B ([7, Section 52, N, O]). The group $\mathrm{Ext}(B, A)$ is algebraically compact for every group A if and only if the torsion part of the group B is a direct sum of cyclic groups (M.J. Schoeman, 1973).

Let us note that for a reduced torsion group G with $G^1 \neq 0$ the group $\mathrm{Ext}(Q/Z, G)$ is cotorsion but not algebraically compact, see [7, Lemma 55.3, Propositions 54.2 and 56.4]. On necessary and sufficient conditions for $\mathrm{Ext}(B, A)$ to be a divisible group or to be a reduced group for fixed A and any B or for fixed B and any A see K.M. Rangaswamy (1964).

3. A complete description of the structure of the group $\mathrm{Ext}(B, A)$ for torsion-free groups A and B of finite rank was obtained by R.B. Warfield, Jr. (1972). L.Ya. Kulikov (1976) has found a complete description of the groups $\mathrm{Ext}(T, A)$ when T is a torsion group, A a torsion-free group, and of the groups $\mathrm{Ext}(T, G)$ (1964), where T is a countable reduced primary group, G an arbitrary group.

K.M. Rangaswamy (1964) described the structure of $\mathrm{Ext}(B, Z)$ for arbitrary B (this gives a solution to the first part of Problem 54 from [6, Ch. X]).

Other results on the structure of concrete groups of extensions or for a given group to be isomorphic to a group of extensions can be found in [27, Section 3], Sections 8 of [28], [29], [30], [31, Section 10], [9, Section 52], [11, Section D.33, Points 1, 2].

4. P.J. Gräbe and G. Viljoen (two papers in 1970) investigated classes $\mathfrak{M}$ of groups such that there is a group H with $G \cong \mathrm{Ext}(H, G)$ (respectively, $G \cong \mathrm{Ext}(G, H)$) for every $G \in \mathfrak{M}$.

$A \cong \mathrm{Ext}(A, Z)$ if and only if A is a finite group (F.R. Beyl and A. Hanna, 1976).

5. If L_1 and L_2 are quasi-isomorphic (Section 6) torsion-free groups, then $\mathrm{Ext}(L_1, G) \cong \mathrm{Ext}(L_2, G)$ for any group G (L. Procházka, 1966). If L_1, L_2 are torsion-free groups of finite rank without infinite cyclic direct summands, then the converse statement also holds ([52, Theorem 1.7]). If A and B are reduced torsion-free groups of finite rank, then $\mathrm{Ext}(C, A) \cong \mathrm{Ext}(C, B)$ for any group C (for any torsion-free group C of finite rank) if and only if the groups A and B are quasi-isomorphic ([53]). This gives a solution to Problem 43 [7, Ch. IX] in the case of finite rank groups A and B.

If $\mathrm{Ext}(X, C) \cong \mathrm{Ext}(Y, C)$ for fixed groups X and Y and an arbitrary group C, then $\mathrm{Ext}(tX, C) \cong \mathrm{Ext}(tY, C)$ and $\mathrm{Ext}(X_p, C) \cong \mathrm{Ext}(Y_p, C)$ for every prime p (for the notations see Section 0). If X and Y are p-groups and X is a countable group of finite Ulm type (Section 1, Point 4), then assuming GCH we see that $\mathrm{Ext}(X, C) \cong \mathrm{Ext}(Y, C)$ $(\forall C)$ implies $X \cong Y$ (A.I. Moskalenko, 1991).

6. We say that the group A is *L-orthogonal* to the group B and the group B is *R-orthogonal* to the group A if $\mathrm{Ext}(B, A) = 0$ (K.M. Rangaswamy, 1964). For the cases when $\mathrm{Ext}(B, A) = 0$ see also [7, Section 52] and [57].

One says that for the classes $\mathfrak{X}$, $\mathfrak{Y}$ of groups the condition $\mathfrak{X} \perp \mathfrak{Y}$ is satisfied if $\mathrm{Ext}(X, Y) = 0$ for any $X \in \mathfrak{X}$, $Y \in \mathfrak{Y}$.

Let $\mathfrak{X}^{\perp} = \{Y \mid \mathfrak{X} \perp Y\}$, ${}^{\perp}\mathfrak{X} = \{Y \mid Y \perp \mathfrak{X}\}$. The couple $(\mathfrak{X}, \mathfrak{Y})$ is said to be a *cotorsion theory* if $\mathfrak{Y} = \mathfrak{X}^{\perp}$, $\mathfrak{X} = {}^{\perp}\mathfrak{Y}$. If $(\mathfrak{F}, \mathfrak{G})$ is a cotorsion theory and $\mathfrak{G} = \mathfrak{X}^{\perp}$, $\mathfrak{F} = {}^{\perp}(\mathfrak{X}^{\perp})$ for a class of groups $\mathfrak{X}$, then $(\mathfrak{F}, \mathfrak{G})$ is said to be *cogenerated* by $\mathfrak{X}$. For cotorsion theories see R. Göbel and R. Prelle (1979), L. Salce (1979).

A group G is called a *Baer group* if $\mathrm{Ext}(G, T) = 0$ for all torsion groups T. Griffith has proved that every Baer group is free, see [7, Section 101].

A group W such that $\mathrm{Ext}(W, Z) = 0$ is called a *Whitehead group*. Assuming ZFC + V = L, we see that every Whitehead group is free (this gives a positive solution in this case for Problem 79 [7, Ch. XIII]). But under ZFC + MA + $\neg$CH we obtain that for all cardinals $\varkappa \geqslant \aleph_1$ there exists a non-free Whitehead group of cardinality $\varkappa$, see P.C. Eklof, 1977 (a result of S. Shelah), and A.H. Mekler, 1983.

7. Some special subgroups of the group $\mathrm{Ext}(B, A)$ play important role in the theory of Abelian groups (in particular, $\mathrm{Pext}(B, A)$ and $\mathrm{Bext}(B, A)$).

The subgroup $\mathrm{Pext}(B, A)$ consists of all the elements of the group $\mathrm{Ext}(B, A)$ determined by those extensions G of A by B in which A is a pure subgroup. Note that

$$\mathrm{Pext}(B, A) = \bigcap_{n=1}^{\infty} n\mathrm{Ext}(B, A)$$

([7, Theorem 53.3]). $\mathrm{Pext}(B, A) = 0$ for any group B (for any group A) if and only if A is algebraically compact (B is the direct sum of cyclic groups, respectively), see [7, Proposition 53.4].

On the groups $\operatorname{Pext}(B, A)$ see also [7, Section 53], M.J. Schoeman (1973).

$\operatorname{Bext}(B, A)$ is the subgroup of $\operatorname{Ext}(B, A)$ consisting of all its elements determined by extensions G in which A is a so called *regular subgroup* (for the definition of regular subgroups see C. Walker (1972) or Ch. Megibben (1980)). $\operatorname{Bext}(B, A) = 0$ for all A if and only if B is completely decomposable (in the sense of Section 4, Point 5). In case G/H is torsion-free, H is a regular subgroup of G if and only if each coset $g + H$ contains an element a for which $\chi_G(a) = \chi_{G/H}(g + H)$ where $\chi_G(a)$ is the *characteristic* of a in G, i.e. the sequence of generalized p-heights of the element a for all prime p (compare Section 3, Point 1). If this condition holds for $H = tG$ then tG is called *balanced* in G (compare [7, Section 86]).

If T is torsion and F torsion-free then $\operatorname{Bext}(F, T) \supseteq t\operatorname{Ext}(F, T)$. $\operatorname{Bext}(F, T) = \operatorname{Ext}(F, T)$ for all torsion groups T if and only if F is homogeneous of type $\tau \ni (0, 0, \dots)$ (see Section 3, Point 3). Torsion-free groups F with the property $\operatorname{Bext}(F, T) = 0$ for any torsion group T are called *B_1-groups* (Section 3, Point 3).

8. Some other questions concerning the group $\operatorname{Ext}(B, A)$ and its subgroups were touched on in the surveys [27–31] and in [7, Ch. IX].

9. Tensor product. Torsion product

I. The *tensor product* $A \otimes B$ of two Abelian groups A and B is the Abelian group determined by taking as generators all couples $a \otimes b$ ($a \in A, b \in B$), and as defining relations

$$(a_1 + a_2) \otimes b = a_1 \otimes b + a_2 \otimes b, \qquad a \otimes (b_1 + b_2) = a \otimes b_1 + a \otimes b_2,$$

where $a, a_1, a_2 \in A, b, b_1, b_2 \in B$ [7, Section 59], [11, Section D.33, Point 6].

The definition of tensor products was introduced by H. Whitney ([11, Section D.33, Point 6]).

1. Note that always $A \otimes B \cong B \otimes A$; $(A \otimes B) \otimes C \cong A \otimes (B \otimes C)$; $(\bigoplus_{i\in I} A_i) \otimes (\bigoplus_{j\in J} B_j) \cong \bigoplus_{i,j}(A_i \otimes B_j)$ and $Z \otimes B \cong B$, $Z(m) \otimes B \cong B/mB$ for all groups B ([7, Section 59]).

2. If A or B is a divisible group, then $A \otimes B$ is also divisible. The tensor product of two algebraically compact torsion-free groups is an algebraically compact group (A.A. Fomin, 1981).

For any group H the group $G = Q_p \otimes H$ is p-local (Section 4, Point 2) and it is called the *localisation* of H at the prime p.

If A or B is a p-group then $A \otimes B$ is a p-group. If A is a p-group and $pB = B$ then $A \otimes B = 0$ (therefore $A \otimes B = 0$ if A is a p-group, B is a q-group for $p \neq q$).

The tensor product of nonzero torsion-free groups A, B is a nonzero torsion-free group ([11, Section D.33, Point 6]). In this case for $a \in A$, $b \in B$ and their types (Section 3, Point 1) we have $\tau(a \otimes b) = \tau(a) + \tau(b)$ (see [7, Section 60, Ex. 9]). But the tensor product of two *p-reduced* (i.e. such that they do not contain nonzero subgroups C with $C = pC$) torsion-free groups may be not a p-reduced group (A.A. Fomin, 1975).

If U and V are pure subgroups of the groups A and B respectively, then the elements $u \otimes v$ ($u \in U$, $v \in V$) generate a subgroup of $A \otimes B$ isomorphic to the group $U \otimes V$. The same holds for any subgroups of A and B if A and B are torsion-free groups ([11, Section D.33, Point 6]), but in general it does not hold. For example, if $Z(p) \cong U \cong A$, $Z(p) \cong V \subset B \cong Z(p^\infty)$, then $U \otimes V \cong Z(p)$ and $A \otimes B = 0$.

If A, B are p-groups and A', B' are their basic subgroups (Section 2, Point 2) then $A \otimes B \cong A' \otimes B'$ (see [7, Theorem 61.1]). Therefore, the tensor product of torsion groups is always the direct sum of finite cyclic groups ([7, Theorem 61.3]). The tensor product of a p-group A and a torsion-free group B is isomorphic to the direct sum of as many copies of the group A as is the p-rank (Section 0) of the group B (L. Fuchs, 1957).

Concerning the torsion part of the tensor product we note that $t(A \otimes B) \cong (tA \otimes tB) \oplus (tA \otimes B/tB) \oplus (A/tA \otimes tB)$ ([7, Theorem 61.5]).

3. A.A. Fomin (1976) has obtained a description of the tensor product of two countable torsion-free groups based on the Kurosh–Derry–Maltsev description (Section 5, Point 2) of the factors (see [6, Ch. XI, Problem 57] and [7, Ch. X, Problem 47]), and also a description (by means of invariants) of the tensor product of algebraically compact torsion-free groups (see A.A. Fomin, 1981).

Let G and K be countable groups of torsion-free rank one (see Section 4, Point 2) with reduced torsion parts. Then one can construct invariants of the group $G \otimes K$ using the invariants [7, Theorem 104.3] of the groups G and K, see Ch. Megibben and E. Toubassi (1976).

In the class of torsion-free groups A with $r_p(A) \leqslant 1$ for all p the tensor multiplication coincides up to the divisible part with multiplication in the ring of universal numbers (Section 5, Point 3; A.A. Fomin, 1981).

4. A necessary and sufficient condition for a torsion-free group of finite rank mn to be isomorphic to the tensor product of two torsion-free groups of ranks m and n respectively was obtained by H. Lausch (1983). H. Lausch (1985) also gave an example of a group A of rank $r(A) = 8$ such that $A \cong B_1 \otimes B_2 \cong C_1 \otimes C_2 \otimes C_3$, where $r(B_1) = 4$, $r(C_i) = 2$ ($i = 1, 2, 3$), but B_1 is not a tensor product of two rank 2 groups.

If M is a torsion-free group and A is a torsion-free group of rank 1 then $M \cong A \otimes G$ for some group G if and only if for any $x \in M$ there exists a homomorphism $f \in \mathrm{Hom}(A, M)$ such that $x \in f(A)$, and, moreover, in this case $M \cong A \otimes \mathrm{Hom}(A, M)$ (see [63]).

Each homogeneous torsion-free group of type τ is the tensor product of a rank 1 torsion-free group of type τ and a homogeneous torsion-free group of zero type (i.e. type $\tau(Z)$) (see L.G. Nongxa, 1984).

5. If $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ is a short exact sequence such that $\alpha(A)$ is a pure subgroup of B (such an exact sequence is called *pure-exact*), then for every group G and induced homomorphisms α_*, β_* the sequence

$$0 \to A \otimes G \xrightarrow{\alpha_*} B \otimes G \xrightarrow{\beta_*} C \otimes G \to 0$$

is pure-exact; the converse statement holds also ([7, Theorem 60.4] and T.J. Head, 1967).

Some other results on the properties of tensor product can be found in [7, Ch. X] and in the surveys [27–31], where, in particular, some works concerning the question, when does the tensor product of the given groups split (Section 4, Point 1), are cited.

II. The *torsion product* of Abelian groups A, B is the Abelian group $\mathrm{Tor}(A, B)$ given by taking as generators all triples (m, a, b), where m is an integer, $a \in A$, $b \in B$ and $ma = 0 = mb$, and as defining relations all possible equalities of the form

$$\begin{aligned}
&(m, a_1, b) + (m, a_2, b) = (m, a_1 + a_2, b),\\
&(m, a, b_1) + (m, a, b_2) = (m, a, b_1 + b_2),\\
&(m, na, b) = (m, a, nb) = (mn, a, b)
\end{aligned}$$

([7, Section 62], [11, Section D.33, Point 8]).

1. In the group $\mathrm{Tor}(A, B)$ we have $(m, 0, b) = 0 = (m, a, 0)$. If A or B is torsion-free, then $\mathrm{Tor}(A, B) = 0$. For all groups A and B, A_i and B_j we have

$$\begin{aligned}
&\mathrm{Tor}(A, B) \cong \mathrm{Tor}(tA, tB) \cong \mathrm{Tor}(tB, tA) \cong \mathrm{Tor}(B, A),\\
&\mathrm{Tor}\Big(\bigoplus_{i\in I} A_i, \bigoplus_{j\in J} B_j\Big) \cong \bigoplus_{i,j} \mathrm{Tor}(A_i, B_j).
\end{aligned}$$

2. The group $\mathrm{Tor}(A, B)$ is torsion for any groups A and B. If A or B is a p-group then $\mathrm{Tor}(A, B)$ is a p-group. If A is a p-group and B a q-group for different primes p and q, then $\mathrm{Tor}(A, B) = 0$.

We have: $\mathrm{Tor}(Z(m), G) \cong G[m]$ for all groups G and all integers $m > 0$; $\mathrm{Tor}(Q/Z, G) \cong tG$; $\mathrm{Tor}(Z(p^\infty), G) \cong G_p$ [7, Section 62, H, I, J].

If A and B are reduced p-groups, then $\mathrm{Tor}(A, B)$ is a reduced p-group of length $\min\{\lambda(A), \lambda(B)\}$ (R.J. Nunke, 1964).

For p-groups A, B, P. Hill (1983) found conditions on $\mathrm{Tor}(A, B)$ to be a direct sum of cyclic groups. The structure of the group $\mathrm{Tor}(A, B)$ is considered in R.J. Nunke (1967), [30, Section 8], [31, Section 10].

3. P.W. Keef (1990) found necessary and sufficient conditions for a p-group G to be isomorphic to the group $\mathrm{Tor}(A, B)$ for some totally projective groups A and B (Section 2, Point 3). For a given group B the class $\mathcal{T}_B$ of all p-groups A with the property $A \cong \mathrm{Tor}(A, B)$ is considered in another paper by P.W. Keef, 1990. A p-group A is called Tor-*idempotent* if $\mathrm{Tor}(A, A) \cong A$. There exist reduced Tor-idempotent groups of arbitrary length.

4. Problem 50 in [7, Ch. X] asks: how are two groups A and B related if $\mathrm{Tor}(A, C) \cong \mathrm{Tor}(B, C)$ for all reduced groups C?

D.O. Cutler (1982) constructed an example of non-isomorphic p-groups A and B with $\mathrm{Tor}(A, G) \cong \mathrm{Tor}(B, G)$ for all reduced p-groups G.

But if $\mathrm{Tor}(A, X) \cong \mathrm{Tor}(B, X)$ for given torsion groups A, B and an arbitrary reduced torsion group X, then A and B have the same cardinality $\mathfrak{m}$, the same Ulm invariants (Section 2, Point 3), their maximal divisible subgroups are isomorphic, and $\bigoplus_{\mathfrak{m}} A \cong \bigoplus_{\mathfrak{m}} B$ (P. Hill, 1971).

If A, B, R, S are totally projective groups, then $\mathrm{Tor}(A, B) \cong \mathrm{Tor}(R, S)$ if and only if the Ulm invariants of $\mathrm{Tor}(A, B)$ and $\mathrm{Tor}(R, S)$ coincide (see P.W. Keef, 1990). The Ulm invariants for $\mathrm{Tor}(A, B)$ were described by R.J. Nunke (1964).

5. If E: $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} C \to 0$ is a short exact sequence then for any group G the induced sequence

$$0 \to \mathrm{Tor}(A, G) \xrightarrow{\alpha'} \mathrm{Tor}(B, G) \xrightarrow{\beta'} \mathrm{Tor}(C, G) \xrightarrow{E_*} A \otimes G$$
$$\xrightarrow{\alpha\otimes 1} B \otimes G \xrightarrow{\beta\otimes 1} C \otimes G \to 0$$

is exact ([7, Theorem 63.1]). If the subgroup $\alpha(A)$ is pure in B then $\alpha'(\mathrm{Tor}(A, G))$ is pure in $\mathrm{Tor}(B, G)$ and $\mathrm{Im}\, E_* = 0$ (see [7, Theorem 63.2]).

Some additional information on the groups $\mathrm{Tor}(A, B)$ can be found in the surveys [27–31] and also in [7, Ch. X].

10. Valuated groups

A function v on a p-group G is called a *p-valuation* if it satisfies the following conditions:

(1) vx is an ordinal or the symbol ∞ ($x \in G$);

(2) $vpx > vx$ (we assume that $\infty < \infty$ and $\alpha < \infty$ for all ordinals α);

(3) $v(x + y) \geqslant \min(vx, vy)$;

(4) $vnx = vx$ if $(n, p) = 1$

(see [32,47]).

A p-group G together with a p-valuation on it is called a *valuated p-group.*

As an example of a p-valuation on a p-group G one can take the height function h^* for which to every element $x \in G$ we let correspond its generalized p-height $h^*(x)$.

By a *p-valuation* v_p on an arbitrary Abelian group G we mean a function with the properties (1)–(4) of a p-valuation on a p-group. The group G is called a *valuated group* if a p valuation v_p is defined on G for every prime p (see [34]).

Any valuated group A can be embedded into a group B in such a way that the restriction to A of the p-height function of B for every p is the p-valuation on A (see [32] and [34, Theorem 23]).

By a *homomorphism of valuated groups* $f: A \to A'$ we mean a homomorphism f such that $v'_p f(x) \geqslant v_p x$ for all $x \in A$ and p. If f is a monomorphism and $v'_p f(x) = v_p x$ for all $x \in A$ and p then f is said to be an *embedding* of A into A'. We say that the valuated group A is a (valuated) *subgroup* of the valuated group B if A is a subgroup of B and the natural embedding of A into B is an embedding of valuated groups. If A is a p-local (Section 4, Point 2) valuated group then by a subgroup of A we mean a $\mathbb{Q}_p$-submodule.

If A is a p-local valuated group and α is an ordinal or ∞, then let $A(\alpha) = \{a \in A \mid va \geqslant \alpha\}$. If $\alpha \neq \infty$ then the *α-th Ulm invariant* $f_A(\alpha)$ of A is defined to be the *rank* (i.e. the dimension of the corresponding vector space over $\mathbb{F}_p$) of the p-bounded group $F_A(\alpha) = \{a \in A(\alpha) \mid pa \in A(\alpha+2)\}/A(\alpha+1)$. $f_A(\infty)$ is the rank of the p-bounded group $F_A(\infty) = \{a \in A(\infty) \mid pa = 0\}$. If G is a valuated group (not necessary p-local) then the Ulm invariants $f_G(p, \alpha)$ of G are the Ulm invariants of the various localizations (Section 9, Part I, Point 2) of G (see [33]).

The *direct sum* of a family of valuated groups is their group direct sum, the value of an element being the minimum of the values of its components. A *free valuated group* F is a direct sum of infinite cyclic valuated groups with the property $v_p px = v_p x + 1$ for all $x \in F$ and p. A complete system of invariants with respect to isomorphism for direct sums of valuated infinite cyclic groups was obtained by D.M. Arnold, R. Hunter and E. Walker (1979). A direct summand of such a group is a direct sum of valuated infinite cyclic groups (R.O. Stanton, 1979). But there exists subgroups of direct sums of valuated prime order p cyclic groups which are not direct sums of valuated cyclic groups (see [34]).

The validity for valuated groups of some results form Abelian group theory is discussed by R. Hunter, F. Richman and E. Walker (1984 and 1987). The Ulm theorem concerning simply presented p-groups ([7, Theorem 83.3]) holds for valuated simply presented p-groups.

On valuated Warfield groups (Section 4, Point 3) and other questions connected with valuated groups see [29, Section 10], [30, Section 10], [31, Section 11], [20, p. 7], [36, p. 6], [23,33].

11. Other questions

1. *Purity.* For pure subgroups (Section 1, Point 5) of Abelian groups the following statements hold:

(0) every direct summand of a group is a pure subgroup,

(1) if $A \subseteq B \subseteq C$ and A is pure in B, B is pure in C then A is pure in C;

(2) if $A \subseteq B \subseteq C$ and A is pure in C then A is pure in B;

(3) if A is pure in B then A/K is pure in B/K for any subgroup $K \subseteq A$;

(4) if $K \subseteq A \subseteq B$ with K pure in B and A/K pure in B/K then A is pure in B.

One says that in the class of all Abelian groups there is given an *ω-purity* (a purity ω) if in each group some subgroups (called ω-pure) are fixed (notation: $A \subseteq_\omega B$) so that for these subgroups the conditions (0)–(4) above are satisfied [12, Section 1], [26, Section 11, Part I].

An example of an ω-purity (nearest to purity and called ε-purity) is the following. For each prime p let us fix a set, possibly empty, of natural numbers $M_p = \{k_{1p}, k_{2p}, \ldots\}$; by definition, $A \subseteq_\varepsilon B$ if and only if $a = p^{k_{ip}} b$ with $a \in A$, $b \in B$ implies $a = p^{k_{ip}} a_1$ for some $a_1 \in A$ (for every prime p and $k_{ip} \in M_p$). The concept of ε-purity was studied by V.S. Rokhlina, see [28, Section 5]. If each M_p is the set of all natural numbers we have purity, if $M_p = \{1\}$ for each p we have *neatness* ([6, Section 28]).

Many other examples of ω-purity can be found in [12, Section 1], Section 5 of [28], [29,30] and [31, Section 7]; in particular, see T. Kepka (1973), S.I. Komarov (1982), C.P. Walker (1966).

An ω-purity is caled *inductive* if the union of an ascending chain of ω-pure subgroups in any group is an ω-pure subgroup. A.A. Manovcev (1975) gave a complete description of all inductive ω-purities in the class of all Abelian groups: they are intersections of ε-purities and Head's η-purities (T.J. Head, 1967). The intersection of two purities ω_1 and ω_2 is the purity $\omega_1 \cap \omega_2$ such that $A \subseteq_{\omega_1 \cap \omega_2} B$ whenever $A \subseteq_{\omega_1} B$ and $A \subseteq_{\omega_2} B$. *η-purity* (originally called subpurity) is defined as follows: $A \subseteq_\eta B$ (or $A \subseteq_{\eta P} B$) if, for a fixed set P of primes, $a = p^k b$ ($a \in A, b \in B$, $p \in P$, k natural) implies $p^l a = p^{k+l} a_1$ for some $a_1 \in A$ and some integer $l \geqslant 0$. In another way this problem has been solved by V.I. Kuz'minov (1976), see also E.G. Sklyarenko (1981).

The extensions G of the group A by the group B (Section 8, Part II) in which A is an ω-pure subgroup of G determine elements of the group $\mathrm{Ext}(B, A)$ forming a subgroup $\omega\mathrm{Ext}(B, A)$ ([12, Point (1.2)]). Concerning $\omega\mathrm{Ext}(B, A)$ see [12, Points (1.3), (1.4)], [26, Section 11, Part I, Point 2].

If $\omega\mathrm{Ext}(B, A) = \mathrm{Ext}(B, A)$ for fixed A and any B (respectively, for fixed B and any A) then A (respectively B) is called *ω-divisible* (*ω-flat*) ([12, p. 26]). A.V. Ivanov (1978) found necessary and sufficient conditions on two classes of groups $\mathfrak{A}$ and $\mathfrak{B}$ to be the class of all ω-divisible and the class of all ω-flat groups for an ω-purity. The result is valid also for modules over arbitrary left hereditary rings.

If $\omega\mathrm{Ext}(B, A) = 0$ for a fixed group A and any group B (respectively, for fixed B and any A) then A is called *ω-injective* (respectively, B is called *ω-projective*), see [12, Point (1.6)]. This is equivalent to the group A (the group B) to be injective (projective) relative to all exact sequences $0 \to K \xrightarrow{\alpha} G \to L \to 0$ where $\alpha(K) \subseteq_\omega G$, see [12, p. 24 and Point (1.6)].

We say that a group G is *quasi-injective* (*quasi-projective*) if it is injective (projective) with respect to any exact sequence $0 \to H \xrightarrow{\alpha} G \to G/H \to 0$ with $\alpha(H) \subseteq_\omega G$, see [29, Section 5], [30, Section 5], [31, Section 7], [16] and Problem 17 from [7, Ch. V].

Many other matters concerning ω-purity are discussed in [12, Section 1], [26, Section 11, Part I], [27, Section 4], Section 5 of [28], [29,30], [31, Section 7] and [35].

2. *Radicals of Abelian groups.* A function R assigning to each group A a subgroup $R(A)$ so that if $f: A \to B$ is a homomorphism then $f(R(A)) \subseteq_b R(B)$ is called a *subfunctor of the identity* or *subfunctor of the identity functor* (R.J. Nunke, 1964). A subfunctor of the identity is also called a *preradical* (P. Hill and Ch. Megibben, 1968). If R is a preradical then the subgroup $R(A)$ is called a preradical or a *functorial subgroup* of the group A (B. Charles, 1968).

R. Marty (1968) calls a subgroup $R(G)$ (a subgroup $S(G)$) of the group G a *radical* (*socle*) if R (respectively S) is a preradical and, moreover, $R(G/R(G)) = 0$ (respectively, $S(S(G)) = S(G)$). A radical which is at the same time a socle is called *idempotent radical* by B.J. Gardner (1972). Let us note that in [12, Section 2] the word "radical" means idempotent radical.

If $R(A) = A$ or $R(A) = 0$ then the group A is called *R-radical* or *R-semisimple*, respectively. The idempotent radical $R(G)$ of a group G is the greatest R-radical subgroup of the group G and at the same time is the intersection of all its subgroups H such that G/H is R-semisimple ([12, Point (2.8)]). It is always the case that $R(A)$ is a pure subgroup in A (B.J. Gardner, 1972).

A radical R such that $A = R(A) \oplus A'$ for some subgroup A' in any group A is called *splitting*. A *non-trivial* radical R (i.e. not such that $R(A) = A$, $\forall A$ or $R(A) = 0$, $\forall A$) is splitting if and only if the class of all R-radical groups coincides with the class of all divisible groups or with one of the classes of all groups of the form $\bigoplus_{p\in\pi} G_p$ where G_p is a divisible p-group, π is a fixed set of primes (B.J. Gardner, ibid).

Let $\mathcal{A}$ be an arbitrary class of Abelian groups and

$$R_{\mathcal{A}}(G) = \bigcap \{\ker f \mid f \in \operatorname{Hom}(G, X),\ X \in \mathcal{A}\}$$

for any group G. Then $R_{\mathcal{A}}$ is a radical (in the sense of Marty, see above) and $\{G \mid R_{\mathcal{A}}(G) = 0\}$ is the least class of groups containing $\mathcal{A}$ and closed under formation of subgroups and direct products (a class of groups with the last two properties is called an annihilator class). If R is an arbitrary radical in the class of all Abelian groups then $R = R_{\mathcal{A}}$ with $\mathcal{A} = \{G \mid R(G) = 0\}$. This gives a bijective correspondence between annihilator classes and radicals (M. Dugas, T.H. Fay, S. Shelah, 1987).

If $\mathcal{A}$ is the class of all $\aleph_1$-free groups (Section 1, Point 3) then the radical $R_{\mathcal{A}}$ is called the *Chase radical*. The Chase radical of a group G is the sum of its countable subgroups that have no free direct summands (see K. Eda, 1987).

Additional information on radicals in Abelian groups can be found in [12, Section 2] and [28, Section 9], [29, Section 10], [30, Section 11], [31, Section 11]. On the relations between radicals and purities see [12, Point (2.47)] and A.V. Ivanov, 1978.

3. *N-high subgroups.* If N is a subgroup of the group A then the subgroup B of A is called *N-high* if it is maximal with respect to the property $N \cap B = 0$. Concerning various properties of such subgroups see [27, Section 5], [28, Section 9], [29, Section 6], [30, Section 6], [31, Introduction]. Subgroups N with all N-high subgroups being ε-pure (Point 1 above) for a given ε-purity were described by V.S. Rokhlina (1971).

4. *Transitive and fully transitive groups.* A torsion-free group G is called *transitive* (respectively *fully transitive* or *fully invariant*) if for any $a, b \in G$ with characteristics (Section 3, Point 1) $\chi(a) = \chi(b)$ ($\chi(a) \leqslant \chi(b)$ respectively) there exists an automorphism (endomorphism) η of the group G such that $\eta(a) = b$. Transitive and fully transitive groups can be defined also in the case of arbitrary groups if one uses the notions of indicator and height matrix ([7, Section 65, 103]), see [31, Section 8]. Concerning properties of transitive and fully transitive groups see [28, Section 1], [29, Section 1], [30, Section 7], [31, Section 8], [16], P.A. Krylov (1984), V.M. Misyakov (1994). A.L. Corner (1976) has constructed an example of a transitive (fully transitive) but not fully transitive (not transitive, respectively) reduced p-group, see also Ch. Megibben, 1969.

5. Interesting results have been obtained for groups regarded as modules over their endomorphism rings, see [28, Section 9], [29, Section 10], [30, Section 11], [31, Section 11] and also [54] where some unsolved problems are formulated.

6. On the additive group of a ring see [5] (correction in Y. Hirano and I. Mogami, 1986), [6, Ch. XII], [7, Ch. XVII], [26, Section 11, Part III], [27, Section 10], [29, Section 9]; [30, Section 9], [31, Section 11].

For topological Abelian groups (in particular, for the duality theory for locally compact Abelian groups) see L.S. Pontryagin [13].

The book by K.R. Goodearl [8] is devoted to partially ordered Abelian groups with applications to algebraic K-theory.

7. Many other questions concerning Abelian groups are discussed in the reviews [27–31] where all Abelian group papers refereed in "Referativnyj Zhurnal Mat." in 1962–1992 are surveyed.

References

Books

[1] D. Arnold, *Finite Rank Torsion Free Abelian Groups and Rings VIII*, Lecture Notes in Math., Vol. 931, Springer, Berlin (1982).
[2] R.A. Beaumont and R.S. Pierce, *Torsion-Free Groups of Rank Two*, Amer. Math. Soc., Providence, RI (1961).
[3] I.Kh. Bekker and S.F. Kozhukhov, *Automorphisms of Torsion-Free Abelian Groups* (in Russian), Tomsk Univ., Tomsk (1988).
[4] P. Eklof, *Set Theoretic Methods in Homological Algebra and Abelian Groups*, Les Presses de l'Université, Montreal (1980).
[5] S. Feigelstock, *Additive Groups of Rings*, Pitman Adv. Publ. Program (1983).
[6] L. Fuchs, *Abelian Groups*, Publ. House Hungar. Acad. Sci., Budapest (1958).
[7] L. Fuchs, *Infinite Abelian Groups*, Vol. I (1970); Vol. II (1973), Academic Press, New York.
[8] K.R. Goodearl, *Partially Ordered Abelian Groups with Interpolation*, Amer. Math. Soc., Providence, RI (1986).
[9] P.A. Griffith, *Infinite Abelian Group Theory*, Chicago Lectures in Math., The Univ. of Chicago, Chicago, London (1970).
[10] I. Kaplansky, *Infinite Abelian Groups*, Univ. of Michigan Press, Ann Arbor, MI (1954 and 1969).
[11] A.G. Kurosh, *Theory of Groups*, 3rd ed. (in Russian), Nauka, Moscow (1967).
[12] A.P. Mishina and L.A. Skornyakov, *Abelian Groups and Modules*, Amer. Math. Soc. Translations, Ser. 2, Vol. 107, Amer. Math. Soc., Providence, RI (1976).
[13] L.S. Pontryagin, *Topogical Groups*, 2nd ed., Gordon and Breach, New York (1966).

Surveys

[14] D.M. Arnold, *Notes on Butler groups and balanced extensions*, Boll. Un. Mat. Ital. A5 **2** (1986), 175–184.
[15] D.M. Arnold and C. Vinsonhaler, *Finite rank Butler groups: A survey of recent results*, Abelian Groups (Curaçao, 1991), Lecture Notes in Pure and Appl. Math., Vol. 146, Dekker, New York (1993), 17–41.
[16] I.Kh. Bekker, P.A. Krylov and A.R. Chekhlov, *Torsion-free Abelian groups close to algebraically compact*, Abelian Groups and Modules, No. 11-12, Tomsk Univ., Tomsk (1994), 3–52 (in Russian).
[17] L. Bican, *Subgroups of Butler groups*, Abelian Groups and Modules. Proc. Padova Conf., 23 June–1 July 1994, Kluwer Acad. Publ. (1995), 49–57.
[18] L. Fuchs, *A survey on Butler groups of infinite rank*, Abelian Group Theory and Relat. Top.: Conf. Abelian Groups, Oberwolfach, 1–7 Aug. 1993, Providence, RI (1994), 121–139.
[19] R. Göbel, *Wie weit sind Moduln vom Satz von Krull–Remak–Schmidt entfernt?*, Jaresber. Deutsch. Math. Verein. **88** (1) (1986), 11–49.
[20] R. Göbel, *Helmut Ulm: His work and its impact on recent mathematics*, Contemp. Math. **87** (1989), 1–10.

[21] P. Hill, *The classification problem*, Abelian Groups and Modules. Proc. Conf., Udine, 9–14 Apr. 1984, Wien, New York (1984), 1–16.
[22] P. Hill and Ch. Megibben, *Mixed groups*, Trans. Amer. Math. Soc. **334** (1) (1992), 121–142.
[23] R.H. Hunter, F. Richman and E. Walker, *Warfield modules*, Lecture Notes in Math., Vol. 616 (1977), 87–123.
[24] A. Mader, *Almost completely decomposable torsion-free Abelian groups*, Abelian Groups and Modules. Proc. Padova Conf., 23 June–1 July 1994, Kluwer Acad. Publ. (1995), 343–366.
[25] W. May, *Endomorphism rings of mixed Abelian groups*, Contemp. Math. **87** (1989), 61–74.
[26] A.V. Mikhalev and A.P. Mishina, *Infinite Abelian groups: Methods and results*, Fund. i Prikl. Mat. **1** (2) (1995), 319–375 (in Russian).
[27] A.P. Mishina, *Abelian groups*, Algebra. Topology. Geometry, 1965, Itogi Nauki i Tekh. VINITI Akad. Nauk SSSR, Moscow (1967), 9–44 (in Russian).
[28] A.P. Mishina, *Abelian groups*, Algebra. Topology. Geometry, Vol. 10, Itogi Nauki i Tekh. VINITI Akad. Nauk SSSR, Moscow (1972), 5–45 (in Russian); J. Soviet Math. **2** (3) (1974), 239–263.
[29] A.P. Mishina, *Abelian groups*, Algebra. Topology. Geometry, Vol. 17, Itogi Nauki i Tekh. VINITI Akad. Nauk SSSR, Moscow (1979), 3–63.
[30] A.P. Mishina, *Abelian groups*, Algebra. Topology. Geometry, Vol. 23, Itogi Nauki i Tekh. VINITI Akad. Nauk SSSR, Moscow (1985), 51–118 (in Russian); J. Soviet Math. **40** (1988), 288–330.
[31] A.P. Mishina, *Abelian groups*, Contemporary Math. and Its Applications, Thematic Survey, Vol. 10, Algebra-2, Moscow (1994), 2721–2792.
[32] F. Richman, *A guide to valuated groups*, Lecture Notes in Math., Vol. 616 (1977), 73–86.
[33] F. Richman, *Mixed groups*, Lecture Notes in Math., Vol. 1006, (1983), 445–470.
[34] F. Richman and E. Walker, *Valuated groups*, J. Algebra **56** (1) (1979), 145–167.
[35] E.G. Sklyarenko, *Relative homological algebra in the category of modules*, Uspekhi Mat. Nauk **33** (3) (1978), 85–120 (in Russian).
[36] C. Vinsonhaler, *The Abelian groups papers of Robert B. Warfield, Jr.*, Abelian Groups and Noncommutat. Rings: Collec. Pap. Mem. Robert B. Warfield (Jr.), Providence, RI (1992), 5–9.
[37] R.B. Warfield, Jr., *The structure of mixed Abelian groups*, Lecture Notes in Math., Vol. 616 (1977), 1–38.
[38] R.B. Warfield, Jr., *Classification theory of Abelian groups. II: Local theory*, Lecture Notes in Math., Vol. 874 (1981), 322–349.

Selected papers

[39] D.M. Arnold, *Pure subgroups of finite rank completely decomposable groups*, Lecture Notes in Math., Vol. 874 (1981), 1–30.
[40] D.M. Arnold and C. Vinsonhaler, *Pure subgroups of finite rank completely decomposable grooups. II*, Lecture Notes in Math., Vol. 1006 (1983), 97–143.
[41] A.L.S. Corner, *Every countable reduced torsion-free ring is an endomorphism ring*, Proc. London Math. Soc. **13** (52) (1963), 687–710.
[42] Yu.L. Ershov, *On algebraically compact groups*, Algebra i Logika (Novosibirsk) **17** (6) (1978), 684–692 (in Russian).
[43] A.A. Fomin, *Invariants and duality in some classes of finite rank torsion-free Abelian groups*, Algebra i Logika (Novosibirsk) **26** (1) (1987), 63–83 (in Russian).
[44] A.A. Fomin, *Abelian groups with one τ-adic relation*, Algebra i Logika (Novosibirsk) **28** (1) (1989), 83–104 (in Russian).
[45] A.A. Fomin, *The category of quasi-homomorphisms of Abelian torsion free groups of finite rank*, Proc. Int. Conf. Algebra Dedicat. Mem. A.I. Mal'cev, Novosibirsk, 21–26 Aug. 1989, Providence, RI (1992), 91–111.
[46] C.F. Gauss, *Werke*, 12 Bände, Göttingen (1870–1927), s. B. 4 (Disquisitiones mathematicae).
[47] R. Hunter and E. Walker, *Valuated p-groups*, Lecture Notes in Math., Vol. 874 (1981), 350–373.
[48] A.V. Ivanov, *Countable direct sums of complete and torsion-complete groups*, Trudy Moskov. Mat. Obshch. **40** (1979), 121–170 (in Russian).
[49] A.V. Ivanov, *On the test problem of Kaplansky*, Trudy Moskov. Mat. Obshch. **42** (1981), 200–220 (in Russian).

[50] A.V. Ivanov, *Abelian groups with self-injective endomorphism rings and endomorphism rings with the annihilator condition*, Abelian Groups and Modules (Tomsk) (1982), 93–109 (in Russian).

[51] B. Jónsson, *On direct decomposition of torsion-free Abelian groups*, Math. Scand. **7** (2) (1959), 361–371.

[52] P.A. Krylov, *On two problems concerning groups of extensions of Abelian groups*, Mat. Sb. **185** (1) (1994), 73–94 (in Russian).

[53] P.A. Krylov, *On isomorphism of groups of extensions of Abelian groups*, Abelian Groups and Modules, No. 13–14 (1996), 105–131 (in Russian).

[54] P.A. Krylov, *Torsion-free Abelian groups as modules over their endomorphism rings*, Abelian Groups and Modules, No. 13–14, Tomsk Univ., Tomsk (1996), 77–104 (in Russian).

[55] L.Ya. Kulikov, *To the theory of Abelian groups of arbitrary power*, Mat. Sb. **16** (1945), 129–162 (in Russian).

[56] L.Ya. Kulikov, *Generalized primary groups*, Trudy Moskov. Mat. Obshch. **1** (1952), 247–326; **2** (1953), 85–167 (in Russian).

[57] L.Ya. Kulikov, *Conditions by which a group of Abelian extensions is zero*, Uchen. Zap. Moskov. Gos. Ped. Inst. **375** (1971), 41–55.

[58] H. Leptin, *Abelschen p-Gruppen und ihre Automorphismengruppen*, Math. Z. **73** (3) (1960), 235–253.

[59] W. Liebert, *Isomorphic automorphism groups of primary Abelian groups*, Abelian Group Theory: Proc. 3rd Conf., Oberwolfach, 11–17 Aug. 1985, New York (1987), 9–31.

[60] A.A. Manovtsev, *Inductive purities in Abelian groups*, Mat. Sb. **96** (3) (1975), 414–446 (in Russian).

[61] A.M. Sebel'din, *Determining of Abelian groups by their endomorphism rings*, Abelian Groups and Modules, No. 8, Tomsk (1989), 113–123 (in Russian).

[62] C. Vinsonhaler and W. Wickless, *Dualities for torsion-free Abelian groups of finite rank*, J. Algebra **128** (2) (1990), 474–487.

[63] R.B. Warfield, Jr., *Homomorphisms and duality for torsion-free groups*, Math. Z. **107** (3) (1968), 189–200.

Section 6C
Representation Theory of 'Continuous Groups' (Linear Algebraic Groups, Lie Groups, Loop Groups, ...) and the Corresponding Algebras

This Page Intentionally Left Blank

Infinite-Dimensional Representations of Quantum Algebras

A.U. Klimyk

Institute for Theoretical Physics, Kiev-143, 252143 Ukraine

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

1. Introduction

Quantum groups appeared first (in a somewhat different form) in the 1960th in papers by G.I. Kac (from Kiev) on group algebras (see, for example, [9]). They did not attract interest of mathematicians at that time. In other words, the discovery of quantum groups by G.I. Kac was ahead of its time. In the 1980th certain algebraic structures appeared in papers (related the to quantization) by L.D. Faddeev and his collaborators, which are now identified with the quantum group $\mathrm{SL}_q(2, \mathbf{C})$ and with the q-deformation of the universal enveloping algebra $U(\mathrm{sl}_2)$. Attempts to grasp the meaning of these examples led to formulation of the contemporary notion of quantum groups and quantum algebras [8,13,27]. Quantum groups and algebras turned out to be of a great importance for contemporary mathematics and for theoretical and mathematical physics. By means of the universal R-matrix and the Yang–Baxter equation they are applied in quantum field theory and in statistical physics. On the way, several meaningful exactly solvable integrable systems were solved.

Quantum groups appear as mathematical objects describing symmetries (in a somewhat different meaning of this word) in non-commutative geometry. They are "motion groups" of such objects of non-commutative geometry as the quantum plane, quantum spheres and quantum hyperboloids. On the basis of quantum groups a non-commutative differential geometry and calculus is developing which is different in spirit from that of A. Connes. By making use of this non-commutative differential geometry physicists try to construct q-quantum mechanics. The theory of the q-analogue of the quantum harmonic oscillator is developed. The quantum Heisenberg group has been constructed.

Quantum groups and algebras can be applied to the investigation of basic hypergeometric functions and q-orthogonal polynomials in the same way as Lie groups and discrete groups are used for investigation of the usual special functions and orthogonal polynomials. In this way, some new addition theorems and product formulas for q-orthogonal polynomials were proved.

In a series of papers quantum algebras were applied to describe some physical results. For example, it has been shown that representations of the quantum algebra $U_q(\mathrm{su}_2)$ describe rotational spectra of nuclei better than representations of the Lie group SU(2) do. However, one has to choose an appropriate value of q and the physical meaning of this choice is not yet understood. Representations of the quantum algebras $U_q(\mathrm{u}_n)$ can be applied to the description of relations between masses of hadrons (strongly interacting elementary particles). This leads to q-analogues of the Gell–Mann–Okubo mass formulas. It is well known that the Gell–Mann–Okubo formulas do not give exact relations between masses. Mass formulas, obtained by means of the quantum algebras $U_q(\mathrm{u}_n)$, give exact mass relations (up to the exactness of experimental data) if q is related to the Alexander polynomials from algebraic geometry describing the topology of surfaces in $\mathbf{R}^3$.

Most applications use representations of quantum groups and quantum algebras rather than these groups and algebras themselves. Finite dimensional irreducible representations of quantum algebras, corresponding to simple complex Lie algebras, for q not coinciding with a root of unity were classified by Lusztig [20] and Rosso [28]. It was shown that every irreducible finite dimensional representation of a simple Lie algebra g can be deformed in an irreducible representation of the corresponding quantum algebra (q-deformed universal enveloping algebra) $U_q(g)$. Moreover, every finite dimensional irreducible representation of $U_q(g)$ is essentially obtained in this way (after possibly tensoring by a one-dimensional representation). Jimbo [14] gave the action formulas for the operators of finite dimensional irreducible representations of the quantum algebra $U_q(\mathrm{u}_n)$ with respect to bases analogous to the Gel'fand–Tsetlin bases for irreducible representations of the classical group $U(n)$. These formulas differ from the corresponding Gel'fand–Tsetlin formulas by replacement of all expressions of the type $(m_{ik}-m_{js}\pm a)$ by $[m_{ik}-m_{js}\pm a]$, where $[b]$ is the q-number defined as

$$[b]=\left(q^{b/2}-q^{-b/2}\right)\left(q^{1/2}-q^{-1/2}\right)^{-1}.$$

Representations of quantum algebras with respect to Gel'fand–Tsetlin bases when q is a root of unity are described in [2,4].

It is a more complicated matter to deal with the irreducible finite dimensional representations of the q-analogue of the orthogonal group. The Gel'fand–Tsetlin formulas for the action of the representation infinitesimal operators for these groups use a basis constructed according to restriction of representations of the orthogonal group $\mathrm{SO}(n)$ successively onto the subgroups

$$\mathrm{SO}(n-1)\supset\mathrm{SO}(n-2)\supset\cdots\supset\mathrm{SO}(2).$$

The quantum orthogonal group $\mathrm{SO}_q(n)$, constructed according to the Drinfeld–Jimbo method, contains the quantum subgroup $\mathrm{SO}_q(n-2)$ and does not contains $\mathrm{SO}_q(n-1)$. Therefore, the Gel'fand–Tsetlin procedure of construction of irreducible finite-dimensional representations cannot be applied to $\mathrm{SO}_q(n)$. To be successful with this procedure, we proposed in [10] another definition of the q-deformed quantum algebras $U_q(\mathrm{so}_n)$ which permits restriction from $U_q(\mathrm{so}_n)$ onto $U_q(\mathrm{so}_{n-1})$. This allows us to construct a q-analogue of the Gel'fand–Tsetlin bases in the carrier spaces of representations of $U_q(\mathrm{so}_n)$. However, there are difficulties with introducing the structure of a Hopf algebra on these algebras.

In this review we mainly concentrate our attention on infinite dimensional representations of quantum algebras. For this case we have not such a close analogy with the theory of infinite-dimensional representations of semisimple Lie algebras as in the case of finite-dimensional representations. The main aim of this chapter is to illustrate the characteristic properties and special features of infinite-dimensional representations of quantum algebras. Since it is not possible to describe all achievements of the theory of infinite-dimensional representations of quantum algebras in this short review, after a general description of the status of this theory we give a more detailed description of the infinite-dimensional representations of the quantum algebra $U_q(\mathrm{su}_{1,1})$, which is one of simplest noncompact quantum algebras. The representations of $U_q(\mathrm{su}_{1,1})$ show the main features of infinite dimensional irreducible representations of quantum algebras with respect to irreducible representations of simple Lie algebras. Then we shortly discuss infinite dimensional representations of some other quantum algebras.

Quantum algebras and quantum groups are equipped with the structure of a Hopf algebras. We refer the reader for the corresponding definitions and statements to the papers [7,12] or to the book [34], Chapter 14.

2. Status of the theory of infinite-dimensional representations of quantum algebras

In the case of infinite-dimensional representations of real semisimple Lie algebras, there is a one-to-one correspondence between irreducible infinite-dimensional representations of a connected simply connected real semisimple Lie group G and those of its Lie algebra g (we consider only those representations of G and g which correspond to Harish-Chandra g-modules). Let $P = MAN$ be a minimal parabolic subgroup in G [19], and let ω be an irreducible finite dimensional representation of P. We induce by ω a representation T_ω of G. The representations T_ω constitute the principal nonunitary series of G. Using the method described in [5] we can construct the corresponding representations T_ω of g induced by representations ω of the subalgebra $p = m + a + n$ (the Lie algebra of P). There is Harish-Chandra's theorem which states that every irreducible representation of g (or of G) is equivalent to some irreducible representation T_ω or to a subrepresentation of a quotient representation of some reducible representation T_ω. By means of this theorem a classification of all irreducible representations of g has been obtained.

An essential detail in the construction of these representations is the Iwasawa decomposition of G (and of g) and parabolic subgroups. In the Iwasawa decomposition $g = k + a + n$ of g the subalgebra k corresponds to a maximal compact subgroup K of G.

In order to extend this method of construction of infinite dimensional representations to quantum algebras $U_q(g)$ we must have the corresponding subalgebras in $U_q(g)$. However, to some subalgebras of g (to some subgroups of G) there correspond no quantum subalgebras in $U_q(g)$ (no quantum subgroups in the quantum group G_q). In particular, some quantum algebras $U_q(g)$ have no quantum subalgebras corresponding to the subalgebra k of g. For example, the quantum group $\mathrm{SL}_q(n, \mathbf{R})$ has no quantum subgroups corresponding to the subgroup $\mathrm{SO}(n)$ of $\mathrm{SL}(n, \mathbf{R})$. In this cases we cannot extend to $U_q(g)$ the methods of construction of induced infinite dimensional representations which are so useful in the classical case. As is shown by the example of the quantum algebra $U_q(\mathrm{sl}(n, \mathbf{R}))$ [32], for these quantum algebras the theory of infinite dimensional representations is absolutely different from that for the corresponding Lie algebras. Almost no results for representations of these quantum algebras are obtained. Only finite dimensional irreducible representations and irreducible representations with highest weights are classified for them.

The quantum algebras $U_q(\mathrm{u}_{m,n})$, $U_q(\mathrm{sp}_{m,n})$ and the quantum algebras $U_q(\mathrm{so}_{r,s})$ with even r and s have quantum subalgebras corresponding to the subalgebra k of g. For these quantum algebras we can construct representations analogous to representations of g. However, up to now we do not yet have a complete theory of infinite dimensional representations for them. But we can state that many elements of the theory of representations of the corresponding Lie algebras $\mathrm{u}_{m,n}$, $\mathrm{sp}_{m,n}$, $\mathrm{so}_{r,s}$ are valid for $U_q(\mathrm{u}_{m,n})$, $U_q(\mathrm{sp}_{m,n})$ and $U_q(\mathrm{so}_{r,s})$.

Along with the method of construction of induced representations of Lie groups and Lie algebras by means of Iwasawa and Gauss decompositions, there are other methods of obtaining infinite dimensional representations of these groups and algebras. Let us mention two of them:

(1) construction of Verma modules;

(2) derivation of explicit formulas of action of representation operators upon basis vectors (formulas of the Gel'fand–Tsetlin type).

Verma modules can be also constructed for quantum algebras. For details of this construction see [1,20,23]. If q is not a root of unity, then the theory of Verma modules is similar to that for semisimple Lie algebras [6,20]. In particular, by means of Verma modules we can construct irreducible representations of a quantum algebra $U_q(g)$ (q not a root of unity) with highest weights. As in the classical case, for every weight Λ there exists a unique irreducible representation of $U_q(g)$ with highest weight Λ. If Λ is a dominant and integral weight, then this irreducible representation is finite-dimensional. Otherwise it is infinite-dimensional. The situation is completely different when q is a root of unity. In these cases all Verma modules are reducible. Their irreducible components are finite-dimensional. Some aspects of the theory of Verma modules, when q is a root of unity, can be found in [1,6,21,22].

Irreducible representations of quantum algebras can be given by explicit formulas for the action of representation operators upon basis elements. The action formulas for operators of representations of Lie algebras can be derived from formulas for induced representations of the corresponding Lie groups [16]. We cannot do this in the quantum case since no induced representations of quantum groups exist. But for quantum case one can use the action formulas for representations of Lie algebras in order to try to deform them for obtaining the formulas for operators of representations of $U_q(g)$. To prove that the formulas thus obtained really give representations of a quantum algebra we must verify that the corresponding defining relations are satisfied.

In this chapter we consider infinite-dimensional representations of some specific quantum algebras. These algebras exhibit the main properties of infinite-dimensional representations of quantum algebras in the case when q is not a root of unity. The main peculiarities of these representations (with respect to the classical case) are

(1) appearance of the so called strange series of representations;

(2) appearance of additional relations of equivalence.

By using the example of the quantum algebra $U_q(\mathrm{su}_{1,1})$ we show that these peculiarities have influence upon harmonic analysis on quantum groups and on quantum homogeneous spaces. Namely, representations from the strange series appear in the decomposition of the regular and quasi-regular representations of $U_q(\mathrm{su}_{1,1})$. Additional equivalences between representations (they are a consequence of the periodicity of q-numbers) lead to the appearance of theta functions in the corresponding Plancherel measure [30,31]. This fact is of great importance for integral transforms with kernels expressed in terms of basic hypergeometric functions.

3. Representations of the quantum algebra $U_q(\mathrm{su}_{1,1})$

3.1. *The quantum algebra $U_q(\mathrm{sl}_2)$ and its real forms*

Let $q = \exp h$ be a fixed complex number. The elements E_+, E_-, H satisfying the relations

$$[H, E_+] = 2E_+, \qquad [H, E_-] = -2E_-, \qquad [E_+, E_-] = \frac{q^{H/2} - q^{-H/2}}{q^{1/2} - q^{-1/2}} \tag{1}$$

generate an associative algebra which is denoted by $U_q(\mathrm{sl}_2)$. Sometimes it is useful to take the elements $k = q^{H/4} = \exp(hH/4)$, k^{-1}, E_+, E_- instead of E_+, E_-, H. For k, k^{-1}, E_+, E_- we have the relations

$$kE_+k^{-1} = q^{1/2}E_+, \qquad kE_-k^{-1} = q^{-1/2}E_-, \qquad kk^{-1} = k^{-1}k = 1, \tag{2}$$

$$[E_+, E_-] = \frac{k^2 - k^{-2}}{q^{1/2} - q^{-1/2}}. \tag{3}$$

It is clear that the monomials $E_+^l k^m E_-^n$, $m \in \mathbf{Z}$, $l, n \in \mathbf{Z}_+$, as well as the monomials $E_-^l k^m E_+^n$, $m \in \mathbf{Z}$, $l, n \in \mathbf{Z}_+$, form a basis of $U_q(\mathrm{sl}_2)$. Here $\mathbf{Z}$ is the set of all integers and $\mathbf{Z}_+$ is the set of nonnegative integers.

The algebra $U_q \equiv U_q(\mathrm{sl}_2)$ comes equipped with the structure of a Hopf algebra. To give this structure we linearly extend the multiplication $(a, b) \to ab$ to a homomorphism $m : U_q \otimes U_q \to U_q$ and introduce a comultiplication $\Delta : U_q \to U_q \otimes U_q$, an antipode $S : U_q \to U_q$ and a counit $\varepsilon : U_q \to \mathbf{C}$ in $U_q \equiv U_q(\mathrm{sl}_2)$ by the formulas

$$\Delta(E_\pm) = E_\pm \otimes q^{H/4} + q^{-H/4} \otimes E_\pm, \qquad \Delta(H) = H \otimes 1 + 1 \otimes H, \tag{4}$$

$$S(E_\pm) = -q^{\pm 1/2}E_\pm, \qquad S(H) = -H, \tag{5}$$

$$\varepsilon(E_\pm) = \varepsilon(H) = 0. \tag{6}$$

Since the operators Δ, S and ε are homomorphisms for the corresponding algebras, they are defined on all elements from $U_q(\mathrm{sl}_2)$. The associative algebra $U_q(\mathrm{sl}_2)$ equipped with this structure of a Hopf algebra is called the quantum algebra $U_q(\mathrm{sl}_2)$.

If $q = \exp \pi i m$, $m \in \mathbf{Z}$, then in $U_q(\mathrm{sl}_2)$ there are non-trivial ideals. If J is a maximal ideal in $U_q(\mathrm{sl}_2)$, then the quotient algebra $U_q(\mathrm{sl}_2)/J$ is finite-dimensional. Below we assume that q is not a root of unity.

By introducing a *-operation into $U_q(\mathrm{sl}_2)$ we can distinguish real forms of this algebra. The formulas $H^* = H$, $E_+^* = E_-$, $E_-^* = E_+$ define the *-structure which distinguishes in $U_q(\mathrm{sl}_2)$ its compact real form, denoted by $U_q(\mathrm{su}_2)$. The formulas

$$H^* = H, \qquad E_+^* = -E_-, \qquad E_-^* = -E_+ \tag{7}$$

define a noncompact real form, denoted by $U_q(\mathrm{su}_{1,1})$. The *-structure $H^* = H$, $E_+^* = E_+$, $E_-^* = E_-$ defines the noncompact real form $U_q(\mathrm{sl}_{2,\mathbf{R}})$. The mapping $^* : U_q(\mathrm{sl}_2) \to U_q(\mathrm{sl}_2)$ must be an anti-automorphism of the associative algebra $U_q(\mathrm{sl}_2)$ and an automorphism of the coalgebra $U_q(\mathrm{sl}_2)$ satisfying the conditions $(a^*)^* = a$, $S(S(a^*)^*) = a$ for all $a \in U_q(\mathrm{sl}_2)$. For this reason we must take $q \in \mathbf{R}$ for $U_q(\mathrm{su}_2)$ and for $U_q(\mathrm{su}_{1,1})$ and $|q| = 1$ for $U_q(\mathrm{sl}_{2,\mathbf{R}})$.

3.2. *Representations of the principal series*

In order to determine a representation of the quantum algebra $U_q(\mathrm{su}_{1,1})$ we give operators $T(H)$, $T(E_+)$, $T(E_-)$ satisfying the relations

$$T(H)^* = T(H), \qquad T(E_+)^* = -T(E_-), \qquad T(E_-)^* = -T(E_+). \tag{8}$$

These operators define also a representation of the quantum algebra $U_q(\mathrm{sl}_2)$.

These representations of the algebra $U_q(\mathrm{su}_{1,1})$ are infinite-dimensional. There are diferent definitions of representations of $U_q(\mathrm{su}_{1,1})$. We use the following one. A linear representation T of $U_q(\mathrm{su}_{1,1})$ is a homomorphism of the associative algebra $U_q(\mathrm{sl}_2)$ into the algebra of linear operators on a Hilbert space V which are defined on an everywhere dense subspace $D \subset V$ such that relations (1) and (8) are fulfilled on D and $T(\exp(2\mathrm{i}\theta H))$, $0 \leqslant \theta < 2\pi$, is a direct sum of one-dimensional representations of the group SO(2). The last condition means that the operator $T(H)$ is diagonal in an appropriate basis and has a discrete spectrum consisting of integral and half-integral eigenvalues. If we remove the condition of fulfilment of relations (8), then we obtain representations of the quantum algebra $U_q(\mathrm{sl}_2)$. Sometimes, as in the case of the Lie algebra $\mathrm{su}_{1,1}$, these representations are called representations of $U_q(\mathrm{su}_{1,1})$. Then representations of $U_q(\mathrm{su}_{1,1})$ defined above are called infinitesimal unitary representations of $U_q(\mathrm{su}_{1,1})$. Below we shall omit the word "infinitesimal".

Let ε be equal to 0 or $1/2$ and let V_ε be the complex Hilbert space with orthonormal basis $\{\mathbf{e}_m \mid m = n + \varepsilon, n = 0, \pm 1, \pm 2, \ldots\}$. We fix a complex number σ and denote by $T_{\sigma\varepsilon}$ the representation of $U_q(\mathrm{sl}_2)$ on V_ε given by the formulas

$$T_{\sigma\varepsilon}(H/2)\mathbf{e}_m = m\mathbf{e}_m, \qquad T_{\sigma\varepsilon}(E_\pm)\mathbf{e}_m = [-\sigma \pm m]\mathbf{e}_{m\pm 1}, \tag{9}$$

where $[a]$ is a q-number. These representations $T_{\sigma\varepsilon}$ constitute the principal series of representations of $U_q(\mathrm{sl}_2)$. When $q \to 1$ then $[-\sigma \pm m] \to -\sigma \pm m$. Therefore, at $q = 1$ the formulas (9) give the representations $T_{\sigma\varepsilon}$ of the principal nonunitary series of the Lie algebra $\mathrm{su}_{1,1}$ [33]. As in the classical case, the following theorem is valid:

THEOREM [17]. *The representation $T_{\sigma\varepsilon}$ is irreducible if and only if 2σ is not an integer of the same evenness as 2ε.*

We have to analyse reducible representations $T_{\sigma\varepsilon}$ and to find equivalence relations. As in the classical case [33], for the irreducible representations $T_{\sigma\varepsilon}$ we have the equivalence

$$T_{\sigma\varepsilon} \sim T_{-\sigma-1,\varepsilon}, \tag{10}$$

which are realized by means of intertwining operators. These relations exhaust all equivalences for irreducible representations of the Lie algebra $\mathrm{su}_{1,1}$. For the quantum algebra we have additional equivalence relations. They are consequences of the periodicities of the function

$$w(z) = [z] \equiv \left(q^{z/2} - q^{-z/2}\right)\left(q^{1/2} - q^{-1/2}\right)^{-1}.$$

We set $q = \exp h$, $h \in \mathbf{R}$. Then $w(z)$ is a periodic function with period $4\pi\mathrm{i}/h$. Moreover, we have $w(z) = -w(z + 2\pi\mathrm{i}/h)$. This leads to the equivalences

$$T_{\sigma\varepsilon} = T_{\sigma+4\pi\mathrm{i}k/h,\varepsilon}, \qquad T_{\sigma\varepsilon} \sim T_{\sigma+2\pi\mathrm{i}k/h,\varepsilon}, \quad k \in \mathbf{Z}, \tag{11}$$

for real q. If $q = \exp \mathrm{i}h,\ h \in \mathbf{R}$, then

$$T_{\sigma\varepsilon} = T_{\sigma+4\pi k/h,\varepsilon}, \qquad T_{\sigma\varepsilon} \sim T_{\sigma+2\pi k/h,\varepsilon}, \quad k \in \mathbf{Z}. \tag{12}$$

3.3. *Irreducible representations of* $U_q(\mathrm{sl}_2)$

From the reducible representations $T_{\sigma\varepsilon}$ we extract irreducible components. As a result, we obtain the irreducible representations $T_l^{\pm}$, $l = \frac{1}{2}, 1, \frac{3}{2}, \ldots$, which are q-analogues of the corresponding irreducible representations of the Lie algebra $\mathrm{su}_{1,1}$ (see, for example, [33]). Taking into account the equivalence relations (10)–(12), we obtain the following classes of nonequivalent infinite-dimensional irreducible representations of $U_q(\mathrm{sl}_2)$, $q = \exp h$, $h \in \mathbf{R}$:

(a) the representations $T_{\sigma\varepsilon}$, where $\varepsilon \in \{0, 1/2\}$, $\mathrm{Re}\,\sigma \geqslant -1/2$, $0 \leqslant \mathrm{Im}\,\sigma < 2\pi/h$, $2\sigma \not\equiv 2\varepsilon \pmod 2$;

(b) the representations T_l^{+}, T_l^{-}, $l = \frac{1}{2}, 1, \frac{3}{2}, \ldots$.

In the case of $U_q(\mathrm{sl}_2)$, $q = \exp \mathrm{i}h$, $h \in \mathbf{R}$ (q is not a root of unity) we must replace class (a) by the representations $T_{\sigma\varepsilon}$, where $\varepsilon \in \{0, 1/2\}$, $-1/2 \leqslant \mathrm{Re}\,\sigma < -1/2+2\pi/h$, $2\sigma \not\equiv 2\varepsilon \pmod 2$. For the algebra $U_q(\mathrm{sl}_2)$, $q = \exp\,(\mathrm{i}h_1 + h_2)$, $h_1 \neq 0$, $h_2 \neq 0$, the class (a) consists of representations $T_{\sigma\varepsilon}$, where $\varepsilon \in \{0, 1/2\}$, $\mathrm{Re}\,\sigma \geqslant -1/2$, $2\sigma \not\equiv 2\varepsilon \pmod 2$.

REMARK. We can remove the condition that the operator $T(H)$ contains only integral or half-integral eigenvalues and just demand that $T(H)$ has a discrete spectrum. In this case we can consider that ε is a complex number and V_ε has the orthonormal basis $\{\mathbf{e}_m \mid m = n+\varepsilon, n = 0, \pm 1, \pm 2, \ldots\}$. Then formulas (9) with complex ε give the representations $T_{\sigma\varepsilon}$ determined by two complex numbers. Detailed description of these representations are given in [3].

3.4. *Irreducible representations of* $U_q(\mathrm{su}_{1,1})$

Now we can verify which of the representations of Section 3.3 satisfy the conditions (8). As a result we obtain the following classes of representations of the quantum algebra $U_q(\mathrm{su}_{1,1})$ (they are also called unitary representations of $U_q(\mathrm{su}_{1,1})$) when q is real (see [17,25,30]):

(a) the representations $T_{\sigma\varepsilon}$, $\sigma = \mathrm{i}\rho - 1/2$, $0 \leqslant \rho < 2\pi/h$ (the principal unitary series);

(b) the representations $T_{\sigma 0}$, $-1/2 \leqslant \sigma < 0$ (the supplementary series);

(c) the representations $T_{\sigma\varepsilon}$, $\mathrm{Im}\,\sigma = \pi/h$, $\mathrm{Re}\,\sigma > -1/2$ (the strange series);

(d) the representations T_l^{+}, T_l^{-}, $l = \frac{1}{2}, 1, \frac{3}{2}, \ldots$ (the discrete series);

(e) the zero representation.

The strange series of unitary representations has no analogue in the classical case. When $q \to 1$ then $\pi/h \to \infty$ and this series disappears. Note that contrary to the classical case, representations of the principal unitary series are parametrized by the points of a compact set.

The operators of the representations of the principal unitary series satisfy relations (8) with respect to the scalar product of the Hilbert space V_ε. In order to satisfy these relations for other classes of representations we have to introduce other scalar products in V_ε. To define these scalar products we note that V_ε can be realized as a space of functions [30]. To obtain this space we define the correspondence $\mathbf{e}_m \to z^{-m}$ and obtain the space $\mathbf{C}[z, z^{-1}]$ of polynomials in z and z^{-1}. Now equip $\mathbf{C}[z, z^{-1}]$ with an appropriate scalar product and obtain the Hilbert space $L^2(\mathbf{C}, d\mu)$. The representation operators satisfy conditions (8) on these Hilbert spaces. To give explicitly these Hilbert spaces we must define the measure $d\mu$. For the principal unitary series

$$d\mu(z) = \delta\big(|z|^2 - 1\big)\, dz\, d\bar{z},$$

where δ is the delta function. For the discrete series representations $T_l^{\pm}$ we have

$$\begin{aligned} d\mu_\pm(z) &= (1-q)^{2l-1} \sum_{k=0}^{\infty} q^{k(-l\mp\varepsilon+1)} \begin{bmatrix} 2l+k-2 \\ k \end{bmatrix}_q \\ &\quad \times \delta\big(|z|^2 - q^{\pm(k+l-1/2)}\big)\, dz\, d\bar{z}, \end{aligned} \tag{13}$$

where $\varepsilon \in \{0, 1/2\}$, $2l \equiv 2\varepsilon \ (\mathrm{mod}\, 2)$ and

$$\begin{bmatrix} m \\ n \end{bmatrix}_q = \frac{[m]!}{[n]![m-n]!}, \qquad [n]! = [1][2]\cdots[n].$$

For the strange series representations $T_{\sigma\varepsilon}$ for which $2\,\mathrm{Re}\,\sigma \in \mathbf{Z}$ we have

$$d\mu(z) = \sum_{k=0}^{2m+1} q^{k(k-2m+2\varepsilon-1)} \begin{bmatrix} 2m+1 \\ k \end{bmatrix}_q \delta\big(|z|^2 - q^{m-k+1/2}\big)\, dz\, d\bar{z}, \tag{14}$$

where $m = \mathrm{Re}\,\sigma$. In these formulas we suppose that $q > 0$.

Considering $U_q(\mathrm{su}_{1,1})$ as an associative algebra with involution one can consider representations of $U_q(\mathrm{su}_{1,1})$ for complex values of q. Operators of these representations must satisfy relations (8). If $q = \exp ih$, $h \in \mathbf{R}$, then $U_q(\mathrm{su}_{1,1})$ has the following classes of irreducible representations [17]:

(a) the representations $T_{\sigma\varepsilon}$, $\sigma = \mathrm{i}\rho - 1/2$, $\rho \in \mathbf{R}$ (the principal unitary series);
(b) the representations $T_{\sigma\varepsilon}$, $\mathrm{Re}\,\sigma = \pi/h$ (the strange series);
(c) the zero representation.

If $q = \exp(h_1 + \mathrm{i}h_2)$, $h_1, h_2 \in \mathbf{R}, h_1 \neq 0$, $h_2 \neq 0$, then the algebra $U_q(\mathrm{su}_{1,1})$ has no irreducible unitary representations except for the zero one.

3.5. *Properties of representation operators*

The operators of the irreducible representations of the Lie algebra $\mathrm{su}_{1,1}$ corresponding to the representations of $U_q(\mathrm{su}_{1,1})$ from Section 3.3 are unbounded. Since for real q we have $[m] \to \pm\infty$ when $m \to \infty$, and for $q = \exp ih$, $h \in \mathbf{R}$, we obtain

$$\big|[m]\big| \leqslant 2\big|q^{1/2} - q^{-1/2}\big|^{-1} \quad \text{when } m \to \infty,$$

then the operators $T_{\sigma\varepsilon}(E_\pm)$ are unbounded for real q and bounded for imaginary q.

We now discuss the problem of self-adjointness for the operators $T(E_+ - E_-)$ and $T(\mathrm{i}E_+ + \mathrm{i}E_-)$ when $q > 0$. Recall that $E_+ - E_-$ and $\mathrm{i}(E_+ + E_-)$ are symmetric elements in the quantum algebra $U_q(\mathrm{su}_{1,1})$.

There is connection between representations of a semisimple Lie group and representations of its Lie algebra. To noncompact generators I there correspond unbounded operators in the infinite-dimensional irreducible representations T of a semisimple Lie algebra g. To every such representation T of g there corresponds an irreducible representation T of the Lie group G with Lie algebra g. The operators of this representation T of G are bounded. If the representation T of G is unitary, then the corresponding representation of g is given by skew-symmetric operators. Unitarity of a representation T of G means that the closures of these skew-symmetric operators multiplied by $\sqrt{-1}$ are selfadjoint operators. Properties of self-adjointness for operators corresponding to symmetric elements of the universal enveloping algebra $U(g)$ of g are also well known.

There is no corresponding theory for infinite-dimensional representations of quantum algebras. Moreover, as we shall see below, the situation for quantum algebras is unlike that what we have in the classical case.

Let $L \equiv T_l^+(E_+ - E_-)$. We have

$$L|m\rangle = a_m|m+1\rangle + a_{m-1}|m-1\rangle, \quad a_m = \big([-l+m][l+m+1]\big)^{1/2}.$$

Let $\overline{L}$ be the closure of the operator L. It is well known from the theory of Jacobi matrices that operators of the type $\overline{L}$ are selfadjoint or have deficiency indices (1,1) (and therefore are not selfadjoint). There is the statement which asserts that if beginning with some positive j we have $a_{n-1}a_{n+1} \leqslant a_n^2$, $n \geqslant j$, and, besides, $\sum_{n=0}^{\infty} a_n^{-1} < \infty$, then $\overline{L}$ is not a selfadjoint operator. These conditions are fulfilled for the operator $\overline{L}$ and hence $\overline{L}$ has deficiency indices (1,1). This means that deficiency subspaces are one-dimensional. These deficiency subspaces can be found explicitly. Coordinates of basis vectors of these subspaces are expressed in terms of q-orthogonal polynomials [15]. The operator $\overline{L}$ has selfadjoint extensions which are constructed by means of deficiency subspaces. The same is true for the closure of the operator $T_l^+(\mathrm{i}E_+ + \mathrm{i}E_-)$.

The closure $\overline{L}$ of the operator $L = T_{\sigma\varepsilon}(E_+ - E_-)$ for other irreducible representations of the quantum algebra $U_q(\mathrm{su}_{1,1})$ is also not selfadjoint and has deficiency indices (1,1) or (2,2). An explicit form for the deficiency subspaces is not known for these cases.

4. The regular representation of $U_q(\mathrm{su}_{1,1})$

4.1. *The space* $F_\beta(\mathrm{SU}_q(1,1))$

The quantum group $\mathrm{SL}_q(2,\mathbf{C})$ is defined by the algebra A of functions on $\mathrm{SL}_q(2,\mathbf{C})$. This algebra is generated by elements x, u, v, y satisfying the relations

$$uv = vu, \qquad ux = \sqrt{q}xu, \qquad vx = \sqrt{q}xv, \qquad yu = \sqrt{q}uy, \qquad yv = \sqrt{q}vy, \tag{15}$$

$$xy - q^{-1/2}uv = I, \qquad yx - \sqrt{q}uv = I. \tag{16}$$

Finite linear combinations of the products $x^n u^m v^k y^r$, $n, m, k, r \in \mathbf{Z}_+$, are elements of A. The structure of a Hopf algebra is introduced into A which is uniquely defined by the formulas

$$\Delta\begin{pmatrix} x & u \\ v & y \end{pmatrix} = \begin{pmatrix} x\otimes 1 & u\otimes 1 \\ v\otimes 1 & y\otimes 1 \end{pmatrix}\begin{pmatrix} 1\otimes x & 1\otimes u \\ 1\otimes v & 1\otimes y \end{pmatrix},$$

$$\varepsilon\begin{pmatrix} x & u \\ v & y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad S\begin{pmatrix} x & u \\ v & y \end{pmatrix} = \begin{pmatrix} y & -\sqrt{q}u \\ -q^{-1/2}v & x \end{pmatrix},$$

where Δ, ε, S are respectively the comultiplication, the counit and the antipode [24].

Let β be a fixed real number such that $0 < \beta < 1$. We suppose that $0 < q \equiv \exp h < 1$ and define the set M_β of the points $q^{k+\beta}$, $k \in \mathbf{Z}$. Let $F(M_\beta)$ be the algebra of all complex functions on M_β [30]. We extend the algebra of functions $F(M_\beta)$ by means of the algebra A. Namely, let $F_\beta(\mathrm{SL}_q(2,\mathbf{C}))$ be the algebra with unity, generated by the commutative algebra $F(M_\beta)$ and by the elements x, u, v, y satisfying relations (15), (16) and the relations

$$\begin{aligned} xf(z) &= f(q^{-1}z)x, \qquad vf(z) = f(z)v, \qquad uf(z) = f(z)u, \\ yf(z) &= f(qz)y, \qquad uv = vu = z, \end{aligned}$$

where $f \in F(M_\beta)$. The formulas

$$x^* = y, \qquad y^* = x, \qquad u^* = q^{-1/2}v, \qquad v^* = q^{1/2}u, \qquad f(z)^* = \overline{f(z)}$$

define a *-structure on $F_\beta(\mathrm{SL}_q(2,\mathbf{C}))$ which gives the algebra $F_\beta(\mathrm{SU}_q(1,1))$. The functions of $F_\beta(\mathrm{SU}_q(1,1))$ can be uniquely expanded in the form [24,30]

$$\begin{aligned} f &= f_{00} + \sum_{m\geqslant 0,\, n>0} f_{mn}(z)x^m u^n + \sum_{m>0,\, n\leqslant 0} f_{mn}(z)x^m v^{-n} \\ &\quad + \sum_{m<0,\, n\geqslant 0} f_{mn}(z)y^{-m}u^n + \sum_{m\leqslant 0,\, n<0} f_{mn}(z)y^{-m}v^{-n}, \end{aligned} \tag{17}$$

where the coefficients $f_{mn}(z)$ belong to $F(M_\beta)$.

The algebra A with the *-structure

$$x^* = y, \qquad y^* = x, \qquad u^* = q^{-1/2}v, \qquad v^* = q^{1/2}u$$

defines the algebra of functions on the quantum group $\mathrm{SU}_q(1,1)$ and is denoted by $A(\mathrm{SU}_q(1,1))$. It is clear that $A(\mathrm{SU}_q(1,1))$ is canonically imbedded into $F_\beta(\mathrm{SU}_q(1,1))$.

4.2. *The regular representation of* $U_q(\mathrm{su}_{1,1})$

Let $U_q(\mathrm{su}_{1,1})'$ be the algebra of linear functionals on $U_q(\mathrm{su}_{1,1})$. The formulas

$$(\Delta' f)(a\otimes b) = f(ab), \qquad (S'f)(a) = f(Sa), \qquad \varepsilon'(f) = f(I) \tag{18}$$

give the comultiplication Δ', the antipode S' and the counit ε', where $f \in U_q(\mathrm{su}_{1,1})'$, $a, b \in U_q(\mathrm{su}_{1,1})$. The algebra $U_q(\mathrm{su}_{1,1})'$ with these operations is a Hopf algebra. We have $A(\mathrm{SU}_q(1,1)) \subset U_q(\mathrm{su}_{1,1})'$.

In order to define the regular representation of the algebra $U_q(\mathrm{su}_{1,1})$ on the space $F_\beta(\mathrm{SU}_q(1,1))$ we give its representations on $U_q(\mathrm{su}_{1,1})'$ and on $A(\mathrm{SU}_q(1,1))$. If $f \in U_q(\mathrm{su}_{1,1})'$ and $a \in U_q(\mathrm{su}_{1,1})$, then we write $f(a) = \langle a, f\rangle$. The regular representations R_0 and L_0 of $U_q(\mathrm{su}_{1,1})$ on $U_q(\mathrm{su}_{1,1})'$ are defined by the formulas

$$\langle b, R_0(a)f\rangle = \langle ba, f\rangle, \qquad \langle b, L_0(a)\rangle = \langle S^{-1}(a)b, f\rangle,$$

where $f \in U_q(\mathrm{su}_{1,1})'$, $a, b \in U_q(\mathrm{su}_{1,1})$. It is evident that

$$R_0(a)f = \langle \mathrm{id} \otimes a, \Delta'(f)\rangle, \qquad L_0(a)f = \langle S^{-1}(a) \otimes \mathrm{id}, \Delta'(f)\rangle.$$

Therefore,

$$\begin{aligned} &R_0 : A(\mathrm{SU}_q(1,1)) \to A(\mathrm{SU}_q(1,1)), \\ &L_0 : A(\mathrm{SU}_q(1,1)) \to A(\mathrm{SU}_q(1,1)). \end{aligned} \tag{19}$$

We provide the space $F(M_\beta)$ with the topology of pointwise convergence and the space $F_\beta(\mathrm{SU}_q(1,1))$ with the weakest topology for which all the mappings $j_{rs} : f \to f_{rs}(z)$ (see formula (17)) are continuous. It is possible to show [30] that the space $A(\mathrm{SU}_q(1,1))$ is everywhere dense in $F_\beta(\mathrm{SU}_q(1,1))$. For this reason the operators $R_0(a)$, $L_0(a)$, $a \in U_q(\mathrm{su}_{1,1})$, from (19) can be uniquely extended to be the operators $R(a)$, $L(a)$ on the space $F_\beta(\mathrm{SU}_q(1,1))$.

If m is multiplication on the algebra $F \equiv F_\beta(\mathrm{SU}_q(1,1))$, i.e. $m : F \otimes F \to F$, and $m^0 = m\sigma$, where $\sigma(f \otimes g) = g \otimes f$, $f, g \in F$, then for the representations R and L we have

$$m(R \otimes R) = Rm, \qquad m^0(L \otimes L) = Lm^0.$$

The operators $R(a)$, $L(a)$, $a \in U_q(\mathrm{su}_{1,1})$, are constructed in the following way [30]. The elements $R(E_\pm)$, $R(k^{\pm 1})$ act on x, u, v, y as

$$R(E_+)\begin{pmatrix} x & u \\ v & y \end{pmatrix} = \begin{pmatrix} 0 & x \\ 0 & v \end{pmatrix}, \qquad R(E_-)\begin{pmatrix} x & u \\ v & y \end{pmatrix} = \begin{pmatrix} u & 0 \\ y & 0 \end{pmatrix}, \tag{20}$$

$$R(k^{\pm 1})\begin{pmatrix} x & u \\ v & y \end{pmatrix} = \begin{pmatrix} q^{1/2}x & q^{-1/2}u \\ q^{1/2}v & q^{-1/2}y \end{pmatrix}. \tag{21}$$

The operators $L(E_\pm)$, $L(k^{\pm 1})$ are of an analogous form. According to formulas (18) and to the definition of the comultiplication in $U_q(\mathrm{su}_{1,1})$, for $f, g \in F(\mathrm{SU}_q(1,1))$ we have

$$R(E_\pm)fg = R(E_\pm)f \cdot R(k)g + R(k^{-1})f \cdot R(E_\pm)g, \tag{22}$$

$$R(k^{\pm 1})fg = R(k^{\pm 1})f \cdot R(k^{\pm 1})g, \tag{23}$$

$$L(E_\pm)fg = L(E_\pm)f \cdot L(k)g + L(k^{-1})f \cdot L(E_\pm)g, \tag{24}$$
$$L(k^{\pm 1})fg = L(k^{\pm 1})f \cdot L(k^{\pm 1})g. \tag{25}$$

Formulas (20)–(25) allow us to extend the action of the operators $R(E_\pm)$, $R(k^{\pm 1})$ upon all elements of the space $F_\beta(\mathrm{SU}_q(1,1))$. We have

$$\big(R(a)f\big)^* = R\big(S(a)^*\big)f^*, \qquad \big(L(a)f\big)^* = L\big(S^{-1}(a)^*\big)f^*,$$

where $a \in U_q(\mathrm{su}_{1,1})$.

4.3. *The Hilbert space* $L^2_\beta(\mathrm{SU}_q(1,1))$

The formula

$$\nu_\beta(f) = 2\big(q^{1/2} - q^{-1/2}\big)\sum_{n\in\mathbf{Z}} q^{n+\beta} f_{00}\big(q^{n+\beta}\big),$$

where f_{00} is defined by formula (17), gives a linear functional on $F_\beta(\mathrm{SU}_q(1,1))$, which is denoted by $\int f\,\mathrm{d}\nu_\beta$. It is proved [30] that the functional $f \to \int f\,\mathrm{d}\nu_\beta$ is an invariant integral on $F_\beta(\mathrm{SU}_q(1,1))$ in the sense that for $a \in U_q(\mathrm{su}_{1,1})$ and $f \in F_\beta(\mathrm{SU}_q(1,1))$ we have

$$\int \big(R(a)f\big)\,\mathrm{d}\nu_\beta = \int \big(L(a)f\big)\,\mathrm{d}\nu_\beta = \varepsilon(a)\int f\,\mathrm{d}\nu_\beta.$$

Let $L^2_\beta(\mathrm{SU}_q(1,1))$ be the subspace of the space $F_\beta(\mathrm{SU}_q(1,1))$ consisting of functions f for which $\int ff^*\,\mathrm{d}\nu_\beta < \infty$. The formula $(f,g)_\beta = \int fg^*\,\mathrm{d}\nu_\beta$ defines the scalar product on the Hilbert space $L^2_\beta(\mathrm{SU}_q(1,1))$.

Restriction of the operators $R(E_\pm)$, $R(k^{\pm 1})$ onto the space $L^2_\beta(\mathrm{SU}_q(1,1))$ gives unbounded closed operators with a common domain of definition. Thus, we have a representation of $U_q(\mathrm{su}_{1,1})$ on $L^2_\beta(\mathrm{SU}_q(1,1))$ which is denoted by R_β. In a similar way the representation L_β is defined. The representations R_β and L_β are unitary, that is $R_\beta(a)^* = R_\beta(a^*)$, $L_\beta(a^*) = L_\beta(a)^*$ for all $a \in U_q(\mathrm{su}_{1,1})$.

Since $F(M_\beta) \subset F_\beta(\mathrm{SU}_q(1,1))$, we can define the space

$$L^2(M_\beta) = F_\beta(M_\beta) \cap L^2_\beta\big(\mathrm{SU}_q(1,1)\big).$$

The functions $f \in L^2(M_\beta)$ are quantum analogues of the functions on the classical group $\mathrm{SU}(1,1)$ that are two-sided invariant with respect to the subgroup of diagonal matrices.

4.4. *Matrix elements of representations of* $U_q(\mathrm{su}_{1,1})$

Matrix elements $t^{\sigma\varepsilon}_{mn}$ of the representations $T_{\sigma\varepsilon}$ of the quantum algebra $U_q(\mathrm{sl}_2)$, $0 < q < 1$, in the basis $\{\mathbf{e}_m\}$ can be defined in the same way as in the case of finite-dimensional representations [25]. They are expressed in terms of the matrix elements x, u, v, y of the two-dimensional representation $T_{1/2}$ of $U_q(\mathrm{su}_{1,1})$. For $t^{\sigma\varepsilon}_{mn}$ we have the expressions [25]

$$\begin{aligned} t^{\sigma\varepsilon}_{mn} &= \sqrt{c_n/c_m}\, q^{(n-m)(\sigma+n)/2} \frac{(q^{\sigma-m+1};q)_{m-n}}{(q;q)_{m-n}} x^{-m-n} v^{m-n} \\ &\quad \times {}_2\Phi_1(\sigma-n+1, -\sigma-n; m-n+1; q, \zeta) \end{aligned} \tag{26}$$

if $n \leqslant m \leqslant -n$,

$$\begin{aligned} t^{\sigma\varepsilon}_{mn} &= \sqrt{c_n/c_m}\, q^{(n-m)(\sigma+m)/2} \frac{(q^{\sigma+m+1};q)_{n-m}}{(q;q)_{n-m}} x^{-m-n} u^{n-m} \\ &\quad \times {}_2\Phi_1(\sigma-m+1, -\sigma-m; n-m+1; q, \zeta) \end{aligned} \tag{27}$$

if $m \leqslant n \leqslant -m$,

$$\begin{aligned} t^{\sigma\varepsilon}_{mn} &= \sqrt{c_n/c_m}\, q^{(m-n)(\sigma-n)/2} \frac{(q^{\sigma+m+1};q)_{n-m}}{(q;q)_{n-m}} \\ &\quad \times {}_2\Phi_1(\sigma+n+1, -\sigma+n; n-m+1; q, \zeta) u^{n-m} y^{m+n} \end{aligned} \tag{28}$$

if $-n \leqslant m \leqslant n$,

$$\begin{aligned} t^{\sigma\varepsilon}_{mn} &= \sqrt{c_n/c_m}\, q^{(n-m)(\sigma-m)/2} \frac{(q^{\sigma-m+1};q)_{m-n}}{(q,q)_{m-n}} \\ &\quad \times {}_2\Phi_1(\sigma+m+1, -\sigma+m; m-n+1; q, \zeta) v^{m-n} y^{m+n} \end{aligned} \tag{29}$$

if $-m \leqslant n \leqslant m$, where $\zeta = -q^{-1/2}uv$, $c_{p+1}/c_p = q^{-p+1/2}$ and the basic hypergeometric function ${}_2\Phi_1$ is defined by formula

$$ {}_2\Phi_1(a,b;c;q,x) \equiv {}_2\varphi_1\big(q^a,q^b;q^c;q,x\big) = \sum_{n=0}^{\infty} \frac{(q^a;q)_n (q^b;q)_n}{(q^c;q)_n} \frac{x^n}{(q;q)_n}, $$

$$ (d;q)_n = \prod_{j=0}^{n-1} \big(1-dq^j\big), \quad (d;q)_0 = 1. $$

By putting in (26)–(29) $\sigma = \mathrm{i}\rho - 1/2$ we obtain the matrix elements of the principal unitary series representations of the quantum group $\mathrm{SU}_q(1,1)$. In order to obtain matrix elements for the other series of representations of $\mathrm{SU}_q(1,1)$ we must multiply the corresponding matrix elements (26)–(29) by appropriate multipliers which do not depend on x, u, v, y (see the paper [17]).

4.5. *The spherical transform of functions on M_β*

By putting $m=n=0$ in formulas (26)–(29) we obtain the zonal spherical functions of the representations $T_{\sigma\varepsilon}$, $\varepsilon=0$:

$$t_{00}^{\sigma 0} = {}_2\varphi_1\big(q^{-l}, q^{l+1}; q; q, \zeta\big) = {}_2\Phi_1(-l, l+1; 1; q, \zeta).$$

The formula

$$\hat{f}_\beta(\sigma) = \int \big(t_{00}^{\sigma 0}\big)^* f \, \mathrm{d}\nu_\beta \tag{30}$$

defines the q-spherical transform of functions f from $L^2(M_\beta)$. It is the q-analogue of the well known Mehler–Fock transform. The inverse transform is given by formulas [29,30]

$$\begin{aligned} f(z) &= \int_0^{\pi/h} \hat{f}_\beta(\mathrm{i}\rho - 1/2) t_{00}^{\mathrm{i}\rho-1/2,0} \mu_\beta(\rho)\,\mathrm{d}\rho \\ &\quad + \sum_{m\in\mathbf{Z}, m>-\beta} [2m+2\beta] \hat{f}_\beta\Big(m+\beta+\frac{\mathrm{i}\pi}{h}-\frac{1}{2}\Big) t_{00}^{m+\beta+\mathrm{i}\pi/h-1/2,0}, \end{aligned} \tag{31}$$

where the Plancherel measure $\mu_\beta(\rho)\,\mathrm{d}\rho$ is given by the formula [30]

$$\begin{aligned} \mu_\beta(\rho) &= \frac{h}{2\pi} \frac{\mathrm{e}^{-h/4}}{q^{1/2}-q^{-1/2}} \frac{\Gamma_q^2(\mathrm{i}\rho-1/2)\Gamma_q^2(-\mathrm{i}\rho+1/2)}{\Gamma_q(2\mathrm{i}\rho)\Gamma_q(-2\mathrm{i}\rho)} \\ &\quad \times \frac{\theta^2(\beta,\tau)}{\theta(\beta+\mathrm{i}\rho-1/2,\tau)\theta(\beta-\mathrm{i}\rho+1/2,\tau)}, \\ q &= \mathrm{e}^{-h}, \ \tau = \frac{2\mathrm{i}\pi}{h}, \ 0<q<1. \end{aligned}$$

Here $\Gamma_q(x)$ is the q-gamma function and $\theta(z,\tau)$ is the Jacobi theta function:

$$\Gamma_q(x) = \frac{(q;q)_\infty}{(q^x;q)_\infty}(1-q)^{1-x}, \qquad \theta(z,\tau) = \sum_{k\in\mathbf{Z}} \exp\big(2\pi i k z + \pi i k^2 \tau\big).$$

The Plancherel formula

$$\begin{aligned} \int \big|f(z)\big|^2 \mathrm{d}\nu_\beta &= \int_0^{\pi/h} \big|\hat{f}_\beta(\mathrm{i}\rho-1/2)\big|^2 \mu_\beta(\rho)\,\mathrm{d}\rho \\ &\quad + \sum_{m\in\mathbf{Z}, m>-\beta} [2m+2\beta] \Big|\hat{f}_\beta\Big(m+\beta+\frac{\pi\mathrm{i}}{h}-\frac{1}{2}\Big)\Big|^2 \end{aligned}$$

is valid for the transform (30).

Let us note that the integral in (31) corresponds to the representations of the principal unitary series of $U_q(\mathrm{su}_{1,1})$ and the sum contains matrix elements of representations of the strange series.

5. Representations of the quantum algebras $U_q(\mathrm{su}_{r,s})$

5.1. *The quantum algebras* $U_q(\mathrm{su}_{r,s})$

The quantum algebra $U_q(\mathrm{sl}_n)$ is generated by elements [26]

$$k_i \equiv q^{h_i/4}, k_i^{-1} \equiv q^{-h_i/4}, e_i, f_i, \quad i = 1, 2, \ldots, n-1,$$

satisfying the relations

$$k_i k_i^{-1} = k_i^{-1} k_i = 1, \qquad k_i k_j = k_j k_i, \tag{32}$$

$$k_i e_j k_i^{-1} = q^{a_{ij}/4} e_j, \qquad k_i f_j k_i^{-1} = q^{-a_{ij}/4} f_j, \tag{33}$$

$$[e_i, f_j] = \delta_{ij} \frac{k_i^2 - k_i^{-2}}{q^{1/2} - q^{-1/2}} \equiv \delta_{ij}[h_i], \tag{34}$$

$$[e_i, e_j] = [f_i, f_j] = 0, \quad |i - j| > 1, \tag{35}$$

$$e_i^2 e_{i\pm 1} - \left(q^{1/2} + q^{-1/2}\right) e_i e_{i\pm 1} e_i + e_{i\pm 1} e_i^2 = 0, \tag{36}$$

$$f_i^2 f_{i\pm 1} - \left(q^{1/2} + q^{-1/2}\right) f_i f_{i\pm 1} f_i + f_{i\pm 1} f_i^2 = 0, \tag{37}$$

where $a_{ii} = 2$, $a_{i,i\pm 1} = -1$ and $a_{ij} = 0$ for $|i - j| > 1$. The last two relations are called the q-Serre relations. The structure of a Hopf algebra is introduced on $U_q(\mathrm{sl}_n)$. The coalgebra structure allows us the composition of representations whereas the universal R-matrix is an intertwiner between the switched tensor products. We shall only deal with representations of $U_q(\mathrm{sl}_n)$ considered as an associative algebra generated by k_i, k_i^{-1}, e_i, f_i, $i = 1, 2, \ldots, n-1$.

We can equip $U_q(\mathrm{sl}_n)$ with $*$-structures. Different $*$-structures define different real forms of $U_q(\mathrm{sl}_n)$. The $*$-structure

$$\left(k_i^{\pm 1}\right)^* = k_i^{\pm 1}, \; e_i^* = f_i, \; f_i^* = e_i$$

gives the compact real form $U_q(\mathrm{su}_n)$ of the quantum algebra $U_q(\mathrm{sl}_n)$. The $*$-structure

$$\left(k_i^{\pm 1}\right)^* = k_i^{\pm 1}, \; e_p^* = -f_p, \; f_p^* = -e_p, \tag{38}$$

$$e_i^* = f_i, \; f_i^* = e_i, \; i \neq p, \tag{39}$$

defines the noncompact real form $U_q(\mathrm{su}_{p,n-p})$. There are other real forms of $U_q(\mathrm{sl}_n)$. They will not be considered below.

The irreducible finite-dimensional representations of the quantum algebra $U_q(\mathrm{su}_n)$ (q is not a root of unity) are deformations of the corresponding irreducible finite-dimensional representations of the Lie algebra su_n. These representations of $U_q(\mathrm{su}_n)$ can be given by explicit formulas for the representation operators with respect to the q-analogue of the Gel'fand–Tsetlin basis. These formulas are obtained from the corresponding Gel'fand–

Tsetlin formulas for representations of su_n by replacement of all expressions of the type $(m_{ik} - m_{js} \pm a)$ by $[m_{ik} - m_{js} \pm a]$ respectively, where $[b]$ is a q-number.

We are interested in representations of the quantum algebra $U_q(\mathrm{su}_{r,s})$ with q not coinciding with a root of unity. The operators of these representations must satisfy relations (38) and (39). These representations will be called infinitesimal unitary (we shall omit the word "infinitesimal"). Along with them we shall consider representations of the associative algebra $U_q(\mathrm{su}_{r,s})$ for which conditions (38) and (39) may be violated. They are not representations of the *-algebra $U_q(\mathrm{su}_{r,s})$. But we need them to separate irreducible unitary representations of $U_q(\mathrm{su}_{r,s})$. Therefore, we give the following definition of representations of $U_q(\mathrm{su}_{r,s})$.

A representation T of the associative algebra $U_q(\mathrm{su}_{r,s})$ is an algebraic homomorphism from $U_q(\mathrm{su}_{r,s})$ into an algebra of linear (bounded or unbounded) operators on a Hilbert space H for which the following conditions are fulfilled:

(a) the restriction of T onto the subalgebra $U_q(\mathrm{u}_r + \mathrm{u}_s)$ decomposes into a direct sum of finite-dimensional irreducible representations (given by integral highest weights) with finite multiplicities;

(b) operators of a representation T are defined on everywhere dense subspace V of H, containing all subspaces which are carrier spaces of irreducible finite-dimensional representations of $U_q(\mathrm{u}_r + \mathrm{u}_s)$ from the restriction of T to this subalgebra.

In other words, our representations T of $U_q(\mathrm{su}_{r,s})$ are Harish-Chandra modules of $U_q(\mathrm{su}_{r,s})$ with respect to $U_q(\mathrm{u}_r + \mathrm{u}_s)$.

A representation T of $U_q(\mathrm{su}_{r,s})$ on H is called irreducible if H has no non-trivial invariant subspaces with closure coinciding with H. If the operators of a representation T satisfy relations (38) and (39) on the common domain V of definition, then T is called a unitary representation. To define a representation T of $U_q(\mathrm{su}_{r,s})$ it is sufficient to give the operators $T(k^{\pm 1})$, $T(e_i)$, $T(f_i)$, $i = 1, 2, \ldots, n-1$, satisfying relations (32)–(37).

5.2. *Representations of* $U_q(\mathrm{su}_{n,1})$

We consider representations of $U_q(\mathrm{su}_{n,1})$ when q is not a root of unity. Let c_1 and c_2 be complex numbers such that $c_1 + c_2 = m_0$ is an integer, and let $\mathbf{m} = (m_1, m_2, \ldots, m_{n-1})$ be a highest weight of an irreducible representation of the subalgebra $U_q(\mathrm{u}_{n-1})$. The numbers $\mathbf{m}$, c_1, c_2 determine a representation $T(\mathbf{m}, c_1, c_2)$ of the quantum algebra $U_q(\mathrm{su}_{n,1})$. The restriction of $T(\mathbf{m}, c_1, c_2)$ to $U_q(\mathrm{u}_n)$ decomposes into the orthogonal sum of all irreducible representations $T_{\mathbf{m}_n}$ with highest weights $\mathbf{m}_n = (m_{1n}, \ldots, m_{nn})$ such that

$$m_{1n} \geqslant m_1 \geqslant m_{2n} \geqslant m_2 \geqslant \cdots \geqslant m_{n-1} \geqslant m_{nn}.$$

Every one of these representations of $U_q(\mathrm{u}_n)$ is contained in $T(\mathbf{m}, c_1, c_2)$ exactly once. This restriction determines the Hilbert space $V_{\mathbf{m}}$ on which $T(\mathbf{m}, c_1, c_2)$ acts. We choose in $V_{\mathbf{m}}$ the orthonormal basis consisting of the Gel'fand–Tsetlin bases of the carrier subspaces of irreducible representations $T_{\mathbf{m}_n}$ of $U_q(\mathrm{u}_n)$. These basis elements are denoted by $|\mathbf{m}_n, \alpha\rangle$, where α are the corresponding Gel'fand–Tsetlin patterns of the representation $T_{\mathbf{m}_n}$. The operators $T(e_n)$, $T(f_n)$ of the representation $T(\mathbf{m}, c_1, c_2)$ act upon $|\mathbf{m}_n, \alpha\rangle$ by the formulas

$$T(e_n)|\mathbf{m}_n,\alpha\rangle = \sum_{s=1}^{n}[l_{sn}-c_1]\omega_s(\mathbf{m},\mathbf{m}_n,\alpha)\big|\mathbf{m}_n^{+s},\alpha\big\rangle,$$

$$T(f_n)|\mathbf{m}_n,\alpha\rangle = \sum_{s=1}^{n}[-l_{sn}+c_2+1]\omega_s\big(\mathbf{m},\mathbf{m}_n^{-s},\alpha\big)\big|\mathbf{m}_n^{-s},\alpha\big\rangle,$$

where

$$\omega_s(\mathbf{m},\mathbf{m}_n,\alpha) = \left(-\frac{\prod_{j=1}^{n-1}[l_{j,n-1}-l_{sn}-1][l_j-l_{sn}]}{\prod_{r\neq s}[l_{sn}-l_{rn}+1][l_{sn}-l_{rn}]}\right)^{1/2},$$

$$l_j = m_j - j - 1,\ j=1,2,\ldots,n-1;\ l_{sk}=m_{sk}-s,\ s=1,2,\ldots,k,$$

$$\mathbf{m}_n^{\pm s} = (m_{1n},\ldots,m_{s-1,n},m_{sn}\pm 1,m_{s+1,n},\ldots,m_{nn}).$$

The operator $T(h_n)$ is given by the formula

$$T(h_n)|\mathbf{m}_n,\alpha\rangle = \left(c_1+c_2+n+2+\sum_{j=1}^{n-1}m_j+\sum_{j=1}^{n-1}m_{j,n-1}-2\sum_{j=1}^{n}m_{jn}\right).$$

The other generators e_i, f_j, h_k belong to the subalgebra $U_q(\mathrm{u}_n)$ and the operators $T(e_i)$, $T(f_j)$, $T(h_k)$ act upon the basis elements $|\mathbf{m}_n,\alpha\rangle$ by Jimbo's formulas [14].

Because of periodicity of the function $w(z)=[z]$ the representations

$$T(\mathbf{m},c_1,c_2) \quad \text{and} \quad T(\mathbf{m},c_1+4\pi \mathrm{i}k/h, c_2-4\pi \mathrm{i}k/h), \qquad k\in\mathbf{Z},$$

coincide and the representations

$$T(\mathbf{m},c_1,c_2) \quad \text{and} \quad T(\mathbf{m},c_1+2\pi \mathrm{i}k/h, c_2-2\pi \mathrm{i}k/h), \qquad k\in\mathbf{Z},$$

are equivalent when $q=\exp h$, $h\in\mathbf{R}$. If $q=\exp \mathrm{i}h$, $h\in\mathbf{R}$, then the representations

$$T(\mathbf{m},c_1,c_2) \quad \text{and} \quad T(\mathbf{m},c_1+4\pi k/h, c_2-4\pi k/h), \qquad k\in\mathbf{Z},$$

coincide and the representations

$$T(\mathbf{m},c_1,c_2) \quad \text{and} \quad T(\mathbf{m},c_1+2\pi k/h, c_2-2\pi k/h), \qquad k\in\mathbf{Z},$$

are equivalent. Therefore, we suppose that

$$0\leqslant \operatorname{Im} c_1 < 2\pi/h$$

if $q=\exp h$, $h\in\mathbf{R}$, and

$$0\leqslant \operatorname{Re} c_1 < 2\pi/h$$

if $q=\exp \mathrm{i}h$, $h\in\mathbf{R}$.

THEOREM [17]. *The representation $T(\mathbf{m}, c_1, c_2)$ is irreducible if and only if c_1 and c_2 are not integers or c_1 and c_2 coincide with some of the numbers $l_1, l_2, \dots, l_{n-1}$.*

The irreducible representations $T(\mathbf{m}, c_1, c_2)$ and $T(\mathbf{m}, c_2, c_1)$ are equivalent. The intertwining operator for these representations is constructed in the same way as in the classical case [16]. Reducible representations $T(\mathbf{m}, c_1, c_2)$ and $T(\mathbf{m}, c_2, c_1)$ contain equivalent irreducible representations of $U_q(\mathrm{su}_{n,1})$. A reducible representation $T(\mathbf{m}, c_1, c_2)$ contains a finite number of irreducible components. Reducible representations $T(\mathbf{m}, c_1, c_2)$ of $U_q(\mathrm{su}_{n,1})$ have exactly the same structure as the corresponding representations of the group $\mathrm{SU}(n, 1)$ [16]. Moreover, there is the one-to-one correspondence between irreducible components of a representation $T(\mathbf{m}, c_1, c_2)$ of $U_q(\mathrm{su}_{n,1})$ and irreducible components of a representation $T(\mathbf{m}, c_1, c_2)$ of $\mathrm{SU}(n, 1)$.

5.3. *Unitary irreducible representations of $U_q(\mathrm{su}_{n,1})$*

If $q = \exp h$, $h \in \mathbf{R}$, then the following irreducible representations $T(\mathbf{m}, c_1, c_2)$ of $U_q(\mathrm{su}_{n,1})$ are unitary [17]:

(a) the representations $T(\mathbf{m}, c_1, c_2)$, $c_1 = \overline{c_2}$ (the principal unitary series);
(b) the representations $T(\mathbf{m}, c_1, c_2)$, $\mathrm{Im}\, c_1 = -\mathrm{Im}\, c_2 = \pi/h$ (the strange series);
(c) the representations $T(\mathbf{m}, c_1, c_2)$, where c_1 and c_2 are real numbers for which there exist numbers $l_k = m_k - k - 1$, $l_s = m_s - s - 1$, $k, s = 1, 2, \dots, n-1$, such that $|l_k - c_1| < 1$, $|l_s - c_2| < 1$ and $m_k = m_{k+1} = \dots = m_s$ (for $c_1 > c_2$) or $m_s = m_{s+1} = \dots = m_k$ (for $c_1 < c_2$) (the supplementary series).

Along with these unitary representations there are unitary irreducible representations of $U_q(\mathrm{su}_{n,1})$ which are irreducible components of reducible representations $T(\mathbf{m}, c_1, c_2)$. There is the one-to-one correspondence between nonequivalent unitary irreducible representations of $U_q(\mathrm{su}_{n,1})$ which are irreducible components of $T(\mathbf{m}, c_1, c_2)$ and nonequivalent unitary irreducible representations of the same type of the group $\mathrm{SU}(n, 1)$. The list of the last representations can be found in [16].

If $q = \exp ih$, $h \in \mathbf{R}$, and q is not a root of unity, then the following irreducible representations of $U_q(\mathrm{su}_{n,1})$ are unitary:

(a) the representations $T(\mathbf{m}, c_1, c_2)$, $c_2 = \overline{c_1}$ (the principal unitary series);
(b) the representations $T(\mathbf{m}, c_1, c_2)$, $\mathrm{Re}\, c_1 = -\mathrm{Re}\, c_2 = \pi/h$ (the strange series).

These representations and the zero representation exhaust all irreducible unitary representations of $U_q(\mathrm{su}_{n,1})$ in this case. Remark that the strange series of representations disappears when $q \to 1$.

5.4. *Representations of the degenerate series of the quantum algebra $U_q(\mathrm{su}_{r,s})$*

Representations of the principal degenerate series of $U_q(\mathrm{su}_{r,s})$ are given by two complex numbers λ_1 and λ_2 such that $\lambda_1 + \lambda_2$ is an integer, or by the numbers

$$\mu = (-\lambda_1 + \lambda_2)/2 \quad \text{and} \quad \Lambda = \lambda_1 + \lambda_2.$$

We denote the corresponding representations by $T_{\Lambda\mu}$. The space $V_{\Lambda\mu}$ of the representation $T_{\Lambda\mu}$ is the orthogonal sum of the subspaces $V(m_{1r}, \mathbf{0}, m_{rr};\ m_{1s}, \mathbf{0}, m_{ss})$ which are the carrier spaces of the finite-dimensional irreducible representations of $U_q(\mathrm{u}_r + \mathrm{u}_s)$ with highest weights $(m_{1r}, \mathbf{0}, m_{rr}; m_{1s}, \mathbf{0}, m_{ss})$ such that

$$m_{1r} + m_{rr} + m_{1s} + m_{ss} = \Lambda. \tag{40}$$

Here $(m_{1r}, \mathbf{0}, m_{rr})$ and $(m_{1s}, \mathbf{0}, m_{ss})$, $m_{1r} \geqslant 0 \geqslant m_{rr}$, $m_{1s} \geqslant 0 \geqslant m_{ss}$, are highest weights of irreducible representations of the subalgebras $U_q(\mathrm{u}_r)$ and $U_q(\mathrm{u}_s)$, respectively, and $\mathbf{0}$ denotes a set of zero coordinates.

If $r = 2$ then instead of $(m_{1r}, \mathbf{0}, m_{rr})$ we have the highest weights (m_{12}, m_{22}), $m_{12} \geqslant m_{22}$, and here m_{12} and m_{22} may be both positive or both negative. The same remark is valid for the highest weights $(m_{1s}, \mathbf{0}, m_{ss})$ if $s = 2$.

We choose in the subspaces $V(m_{1r}, \mathbf{0}, m_{rr};\ m_{1s}, \mathbf{0}, m_{ss})$ orthonormal bases which are products of the Gel'fand–Tsetlin bases corresponding to irreducible representations of the quantum subalgebras $U_q(\mathrm{u}_r)$ and $U_q(\mathrm{u}_s)$. Elements of such bases are labelled by double Gel'fand–Tsetlin patterns

$$|M\rangle \equiv \begin{pmatrix} m_{1r} & & \mathbf{0} & & m_{rr} \\ & j & \mathbf{0} & j' & \\ & a & \mathbf{0} & a' & \\ & & \cdots & & \end{pmatrix} \begin{pmatrix} m_{1s} & & \mathbf{0} & & m_{ss} \\ & k & \mathbf{0} & k' & \\ & b & \mathbf{0} & b' & \\ & & \cdots & & \end{pmatrix}. \tag{41}$$

If $r = 2$ or $s = 2$ then the corresponding Gel'fand–Tsetlin pattern is replaced by

$$\begin{pmatrix} m_{12} & & m_{22} \\ & j & \end{pmatrix}.$$

It is clear that the numbers from pattern (41) satisfy the conditions

$$m_{1r} \geqslant j \geqslant a \geqslant \cdots \geqslant a' \geqslant j' \geqslant m_{rr}, \tag{42}$$
$$m_{1s} \geqslant k \geqslant b \geqslant \cdots \geqslant b' \geqslant k' \geqslant m_{ss} \tag{43}$$

or the condition $m_{12} \geqslant j \geqslant m_{22}$ if $r = 2$ or $s = 2$. Thus, elements of the orthonormal basis of the carrier space of the representation $T_{\Lambda\mu}$ are labelled by all patterns (41) satisfying the betweenness conditions (42), (43) and equality (40).

The operators corresponding to elements of the subalgebras $U_q(\mathrm{u}_r)$ and $U_q(\mathrm{u}_s)$ act upon basis elements by Jimbo's formulas [14]. To define completely the representation $T_{\Lambda\mu}$ we have to give the operators $T_{\Lambda\mu}(e_r)$, $T_{\Lambda\mu}(f_r)$. They are of the form, [18]:

$$\begin{aligned} T_{\Lambda\mu}(e_r)|M\rangle &= [\lambda_2 - m_{rr} - m_{1s} - s + 1]A(m_{1r}, m_{rr})B(m_{1s}, m_{ss}) \\ &\quad \times \left|M\left(m_{1r}^{+1}, m_{1s}^{-1}\right)\right\rangle + [-\lambda_1 + m_{1r} + m_{1s}]A(m_{1r}, m_{rr}) \\ &\quad \times D(m_{1s}, m_{ss})\left|M\left(m_{1r}^{+1}, m_{ss}^{-1}\right)\right\rangle \\ &\quad + [\lambda_2 - m_{1r} - m_{1s} - r - s + 2]C(m_{1r}, m_{rr}) \end{aligned}$$

$$
\begin{aligned}
&\times B(m_{1s}, m_{ss})\big|M\big(m_{rr}^{+1}, m_{1s}^{-1}\big)\big\rangle \\
&+ [\lambda_1 + m_{rr} + m_{1s} - r + 1]C(m_{1r}, m_{rr})D(m_{1s}, m_{ss}) \\
&\times \big|M\big(m_{rr}^{+1}, m_{ss}^{-1}\big)\big\rangle,
\end{aligned} \tag{44}
$$

$$
\begin{aligned}
T_{\Lambda\mu}(f_r)|M\rangle = {} & [-\lambda_1 + m_{rr} + m_{1s} - r + 1]A\big(m_{1r}^{-1}, m_{rr}\big)B\big(m_{1s}^{+1}, m_{ss}\big) \\
&\times \big|M\big(m_{1r}^{-1}, m_{1s}^{+1}\big)\big\rangle + [\lambda_2 - m_{1r} - m_{1s} - r - s + 2] \\
&\times A\big(m_{1r}^{-1}, m_{rr}\big)D\big(m_{1s}, m_{ss}^{+1}\big)\big|M\big(m_{1r}^{-1}, m_{ss}^{+1}\big)\big\rangle \\
&+ [-\lambda_1 + m_{1r} + m_{1s}]C\big(m_{1r}, m_{rr}^{-1}\big) \\
&\times B\big(m_{1s}^{+1}, m_{ss}\big)\big|M\big(m_{rr}^{-1}, m_{1s}^{+1}\big)\big\rangle + [\lambda_2 - m_{rr} - m_{1s} - s + 1] \\
&\times C\big(m_{1r}, m_{rr}^{-1}\big)D\big(m_{1s}, m_{ss}^{+1}\big)\big|M\big(m_{rr}^{-1}, m_{ss}^{+1}\big)\big\rangle,
\end{aligned} \tag{45}
$$

where $M(m_{ir}^{\pm 1}, m_{js}^{\mp 1})$ is the pattern obtained from pattern (41) by replacement of m_{ir} and m_{js} respectively by $m_{ir} \pm 1$ and $m_{js} \mp 1$. For $r > 2$ and $s > 2$ the coefficients A, B, C, D are given by the formulas

$$
\begin{aligned}
A(m_{1r}, m_{rr}) &= \left(\frac{[m_{1r} - j' + r - 1][m_{1r} - j + 1]}{[m_{1r} - m_{rr} + r - 1][m_{1r} - m_{rr} + r]}\right)^{1/2}, \\
B(m_{1s}, m_{ss}) &= \left(\frac{[m_{1s} - k][m_{1s} - k' + s - 2]}{[m_{1s} - m_{ss} + s - 1][m_{1s} - m_{ss} + s - 2]}\right)^{1/2}, \\
C(m_{1r}, m_{rr}) &= \left(\frac{[j' - m_{rr}][j - m_{rr} + r - 2]}{[m_{1r} - m_{rr} + r - 1][m_{1r} - m_{rr} + r - 2]}\right)^{1/2}, \\
D(m_{1s}, m_{ss}) &= \left(\frac{[k - m_{ss} + s - 1][k' - m_{ss} + 1]}{[m_{1s} - m_{ss} + s - 1][m_{1s} - m_{ss} + s]}\right)^{1/2},
\end{aligned}
$$

where the $[b]$ are q-numbers. The formula for $A(m_{1r}^{-1}, m_{rr})$ is obtained from the formula for $A(m_{1r}, m_{rr})$ by replacement of m_{1r} by $m_{1r} - 1$. The corresponding notations for B, C and D from formulas (44) and (45) are of the analogous meaning. For $r = 2$ the expressions for $A(m_{1r}, m_{rr})$ and $C(m_{1r}, m_{rr})$ are of the form

$$
\begin{aligned}
A(m_{12}, m_{22}) &= \left(\frac{[m_{12} - j + 1]}{[m_{12} - m_{22} + 1][m_{12} - m_{22} + 2]}\right)^{1/2}, \\
C(m_{12}, m_{22}) &= \left(\frac{[j - m_{22}]}{[m_{12} - m_{22} + 1][m_{12} - m_{22}]}\right)^{1/2}.
\end{aligned}
$$

If $s = 2$ then we have

$$
B(m_{12}, m_{22}) = \left(\frac{[m_{12} - k]}{[m_{12} - m_{22} + 1][m_{12} - m_{22}]}\right)^{1/2},
$$

$$D(m_{12}, m_{22}) = \left(\frac{[k - m_{22} + 1]}{[m_{12} - m_{22} + 1][m_{12} - m_{22} + 2]} \right)^{1/2}.$$

THEOREM. *If q is not a root of unity, then the representation $T_{\Lambda\mu}$ is irreducible if and only if 2μ is not an integer such that $2\mu \equiv \Lambda$ (mod 2).*

The structure of reducible representations $T_{\Lambda\mu}$ is described in paper [18]. The irreducible representations

$$T_{\Lambda\mu} \quad \text{and} \quad T_{\Lambda, -\mu+r+s-1}$$

are equivalent. If these representations are reducible, then they consist of the same irreducible components. The periodicity of the function $w(z) = [z]$ leads to the equivalence relation

$$T_{\Lambda\mu} \sim T_{\Lambda, \mu+2\pi \mathrm{i} k/h}, \quad k \in \mathbf{Z},$$

if $q = \exp h$, $h \in \mathbf{R}$, and to the relation

$$T_{\Lambda\mu} \sim T_{\Lambda, \mu+2\pi k/h}, \quad k \in \mathbf{Z},$$

if $q = \exp \mathrm{i}h$, $h \in \mathbf{R}$.

The following irreducible representations $T_{\Lambda\mu}$ are unitary:

(a) the representations $T_{\Lambda\mu}$, $\mu = \mathrm{i}\rho + (r + s - 1)/2$, $0 < \rho < 2\pi/h$ (the principal degenerate unitary series);

(b) the representations $T_{\Lambda\mu}$ for which $\Lambda = (r + s)$ (mod 2), $(r + s - 1)/2 < \mu < (r + s)/2$ (the supplementary degenerate series);

(c) the representations $T_{\Lambda\mu}$, $\operatorname{Im} \mu = \pi/h$, $\operatorname{Re} \mu \leqslant (r + s - 1)/2$ (the strange series).

There are unitary irreducible representations of $U_q(\mathrm{su}_{r,s})$ which are irreducible components of reducible representations $T_{\Lambda\mu}$ [18].

6. Representations of the q-deformed algebras $U_q(\mathrm{so}_{n,1})$

6.1. *Representations of $U_q(\mathrm{so}_{2,1})$*

We define the q-deformed algebra $U_q(\mathrm{so}(3, \mathbf{C}))$ as the associative algebra generated by elements I_1, I_2, I_3 satisfying the commutation relations

$$[I_1, I_2]_q \equiv q^{1/4} I_1 I_2 - q^{-1/4} I_2 I_1 = I_3,$$
$$[I_3, I_1]_q = I_2, \qquad [I_2, I_3]_q = I_1.$$

The first relation shows that $U_q(\mathrm{so}(3, \mathbf{C}))$ is generated by two elements I_1 and I_2 satisfying the relations

$$I_1^2 I_2 - (q^{1/2} + q^{-1/2}) I_1 I_2 I_1 + I_2 I_1^2 = -I_2,$$
$$I_1 I_2^2 - (q^{1/2} + q^{-1/2}) I_2 I_1 I_2 + I_2^2 I_1 = -I_1.$$

The q-deformed algebra $U_q(\mathrm{so}_{2,1})$ is obtained from $U_q(\mathrm{so}(3,\mathbf{C}))$ by introducing the involution determined uniquely by the formulas $I_1^* = -I_1$, $I_2^* = I_2$. The involution relations $I_1^* = -I_1$, $I_2^* = -I_2$ determine the q-deformed algebra $U_q(\mathrm{so}_3)$.

REMARK. The Lie groups SU(1, 1) and $\mathrm{SO}_0(2,1)$ are locally isomorphic and their Lie algebras are isomorphic. The q-deformed algebras $U_q(\mathrm{su}_{1,1})$ and $U_q(\mathrm{so}_{2,1})$ are not isomorphic. Moreover, they have distinct sets of unitary irreducible representations.

Let ε be a fixed complex number and let V_ε be the complex Hilbert space with orthonormal basis $\{|m\rangle, m = \varepsilon + n, n = 0, \pm 1, \pm 2, \ldots\}$. With every complex number σ we associate the representation $T_{\sigma\varepsilon}$ of the algebra $U_q(\mathrm{so}_{2,1})$ on the Hilbert space V_ε acting by the formulas [11]

$$T_{\sigma\varepsilon}(I_1)|m\rangle = \mathrm{i}[m]|m\rangle,$$
$$T_{\sigma\varepsilon}(I_2)|m\rangle = d(m)[a-m]|m+1\rangle + d(m-1)[-a-m]|m-1\rangle,$$

where $[b]$ is a q-number and

$$d(m) = \big([m][m+1]\big)^{1/2}\big([2m][2m+2]\big)^{-1/2}. \tag{46}$$

If $q \neq 1$ then the operator $T_{\sigma\varepsilon}(I_2)$ is bounded.

In the set of representations $T_{\sigma\varepsilon}$ there exist equivalence relations. First of all, one can see that the matrices of the representations $T_{\sigma\varepsilon}$ and $T_{\sigma,\varepsilon+k}$, $k \in \mathbf{Z}$, coincide. For this reason we assume below that

$$0 \leqslant \operatorname{Re}\varepsilon < 1.$$

We also have the equivalence relations (10)–(12).

THEOREM. *If q is not a root of unity, then the representation $T_{\sigma\varepsilon}$ is irreducible if and only if $a \not\equiv \pm\varepsilon \pmod{\mathbf{Z}}$ or if $\varepsilon = \pm\mathrm{i}\pi + 1$ or if (a, ε) does not coincide with one of the couples $(\pm\mathrm{i}\pi/h, \pm\mathrm{i}\pi/h + 1)$, where $q = \mathrm{e}^h$.*

Reducible representations $T_{\sigma\varepsilon}$ give irreducible components which are irreducible representations of $U_q(\mathrm{so}_{2,1})$ with highest or lowest weights.

The following irreducible representations $T_{\sigma\varepsilon}$ of $U_q(\mathrm{so}_{2,1})$, $q = \exp h$, $h \in \mathbf{R}$, are unitary:

(a) the representations $T_{\sigma\varepsilon}$ with $\sigma = \mathrm{i}\rho - 1/2$, $2\pi/h > \rho \geqslant 0$, $\varepsilon = c + 2\mathrm{i}k\pi/|h|$, $0 \leqslant c < 1$, $k \in \mathbf{Z}$ (the principal unitary series);

(b) the representations $T_{\sigma\varepsilon}$, $0 \leqslant \operatorname{Re}\varepsilon < 1$, $\operatorname{Im}\varepsilon = 2k\pi/h$, $k \in \mathbf{Z}$, $\operatorname{Re}(\varepsilon - 1) > \sigma > -\operatorname{Re}\varepsilon$ for $\operatorname{Re}\varepsilon > 1/2$ and $-\operatorname{Re}\varepsilon > \sigma > \operatorname{Re}\varepsilon - 1$ for $\operatorname{Re}\varepsilon < 1/2$ (the supplementary series);

(c) the representations $T_{\sigma\varepsilon}$, $0 \leqslant \operatorname{Re}\varepsilon < 1$, $\operatorname{Im}\varepsilon = 2k\pi/h$, $k \in \mathbf{Z}$, $\operatorname{Im}\sigma = \pi/h$, $\operatorname{Re}\sigma > -1/2$ (the strange series);

(d) the zero representation.

The representations T_σ^+, $\sigma > 1/2$, and T_σ^-, $\sigma < -1/2$, with lowest (highest) weights are also unitary.

If $q = \exp \mathrm{i}h$, $h \in \mathbf{R}$, then the following irreducible representations are unitary:

(a) the representations $T_{\sigma\varepsilon}$, $0 \leqslant \varepsilon < 1$, $\sigma = \mathrm{i}\rho - 1/2$, $\rho \in \mathbf{R}$ (the principal unitary series);

(b) the representations $T_{\sigma\varepsilon}$, $0 \leqslant \varepsilon < 1$, $\mathrm{Re}\,\sigma = \pi/h$ (the strange series);

(c) the zero representation.

The irreducible representations T_l of $U_q(\mathrm{so}_3)$ are given by a nonnegative integer or half-integer l and act on the linear spaces with the orthonormal bases $\{|m\rangle, m = -l, -l+1, \ldots, l\}$. The operators $T_l(I_1)$ and $T_l(I_2)$ are given by the formulas [11]

$$T_l(I_1)|m\rangle = \mathrm{i}[m]|m\rangle, \tag{47}$$

$$\begin{aligned} T_l(I_2)|m\rangle &= d(m)\big([l-m][l+m+1]\big)^{1/2}|m+1\rangle \\ &\quad - d(m-1)\big([l-m+1][l+m]\big)^{1/2}|m-1\rangle, \end{aligned} \tag{48}$$

where $d(m)$ is defined by formula (46).

6.2. *The quantum algebras* $U_q(\mathrm{so}(n, \mathbf{C}))$ *and their real forms*

It is well known that in the universal enveloping algebra $U(\mathrm{so}(n, \mathbf{C}))$ of the classical complex Lie algebra $\mathrm{so}(n, \mathbf{C})$ the following trilinear relations are valid ($i = 2, 3, \ldots, n$):

$$I_{i+1,i}I_{i,i-1}^2 - 2I_{i,i-1}I_{i+1,i}I_{i,i-1} + I_{i,i-1}^2 I_{i+1,i} = -I_{i+1,i}, \tag{49}$$

$$I_{i,i-1}I_{i+1,i}^2 - 2I_{i+1,i}I_{i,i-1}I_{i+1,i} + I_{i+1,i}^2 I_{i,i-1} = -I_{i,i-1}, \tag{50}$$

$$[I_{i,i-1}, I_{j,j-1}] = 0, \quad |i-j| > 1 \tag{51}$$

(they follow from the commutation relations for the generators I_{ij} of the Lie algebra $\mathrm{so}(n, \mathbf{C})$).

In our approach to the q-deformation of orthogonal algebras we define q-deformed associative algebras $U_q(\mathrm{so}(n, \mathbf{C}))$ by deforming relations (49)–(51) for generators $I_{i,i-1}$, $i = 2, 3, \ldots, n$. The q-deformed relations are of the form

$$I_{i+1,i}I_{i,i-1}^2 - \big(q^{1/2} + q^{-1/2}\big)I_{i,i-1}I_{i+1,i}I_{i,i-1} + I_{i,i-1}^2 I_{i+1,i} = -I_{i+1,i}, \tag{52}$$

$$I_{i,i-1}I_{i+1,i}^2 - \big(q^{1/2} + q^{-1/2}\big)I_{i+1,i}I_{i,i-1}I_{i+1,i} + I_{i+1,i}^2 I_{i,i-1} = -I_{i,i-1}, \tag{53}$$

$$[I_{i,i-1}, I_{j,j-1}] = 0, \quad |i-j| > 1, \tag{54}$$

where $[.,.]$ denotes the usual commutator. Obviously, in the limit $q \to 1$ formulas (52)–(54) give relations (49)–(51). Note also that relations (52) and (53) differ completely from the q-deformed Serre relations in the approach of Jimbo [13] and Drinfeld [8] to quantum orthogonal algebras by presence of a nonzero right hand side and by the possibility of the reduction

$$U_q\big(\mathrm{so}(n, \mathbf{C})\big) \supset U_q\big(\mathrm{so}(n-1, \mathbf{C})\big).$$

Recall that in the standard Jimbo–Drinfeld approach the quantum algebras $U_q(\mathrm{so}(2m, \mathbf{C}))$ and the quantum algebras $U_q(\mathrm{so}(2m+1, \mathbf{C}))$ are separate series of quantum algebras of the types D_m and B_m respectively which are constructed independently from each other. In this chapter, by the algebra $U_q(\mathrm{so}(n, \mathbf{C}))$ we mean the q-deformed algebra determined by the defining relations (52)–(54).

The compact real form $U_q(\mathrm{so}_n)$ of $U_q(\mathrm{so}(n, \mathbf{C}))$ is defined by the involution $I_{i,i-1}^* = -I_{i,i-1}$, $i = 2, 3, \ldots, n$. The noncompact real forms $U_q(\mathrm{so}_{p,r})$, $r = n - p$, are distinguished respectively by the formulas $I_{i,i-1}^* = -I_{i,i-1}$, $i \neq p+1$, $I_{p+1,p}^* = I_{p+1,p}$. In particular, for $r = 1$ we obtain the q-analogue of the Lorentz algebras.

6.3. *Representations of the algebra $U_q(\mathrm{so}_{3,1})$*

The irreducible representations of $U_q(\mathrm{so}_{3,1})$ are given in the same way as in the case of the Lorentz group $\mathrm{SO}_0(3,1)$. These representations $T_{\sigma s}$ are defined by a complex number σ and by an integral or half-integral number s. To give these representations it is sufficient to have the operators $T_{\sigma s}(I_{i,i-1})$, $i = 2, 3, 4$. The representation $T_{\sigma s}$ acts on the Hilbert space V_s with orthonormal basis $|l, m\rangle$, $l = |s|, |s|+1, |s|+2, \ldots$, $m = -l, -l+1, \ldots, l$. On the subspace V_{ls} spanned by the basis elements $|l, m\rangle$ with fixed l the irreducible representation T_l of the subalgebra $U_q(\mathrm{so}_3)$ acts as given by formulas (47) and (48). The operator $T_{\sigma s}(I_4)$, $I_4 \equiv I_{43}$, is of the form [11]

$$T_{\sigma s}(I_4)|l, m\rangle = [\sigma + l + 1]A_{l+1}|l+1, m\rangle - [\sigma - l]A_l|l-1, m\rangle + \mathrm{i}C_l|l, m\rangle,$$

where

$$A_l = \left(\frac{[l+s][l-s][l+m][l-m]}{[l]^2[2l-1][2l+1]}\right)^{1/2}, \qquad C_l = \frac{[\sigma][s][m]}{[l][l+1]}.$$

THEOREM. *If q is not a root of unity, then the representation $T_{\sigma s}$ of the algebra $U_q(\mathrm{so}_{3,1})$ is irreducible if and only if 2σ is not an integer of the same evenness as $2s$ or if $|\sigma| \leqslant |s|$.*

The irreducible representations $T_{\sigma s}$ and $T_{-\sigma,-s}$ of the algebra $U_q(\mathrm{so}_{3,1})$ are equivalent. There also exist equivalence relations connected with the periodicity of the function $w(z) = [z]$.

The following irreducible representations $T_{\sigma s}$ of the algebra $U_q(\mathrm{so}_{3,1})$, $q = \exp h$, $h \in \mathbf{R}$, are unitary:

(a) the representations $T_{\sigma s}$ with nonnegative integral or half-integral s and with $\sigma = \mathrm{i}\rho$, where $0 \leqslant \rho < 4\pi/|h|$ for $s > 0$ and $0 \leqslant \rho < 2\pi/|h|$ for $s = 0$ (the principal unitary series);

(b) the representations $T_{\sigma s}$ where $s = 0$ and $0 < \sigma < 1$ (the supplementary series);

(c) the representations $T_{\sigma s}$ where $\mathrm{Im}\,\sigma = \pi/h$, $\mathrm{Re}\,\sigma > 0$ (the strange series);

(d) the zero representation.

If $q = \exp \mathrm{i}h$, $h \in \mathbf{R}$, then we have the following unitary representations:

(a) the representations $T_{\sigma s}$, $\sigma = \mathrm{i}\rho$, $0 < \rho < \infty$ (the principal unitary series);

(b) the representations $T_{\sigma s}$ with $\mathrm{Re}\,\sigma = \pi/|h|$ (the strange series);
(c) the zero representation.

6.4. *Representations of the algebra $U_q(\mathrm{so}_{n,1})$*

The irreducible finite-dimensional representations T_{m_n} of the algebra $U_q(\mathrm{so}_n)$, which are of class 1 with respect to the subalgebra $U_q(\mathrm{so}_{n-1})$, are given by a nonnegative integer m_n. As in the classical case, we choose Gel'fand–Tsetlin bases in the carrier spaces of these representations. They correspond to successive restrictions of a representation onto the subalgebras

$$U_q(\mathrm{so}_n) \supset U_q(\mathrm{so}_{n-1}) \supset U_q(\mathrm{so}_{n-2}) \supset \cdots \supset U_q(\mathrm{so}_2).$$

The basis elements of these spaces are denoted by $|m_n, m_{n-1}, \ldots, m_3, m_2\rangle$, where $m_{n-1}, \ldots, m_3, m_2$ are integers such that

$$m_n \geqslant m_{n-1} \geqslant \cdots \geqslant m_3 \geqslant m_2 \geqslant 0.$$

With respect to this basis the operator $T_{m_n}(I_{n,n-1})$ of the representation T_{m_n} of $U_q(\mathrm{so}_n)$ is given by formula [11]:

$$\begin{aligned} &T_{m_n}(I_{n,n-1})|m_n, m_{n-1}, m_{n-2}, \ldots, m_2\rangle \\ &\quad = \big([m_n + m_{n-1} + n - 2][m_n - m_{n-1}]\big)^{1/2} R(m_{n-1}) \\ &\qquad \times |m_n, m_{n-1} + 1, m_{n-2}, \ldots, m_2\rangle \\ &\qquad - \big([m_n + m_{n-1} + n - 3][m_n - m_{n-1} + 1]\big)^{1/2} \\ &\qquad \times R(m_{n-1} - 1)|m_n, m_{n-1} - 1, m_{n-2}, \ldots, m_2\rangle, \end{aligned} \tag{55}$$

where $n \geqslant 3$ and

$$R(m_{n-1}) = \left(\frac{[m_{n-1} + m_{n-2} + n - 3][m_{n-1} - m_{n-2} + 1]}{[2m_{n-1} + n - 1][2m_{n-1} + n - 3]}\right)^{1/2}. \tag{56}$$

As it is easily seen, the relations just presented reduce for $n = 3$ to the action formula (48) for $T(I_{32})$.

The representations T_σ of the algebra $U_q(\mathrm{so}_{n,1})$, which are of class 1 with respect to the subalgebra $U_q(\mathrm{so}_n)$, are given by a complex number σ. They act on the Hilbert space V with orthonormal basis

$$|m_n, m_{n-1}, \ldots, m_3, m_2\rangle, \quad m_n \geqslant m_{n-1} \geqslant \cdots \geqslant m_3 \geqslant m_2 \geqslant 0.$$

The representation operators $T_\sigma(I_{i,i-1})$, $i \leqslant n$, corresponding to the subalgebra $U_q(\mathrm{so}_n)$, act upon the basis elements according to formula (55). For the operator $T_\sigma(I_{n+1,n})$ we have

$$\begin{aligned} T_\sigma(I_{n+1,n})|m_n, m_{n-1}, \ldots\rangle = {} & [\sigma + m_n + n - 1]R(m_n)|m_n + 1, m_{n-1}, \ldots\rangle \\ & - [\sigma - m_n + 1]R(m_n - 1)|m_n - 1, m_{n-1}, \ldots\rangle, \end{aligned}$$

where $R(m_n)$ is given (up to the shift $n-1 \to n$) by formula (56). There are equivalence relations between these representations.

THEOREM. *If q is not a root of unity, then the representation T_σ of the algebra $U_q(\mathrm{so}_{n,1})$ is irreducible for all non-integral σ and for integral σ from the interval $-n+1<\sigma<0$.*

There are unitary representations in the set of representations T_σ. In particular, there are the principal unitary series, the supplementary series, and the strange series of unitary representations.

Some of the representations T_σ are reducible. There are unitary representations in the set of reducible components of these representations. For a list of them see [11].

References

[1] N.N. Anderson, P. Polo and K. Wen, *Representations of quantum algebras*, Invent. Math. **104** (1991), 1–59.

[2] D. Arnaudon and A. Chakrabarti, *q-Analogue of $IU(n)$ for q a root of unity*, Phys. Lett. B **255** (1991), 242–248.

[3] I.M. Burban and A.U. Klimyk, *Representations of the quantum algebra $U_q(\mathrm{su}_{1,1})$*, J. Phys. A **26** (1993), 2139–2151.

[4] A. Chakrabarti, *q-Analogs of $IU(n)$ and $U(n,1)$*, J. Math. Phys. **32** (1991), 1227–1234.

[5] J. Dixmier, *Algèbres Enveloppantes*, Gauthier-Villars, Paris (1974).

[6] V.K. Dobrev, *Representations of quantum groups*, Symmetries in Science V, B. Gruber et al., eds, Plenum Press, New York (1991), 93–136.

[7] H.D. Doebner, J.D. Hennig and W. Lücke, *Mathematical guide to quantum groups*, Lecture Notes in Phys., Vol. 370 (1990), 29–63.

[8] V.G. Drinfeld, *Hopf algebras and the Yang–Baxter equations*, Soviet Math. Dokl. **32** (1985), 254–258.

[9] V.G. Drinfeld, *Quantum groups*, Proc. Intern. Cong. Math., Berkeley, 1986, Amer. Math. Soc., Providence, RI (1987), 798–820.

[10] A.M. Gavrilik and A.U. Klimyk, *q-Deformed orthogonal and pseudo-orthogonal algebras and their representations*, Lett. Math. Phys. **21** (1991), 215–220.

[11] A.M. Gavrilik and A.U. Klimyk, *The representations of the q-deformed algebras $U_q(\mathrm{so}_{2,1})$ and $U_q(\mathrm{so}_{3,1})$*, J. Math. Phys. **35** (1994), 3670–3686.

[12] M. Hazewinkel, *Introductory recommendations for the study of Hopf algebras in mathematics and physics*, CWI Quarterly **4** (1991), 3–26.

[13] M. Jimbo, *A q-difference analogue of $U(g)$ and the Yang–Baxter equation*, Lett. Math. Phys. **10** (1985), 63–69.

[14] M. Jimbo, *Quantum R matrix related to the generalized Toda system: An algebraic approach*, Lecture Notes in Phys., Vol. 246 (1986), 335–360.

[15] I.I. Kachurik and A.U. Klimyk, *Spectra of representation operators of q-deformed algebras and q-orthogonal polynomials*, Preprint of Inst. for Theor. Phys., Kiev (1993).

[16] A.U. Klimyk, *Matrix Elements and Clebsch–Gordan Coefficients of Group Representations* (in Russian), Naukova Dumka, Kiev (1979).

[17] A.U. Klimyk and V.A. Groza, *Representations of the quantum pseudo-unitary algebras*, Preprint of Inst. for Theor. Phys., Kiev (1989).

[18] A.U. Klimyk and S.Z. Pakuliak, *Representations of the quantum algebras $U_q(\mathrm{u}_{r,s})$ and $U_q(\mathrm{u}_{r+s})$ related to the quantum hyperboloid and sphere*, J. Math. Phys. **33** (1992), 1937–1947.

[19] A. Knapp, *Representation Theory of Semisimple Groups. An Overview Based on Examples*, Princeton Univ. Press, Princeton (1986).

[20] G. Lusztig, *Quantum deformations of certain simple modules over enveloping algebras*, Adv. in Math. **70** (1988), 237–249.

[21] G. Lusztig, *Modular representations and quantum groups*, Contemp. Math. **82** (1989), 59–77.

[22] G. Lusztig, *Quantum groups at root of* 1, Geom. Dedicata **35** (1990), 89–113.

[23] G. Lusztig, *Canonical bases arising from quantized enveloping algebras*, J. Amer. Math. Soc. **4** (1991), 447–498.

[24] T. Masuda, K. Mimachi, Y. Nakagami, M. Noumi, Y. Saburi and K. Ueno, *Unitary representations of the quantum group* $\mathrm{SU}_q(1,1)$: *Structure of the dual space of* $U_q(\mathrm{sl}(2))$, Lett. Math. Phys. **19** (1990), 187–194.

[25] T. Masuda, K. Mimachi, Y. Nakagami, M. Noumi, Y. Saburi and K. Ueno, *Unitary representations of the quantum group* $\mathrm{SU}_q(1,1)$: *II. Matrix elements of unitary representations and the basic hypergeometric functions*, Lett. Math. Phys. **19** (1990), 195–204.

[26] B. Parshall and J.-P. Wang, *Quantum linear groups*, Mem. Amer. Math. Soc. **439** (1991).

[27] N.Yu. Reshetikhin, L.A. Takhtadjan and L.D. Faddeev, *Quantization of Lie groups and Lie algebras*, Leningrad Math. J. **1** (1990), 193–225.

[28] M. Rosso, *Finite dimensional representations of the quantum analog of the enveloping algebra of a complex simple Lie algebra*, Comm. Math. Phys. **117** (1988), 581–593.

[29] K. Ueno, *Spectral analysis for the Casimir operator on the quantum group* $\mathrm{SU}_q(1,1)$, Proc. Japan Acad. Ser. A **66** (1990), 42–44.

[30] L.L. Vaksman and L.I. Korogodsky, *Harmonic analysis on quantum hyperboloids*, Preprint of Inst. for Theor. Phys., Kiev (1990).

[31] L.L. Vaksman and L.I. Korogodsky, *Spherical functions on the quantum group* $\mathrm{SU}(1,1)$ *and a* q*-analogue of the Mehler–Fock formula*, Funct. Anal. Appl. **25** (1991), 48–49.

[32] E.Ye. Vaysleb, *Representations of quantum *-algebras* $\mathrm{sl}_t(N+1,\mathbf{R})$, Preprint of Inst. of Math., Kiev (1991).

[33] N.Ya. Vilenkin and A.U. Klimyk, *Representation of Lie Groups and Special Functions*, Vol. 1, Kluwer Acad. Publ., Dordrecht (1991).

[34] N.Ya. Vilenkin and A.U. Klimyk, *Representation of Lie Groups and Special Functions*, Vol. 3, Kluwer Acad. Publ., Dordrecht (1992).

This Page Intentionally Left Blank

Section 6E
Abstract and Functorial Representation Theory

This Page Intentionally Left Blank

Burnside Rings

Serge Bouc

Université Paris 7-Denis Diderot, 2 place Jussieu, 75251 Paris Cedex 05, France
E-mail: bouc@math.jussieu.fr

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

1. Introduction

Let G be a finite group. The Burnside ring $B(G)$ of the group G is one of the fundamental representation rings of G, namely the ring of *permutation* representations.

It is in many ways the universal object to consider when looking at the category of G-sets. It can be viewed as an analogue of the ring $\mathbb{Z}$ of integers for this category.

It can be studied from different points of view. First $B(G)$ is a *commutative ring*, and one can look at is prime spectrum and primitive idempotents. This leads to various induction theorems (Artin, Conlon, Dress, Brauer): the typical statement here is that any (virtual) RG-module is a linear combination with suitable coefficients of modules induced from certain subgroups of G (cyclic, hypoelementary, or Dress subgroups).

The Burnside ring is the natural framework to study the *invariants* attached to structured G-sets (such as G-posets, or more generally simplicial G-sets). Those invariants are generalizations for the category of G-sets of classical notions, such as the Möbius function of a poset, or the Steinberg module of a Chevalley group. They have properties of projectivity, which lead to congruences on the values of Euler–Poincaré characteristic of some sets of subgroups of G.

The ring $B(G)$ is also functorial with respect to G and subgroups of G, and this leads to the *Mackey functor* or *Green functor* point of view. There are close connections between the Burnside ring and the Mackey algebra. The Burnside Mackey functor is a typical example of projective Mackey functor. It is also a universal object in the category of Green functors. This leads to a decomposition of the category of Mackey functors for G as a direct sum of smaller Abelian categories.

Finally $B(G)$ is also functorial with respect to bisets, and this is leads to the definition of *double Burnside rings*. Those rings are connected to stable homotopy theory via the Segal conjecture, and they provide tools to study the stable splittings of the classifying spaces of finite groups.

2. Basic properties of G-sets

2.1. *Notation*

Let G be a finite group. The category of finite G-sets will be denoted by G-**set**. The objects of G-**set** are the finite sets with a left G-action, morphisms are G-equivariant maps, and composition of morphisms is the composition of maps.

If H is a subgroup of G, and g is an element of G, the notation ${}^{g}H$ stands for $g.H.g^{-1}$, and the notation H^g for $g^{-1}.H.g$. The normalizer of H in G is denoted by $N_G(H)$. The set of classes gH, for $g \in G$, is denoted by G/H. It is a G-set by left multiplication. A set of representatives in G of G/H is denoted by $[G/H]$.

If X is a G-set, the stabilizer in G of an element x of X is denoted by G_x. The set of orbits of a subgroup H of G on X is denoted by $H\backslash X$, and $[H\backslash X]$ denotes a set of representatives in X of $H\backslash X$.

The set of conjugacy classes of subgroups of G is denoted by s_G, and a set of representatives of s_G is denoted by $[s_G]$. If H and K are subgroups of G, then $H\backslash G/K$ is the set of double cosets HxK, for $x \in G$, and $[H\backslash G/K]$ denotes a set of representatives in G of those double cosets. The notation $H =_G K$ (resp. $H \subseteq_G K$) means that there is an element $x \in G$ such that $H^x = K$ (resp. $H^x \subseteq K$).

The cardinality of a set S is denoted by $|S|$.

If p is a prime number, the smallest normal subgroup N of G such that G/N is a p-group is denoted by $O^p(G)$. It is the subgroup of G generated by the p'-elements, i.e. the elements of order coprime to p.

More generally, if π is a set of primes, the notation $O^\pi(G)$ stands for the smallest normal subgroup N of G such that G/N is a *solvable* π-group. The group G is called π-perfect if $O^\pi(G) = G$. If G itself is a π-group, then the group $O^\pi(G)$ is the limit of the derived series of G. In particular, if π is the set of all primes, then a group is π-perfect if and only if it is *perfect*, i.e. equal to its derived subgroup.

The trivial group will be denoted by $\mathbb{1}$.

2.2. *Operations on G-sets*

When X is a G set, and H is a subgroup of G, one can view X as an H-set by restriction of the action. This H-set is denoted by $\mathrm{Res}_H^G X$. If $f: X \to Y$ is a morphism of G-sets, let $\mathrm{Res}_H^G f$ denote the map f viewed as a morphism of H-sets. This defines a *restriction functor* $\mathrm{Res}_H^G : G\textbf{-set} \to H\textbf{-set}$.

Now if Z is an H-set, the induced G-set $\mathrm{Ind}_H^G Z$ is defined as $G \times_H X$, i.e. the quotient of the Cartesian product $G \times X$ by the right action of H given by $(g,x).h = (gh, h^{-1}x)$ for $g \in G$, $h \in H$, $x \in X$. The left action of G on $G \times_H X$ is induced by its left action on $G \times X$ given by $g'.(g,x) = (g'g, x)$, for $g', g \in G$ and $x \in X$. If $f: Z \to T$ is a morphism of H-sets, then $\mathrm{Ind}_H^G f$ is the morphism of G-sets from $\mathrm{Ind}_H^G Z$ to $\mathrm{Ind}_H^G T$ defined by $(\mathrm{Ind}_H^G f)((g,x)) = (g, f(x))$. This defines an *induction functor* $\mathrm{Ind}_H^G : H\textbf{-set} \to G\textbf{-set}$.

Note that if Z is isomorphic to H/K for some subgroup K of H, then $\mathrm{Ind}_H^G Z$ is isomorphic to G/K.

The set of fixed points of H on X is denoted by X^H. It is viewed as a $N_G(H)/H$-set. If $f: X \to Y$ is a morphism of G-sets, then the restriction of f to X^H is denoted by f^H. It is a morphism of $N_G(H)/H$-sets from X^H to Y^H, and this defines a *fixed points functor* from G**-set** to $N_G(H)/H$**-set**.

When H is a normal subgroup of G, any G/H-set can be viewed as a G-set by inflation, and any morphism of G/H-sets can be viewed as a morphism of G-sets. This operation defines an *inflation functor* from G/H-set to G-set.

Under the same conditions, if X is a G-set, then the set of orbits $H\backslash X$ of H on X can be viewed as a G/H-set, and any morphism of G-sets induces a morphism of G/H-sets between the corresponding sets of orbits. This defines a *deflation functor* $\mathrm{Def}_{G/H}^G : G\textbf{-set} \to G/H\textbf{-set}$.

Finally, let H be a subgroup of G, and x be an element of G. If Z is an H-set, then the group ${}^xH$ acts on Z by $h.z = h^x z$, where $h \in {}^xH$ and $z \in Z$, and $h^x z$ is computed in the H-set Z. This gives an ${}^xH$-set denoted by ${}^xZ$ or $c_{x,H}(Z)$. If $f: Z \to T$ is a morphism of H-sets, then $c_{x,H}(f)$ is the map f, viewed as a morphism of ${}^xH$-sets from ${}^xZ$ to ${}^xT$. This defines a *conjugation* functor $c_{x,H} : H\textbf{-set} \to {}^xH\textbf{-set}$.

Those functors have obvious properties of transitivity. They are moreover related by various identities. Among them:

PROPOSITION 2.2.1. *Let G be a finite group, and H and K be subgroups of G.*

(1) (Mackey formula) *If Z is an H-set, then there is an isomorphism of K-sets*

$$\mathrm{Res}_K^G \mathrm{Ind}_H^G Z \simeq \bigsqcup_{x \in [K\backslash G/H]} \mathrm{Ind}_{K \cap {}^x H}^K \, {}^x\mathrm{Res}_{K^x \cap H}^H Z.$$

(2) (Frobenius identity) *If X is a G-set and Z is an H-set, then there is an isomorphism of G-sets*

$$X \times \mathrm{Ind}_H^G Z \simeq \mathrm{Ind}_H^G \big((\mathrm{Res}_H^G X).Z \big)$$

and in particular for any G-set X, there is an isomorphism of G-sets

$$X \times (G/H) \simeq \mathrm{Ind}_H^G \mathrm{Res}_H^G X.$$

(3) *If Z is an H-set, then there is an isomorphism of $N_G(K)/K$-sets*

$$\big(\mathrm{Ind}_H^G Z\big)^K \simeq \bigsqcup_{\substack{x \in [N_G(K)\backslash G/H] \\ K^x \subseteq H}} \mathrm{Ind}_{N_{^x H}(K)/K}^{N_G(K)/K} \big({}^x Z\big)^K.$$

PROOF (sketch). (1) Let $S = [K\backslash G/H]$ denote the chosen set of representatives of double cosets $K\backslash G/H$, and consider the map

$$\mathrm{Res}_K^G \mathrm{Ind}_H^G Z \to \bigsqcup_{x \in S} \mathrm{Ind}_{K \cap {}^x H}^K \, {}^x\mathrm{Res}_{K^x \cap H}^H Z$$

sending the element (g, z) of the left hand side, with $g \in G$ and $z \in Z$, to the element (k, hz) of the component $x \in S$ of the right hand side, if g can be written $g = kxh$, for some $k \in K$ and $h \in H$. This is the required isomorphism of K-sets.

(2) Consider the map

$$X \times \mathrm{Ind}_H^G Z \to \mathrm{Ind}_H^G \big((\mathrm{Res}_H^G X) \times Z \big)$$

sending the element $(x, (g, z))$ of the left hand side, with $x \in X$, $g \in G$, and $z \in Z$, to the element $(g, (g^{-1}x, z))$ of the right hand side. This is the required isomorphism of G-sets. The other isomorphism in assertion (2) is the special case $Z = H/H$.

(3) Note that $(\mathrm{Ind}_H^G Z)^K = (\mathrm{Res}_{N_G(K)}^G \mathrm{Ind}_H^G Z)^K$. Using the Mackey formula, it is enough to consider the case $K \trianglelefteq G$.

In this case, the element (g, z) of $\mathrm{Ind}_H^G Z$ is invariant by K if and only if $K^g = K \subseteq H$ and $x \in X^K$. Thus $(\mathrm{Ind}_H^G Z)^K$ is empty if $K \nsubseteq H$, and it it equal to $\mathrm{Ind}_{H/K}^{G/K} X^K$ otherwise. □

2.3. *Characterization of G-sets*

LEMMA 2.3.1. *Let G be a finite group.*

(1) *Any G-set is a disjoint union of transitive ones. If X is a transitive G-set, and if $x \in X$, then the map*

$$gG_x \in G/G_x \mapsto g.x \in X$$

is an isomorphism of G-sets.

(2) *If H and K are subgroups of G, then the map $f \mapsto f(H)$ is a one to one correspondence between the set of G-set homomorphisms from G/H to G/K and the set of cosets $xK \in G/K$ such that $H \subseteq {}^xK$. In particular, the G-sets G/H and G/K are isomorphic if and only if H and K are conjugate in G.*

PROOF. Both assertions are obvious. □

One can characterize a G-set up to isomorphism using the following fundamental theorem of Burnside ([15, Ch. XII, Theorem I]):

THEOREM 2.3.2 (Burnside). *Let G be a finite group, and X and Y be finite G-sets. Then the following are equivalent:*

(1) *The G-sets X and Y are isomorphic.*

(2) *For any subgroup H of G, the sets X^H and Y^H have the same cardinality.*

PROOF. It is clear that (1) implies (2), since any G-set isomorphism $X \to Y$ induces a bijection $X^H \to Y^H$ on the sets of fixed points by any subgroup H of G.

To show the converse, observe that it follows from Lemma 2.3.1 that any finite G-set X can be written up to isomorphism as

$$X = \bigsqcup_{K \in [s_G]} a_K(X) G/K \tag{2.3.3}$$

for some $a_K(X) \in \mathbb{N}$, where $a_K(X)G/K$ denotes the disjoint union of $a_K(X)$ copies of G/K.

Now if (2) holds, for any $H \in [s_G]$, there is an equation

$$\sum_{K \in [s_G]} \big(a_K(X) - a_K(Y)\big)\big|(G/K)^H\big| = 0.$$

The matrix m of this system of equations is given by

$$m(H, K) = \big|(G/K)^H\big| = \big|\{x \in G/K \mid H^x \subseteq K\}\big|$$

for $K, H \in [s_G]$. In particular the entry $m(H, K)$ is non-zero if and only if some conjugate of H is contained in K.

If the set $[s_G]$ is given a total ordering $\preceq$ such that $H \preceq K$ implies $|H| \leqslant |K|$, then the matrix m is upper triangular, with non-zero diagonal coefficient $m(H, H) = |N_G(H) : H|$. In particular m is non-singular, and it follows that $a_K(X) = a_K(Y)$, for any $K \in [s_G]$, and the G-sets X and Y are isomorphic. It also follows that the integers $a_K(X)$ in decomposition (2.3.3) are uniquely determined by X. □

DEFINITION 2.3.4. The above matrix m (or sometimes its transpose) is called the *table of marks* of the group G.

3. The ring structure

3.1. *Definition*

The following definition of the Burnside ring of the group G is an axiomatic generalization of Burnside's ideas and techniques. It appears in an article of Solomon ([43]):

DEFINITION 3.1.1 (Solomon). The *Burnside ring* $B(G)$ of G is the Grothendieck group of the category G-**set**, for the relations given by decomposition in disjoint union of G-sets. The multiplication on $B(G)$ is induced by the direct product of G-sets.

This means that $B(G)$ is the free $\mathbb{Z}$-module with basis the set of equivalence classes of finite G-sets, quotiented by relations identifying the class of the disjoint union $X \sqcup Y$ of two G-sets X and Y to the sum of the class of X and the class of Y.

The direct product of G-sets is commutative and distributive with respect to disjoint union, up to canonical isomorphisms. Hence it induces by bilinearity a commutative ring structure on $B(G)$. The class of the set • of cardinality 1 is a unit for this ring structure.

Two finite G-sets A and B have the same image in $B(G)$ if and only if there are sequences of finite G-sets X_i, Y_i, for $1 \leqslant i \leqslant m$, and Z_j, T_j, for $1 \leqslant j \leqslant n$ and an isomorphism of G-sets

$$\begin{aligned} A \sqcup \Big(\bigsqcup_{i=1}^{m} X_i\Big) \sqcup \Big(\bigsqcup_{i=1}^{m} Y_i\Big) \sqcup \Big(\bigsqcup_{j=1}^{n} (Z_j \sqcup T_j)\Big) \\ \simeq B \sqcup \Big(\bigsqcup_{i=1}^{m} (X_i \sqcup Y_i)\Big) \sqcup \Big(\bigsqcup_{j=1}^{n} Z_j\Big) \sqcup \Big(\bigsqcup_{j=1}^{n} T_j\Big). \end{aligned}$$

Taking fixed points of both sides shows that for any subgroup H of G, one has $|A^H| = |B^H|$, and Burnside's Theorem 2.3.2 now implies that A and B are isomorphic as G-sets. In the sequel, a G-set A and its image in $B(G)$ will be identified.

It follows from Burnside's Theorem 2.3.2 that any finite G-set X can be written uniquely up to isomorphism as

$$X \simeq \bigsqcup_{H \in [s_G]} a_H(X) G/H.$$

Hence $B(G)$ is a free $\mathbb{Z}$-module, with basis indexed by elements G/H, for $H \in [s_G]$. In this basis, the multiplication law can be recovered by

$$(G/H).(G/K) = \sum_{x \in [H\backslash G/K]} G/(H \cap {}^xK). \tag{3.1.2}$$

This follows from Proposition 2.2.1, since by the Frobenius identities

$$(G/H) \times (G/K) \simeq \mathrm{Ind}_H^G \mathrm{Res}_H^G \mathrm{Ind}_K^G K/K$$

and since by the Mackey formula

$$\mathrm{Res}_H^G \mathrm{Ind}_K^G K/K \simeq \bigsqcup_{x \in [H\backslash G/K]} H/(H \cap {}^xK).$$

Finally, the operations on G-sets defined in Section 2.2 all commute with disjoint unions. Hence they can be extended to the Burnside ring: the elements of $B(G)$ can be viewed as a formal differences $X - Y$ of two finite G-sets. If $F : G\textbf{-set} \to H\textbf{-set}$ denotes one of the functors of restriction, induction, fixed points, inflation, deflation, or conjugation, then F induces a group homomorphism, still denoted by F, from $B(G)$ to $B(H)$, defined by

$$F(X - Y) = F(X) - F(Y)$$

for any finite G-sets X and Y.

Thus for example, if H is a subgroup of G, there is a restriction homomorphism

$$\mathrm{Res}_H^G : B(G) \to B(H).$$

This homomorphism is actually a morphism of rings (with unit).

In the special case $H = \mathbb{1}$, since $B(H) \simeq \mathbb{Z}$, this gives an extension of the cardinality to a map $X \mapsto |X| = \mathrm{Res}_{\mathbb{1}}^G X$ from $B(G)$ to $\mathbb{Z}$.

Similarly, there is an induction homomorphism

$$\mathrm{Ind}_H^G : B(H) \to B(G).$$

This morphism is not a ring homomorphism in general.

If H is a subgroup of G, there is a fixed points homomorphism $X \mapsto X^H$ from $B(G)$ to $B(N_G(H)/H)$, which is actually a ring homomorphism. When H is a normal subgroup of G, there is an inflation homomorphism

$$\mathrm{Inf}_{G/H}^G : B(G/H) \to B(G)$$

which is a ring homomorphism.

Finally, if x is an element of G, there is a conjugation homomorphism $Z \mapsto {}^xZ$ from $B(H)$ to $B({}^xH)$, which is a ring isomorphism.

The following is an obvious extension of Proposition 2.2.1:

PROPOSITION 3.1.3. *Let G be a finite group, and H and K be subgroups of G.*

(1) (Mackey formula) *If $Z \in B(H)$, then in $B(K)$*

$$\mathrm{Res}_K^G \mathrm{Ind}_H^G Z = \sum_{x \in [K \backslash G/H]} \mathrm{Ind}_{K \cap {}^x H}^K \, {}^x \mathrm{Res}_{K^x \cap H}^H Z.$$

(2) (Frobenius identity) *If $X \in B(G)$ and $Z \in B(H)$, then in $B(G)$*

$$X.\mathrm{Ind}_H^G Z = \mathrm{Ind}_H^G \big((\mathrm{Res}_H^G X).Z \big)$$

and in particular for any $X \in B(G)$

$$X.(G/H) = \mathrm{Ind}_H^G \mathrm{Res}_H^G X.$$

(3) *If $Z \in B(H)$, then in $B(N_G(K)/K)$*

$$\big(\mathrm{Ind}_H^G Z\big)^K = \sum_{\substack{x \in [N_G(K) \backslash G/H] \\ K^x \subseteq H}} \mathrm{Ind}_{N_{{}^x H}(K)/K}^{N_G(K)/K} \big({}^x Z\big)^K.$$

3.2. *Fixed points as ring homomorphisms*

The ring $B(G)$ is finitely generated as a $\mathbb{Z}$-module, hence it is a Noetherian ring. Burnside's Theorem 2.3.2 can be interpreted as follows: each subgroup H of G defines a *ring* homomorphism $\phi_H^G : B(G) \to \mathbb{Z}$ by $\phi_H^G(X) = |X^H|$. The kernel of ϕ_H^G is a prime ideal, since $\mathbb{Z}$ is an integral domain, and the intersection of all those kernels for subgroups H of G is zero. In particular, the ring $B(G)$ is reduced.

Since $\phi_H^G = \phi_K^G$ if H and K are conjugate in G, it follows that the product map

$$\Phi = \prod_{H \in [s_G]} \phi_H^G : B(G) \to \prod_{H \in [s_G]} \mathbb{Z}$$

is injective. Moreover this map Φ is a map between free $\mathbb{Z}$-modules having the same rank. Hence the cokernel of Φ is finite.

The matrix m of Φ with respect to the basis $\{G/H\}_{H \in [s_G]}$ and to the canonical basis $\{u_H\}_{H \in [s_G]}$ of $\prod_{H \in [s_G]} \mathbb{Z}$ is the table of marks of the group G. Recall from Definition 2.3.4 that for $H, K \in [s_G]$

$$m(H, K) = \left|G/K^H\right| = \left|\left\{x \in G/K \mid H^x \subseteq K\right\}\right|.$$

The cardinality of the cokernel of Φ is the determinant of m, hence it is equal to

$$\left|\mathrm{Coker}(\Phi)\right| = \prod_{H \in [s_G]} \left|N_G(H) : H\right|.$$

This cokernel has been described by Dress ([23]):

THEOREM 3.2.1 (Dress). *Let G be a finite group. For H and K in $[s_G]$, set*

$$n(K,H) = \left|\left\{x \in N_G(K)/K \mid \langle x, K\rangle =_G H\right\}\right|.$$

Then the element $y = \sum_{H\in[s_G]} y_H u_H$ of $\prod_{H\in[s_G]} \mathbb{Z}$ is in the image of Φ if and only if for any $K \in [s_G]$

$$\sum_{H\in[s_G]} n(K,H)y_H \equiv 0\ \left(\left|N_G(K)/K\right|\right).$$

PROOF. First let X be a finite G-set, and let $y = \Phi(X)$. Then with the notation of the theorem, one has $y_H = |X^H|$ for all $H \in [s_G]$, thus for any $K \in [s_G]$

$$\begin{aligned}\sum_{H\in[s_G]} n(K,H)y_H &= \sum_{H\in[s_G]} \left|X^H\right| \left|\left\{x \in N_G(K)/K \mid \langle x, K\rangle =_G H\right\}\right| \\ &= \sum_{x\in N_G(K)/K} \left|X^{\langle x,K\rangle}\right| \\ &= \sum_{x\in N_G(K)/K} \left|(X^K)^x\right|.\end{aligned}$$

Now for any finite group L acting on a finite set Z one has

$$|L| \times |L\backslash Z| = |L| \sum_{z\in Z} \frac{|L_z|}{|L|} = \left|\left\{(l,z) \in L \times Z \mid l.z = z\right\}\right| = \sum_{l\in L} \left|Z^l\right|.$$

Applying this for $L = N_G(K)/K$ and $Z = X^K$ gives

$$\sum_{H\in[s_G]} n(K,H)y_H = \left|N_G(K)/K\right| \left|N_G(K)\backslash X^K\right| \equiv 0\ \left(\left|N_G(K)/K\right|\right).$$

By linearity, this proves the "only if" part of the theorem.

Applying this for $X = G/M$, for some $M \in [s_G]$, shows that there is a matrix t indexed by $[s_G] \times [s_G]$, with entries in $\mathbb{Z}$, such that for $K \in [s_G]$

$$\sum_{H\in[s_G]} n(K,H)m(H,M) = \left|N_G(K)/K\right| t(K,M).$$

Now the matrix n is upper triangular, and its diagonal coefficients are equal to 1. Thus the matrix t is upper triangular, and

$$t(K,K) = m(K,K)/\left|N_G(K)/K\right| = 1.$$

In particular t is invertible (over $\mathbb{Z}$).

Now suppose that the element $y = \sum_{H\in[s_G]} y_H u_H \in \prod_{H\in[s_G]} \mathbb{Z}$ satisfies all the congruences of the theorem. Since Coker(Φ) is finite, there exist rational numbers r_M, for $M \in [s_G]$, such that

$$y = \sum_{M\in[s_G]} r_M \Phi(G/M).$$

In other words for each $H \in [s_G]$

$$y_H = \sum_{M\in[s_G]} |G/M^H| r_M.$$

Thus for each $K \in [s_G]$

$$\begin{aligned}\sum_{H\in[s_G]} n(K,H) y_H &= \sum_{H\in[s_G]} \sum_{M\in[s_G]} n(K,H) m(H,M) r_M \\ &= |N_G(K):K| \sum_{M\in[s_G]} t(K,M) r_M.\end{aligned}$$

The left hand side is a multiple of $|N_G(K):K|$ by assumption, hence there exist integers z_K such that for all $K \in [s_G]$

$$\sum_{M\in[s_G]} t(K,M) r_M = z_K.$$

Since t is invertible, it follows that $r_M \in \mathbb{Z}$ for all M, and $y \in \mathrm{Im}(\Phi)$. □

3.3. *Idempotents*

The map Φ of the previous section is an injective map between free $\mathbb{Z}$-modules having the same rank. Hence tensoring with $\mathbb{Q}$ gives a $\mathbb{Q}$-algebra isomorphism

$$\mathbb{Q} \otimes_{\mathbb{Z}} \Phi : \mathbb{Q} \otimes_{\mathbb{Z}} B(G) \xrightarrow{\cong} \prod_{H\in[s_G]} \mathbb{Q}$$

and in particular the algebra $\mathbb{Q} \otimes_{\mathbb{Z}} B(G)$ is semi-simple. It will be denoted by $\mathbb{Q}B(G)$. The component $\mathbb{Q}\phi_H^G$ of $\mathbb{Q}\Phi$ will still be written $X \mapsto |X^H|$, so in general $|X^H|$ will be a rational number for $X \in \mathbb{Q}B(G)$. More generally, all the notations defined for $B(G)$ will be extended without change to $\mathbb{Q}B(G)$.

The inverse image by $\mathbb{Q}\Phi$ of the element indexed by H of the canonical $\mathbb{Q}$-basis of $\prod_{H\in[s_G]} \mathbb{Q}$ is denoted by e_H^G. If H' is conjugate to H in G, one also sets $e_{H'}^G = e_H^G$. With this notation for any subgroups H and K of G, one has

$$|(e_H^G)^K| = \begin{cases} 1 & \text{if } H =_G K, \\ 0 & \text{otherwise.} \end{cases}$$

The set of elements e_H^G, for $H \in [s_G]$, is the set of primitive idempotents of $\mathbb{Q}B(G)$. Those idempotents have been computed explicitly by Gluck ([26]), and later independently by Yoshida ([53]). This can be done using the following lemma:

LEMMA 3.3.1. *Let G be a finite group.*

(1) *Let H be a subgroup of G. Then for any $X \in \mathbb{Q}B(G)$*

$$X.e_H^G = |X^H| e_H^G.$$

Conversely, if $Y \in \mathbb{Q}B(G)$ is such that $X.Y = |X^H| Y$ for any $X \in \mathbb{Q}B(G)$, then $Y \in \mathbb{Q}e_H^G$.

(2) *Let H be a proper subgroup of G. Then $\mathrm{Res}_H^G e_G^G = 0$. Conversely, if $Y \in \mathbb{Q}B(G)$ is such that $\mathrm{Res}_H^G Y = 0$ for any proper subgroup H of G, then $Y \in \mathbb{Q}e_G^G$.*

(3) *Let H be a subgroup of G, and let $y \in G$. Then*

$$\mathrm{Res}_H^G e_H^G = e_H^H, \qquad e_H^G = \frac{1}{|N_G(H) : H|} \mathrm{Ind}_H^G e_H^H, \qquad {}^y e_H^H = e_{{}^yH}^{{}^yH}.$$

PROOF. (1) The set of elements e_H^G, for $H \in [s_G]$, is a $\mathbb{Q}$-basis of $\mathbb{Q}B(G)$, thus for any $X \in \mathbb{Q}B(G)$, there are rational numbers r_H, for $H \in [s_G]$, such that

$$X = \sum_{H \in [s_G]} r_H e_H^G.$$

Taking fixed points of both sides by a subgroup $K \in [s_G]$ shows that $r_K = |X^K|$. It follows that for all $K \in [s_G]$

$$X.e_K^G = |X^K| e_K^G.$$

Now let Y be an element of $\mathbb{Q}B(G)$ verifying $X.Y = |X^H| Y$ for any $X \in \mathbb{Q}B(G)$. Then in particular $e_K^G.Y = 0$ if $K \neq_G H$, thus $Y = |Y^H| e_H^G$ is a rational multiple of e_H^G.

(2) Let H be a proper subgroup of G. Then for any subgroup of K of H

$$|(\mathrm{Res}_H^G e_G^G)^K| = |(e_G^G)^K| = 0$$

since obviously $|(\mathrm{Res}_H^G X)^K| = |X^K|$ for any G-set X, hence for any X in $\mathbb{Q}B(G)$. It shows that the restriction of e_G^G to any proper subgroup of G is zero.

Conversely, if the restriction of an element Y to any proper subgroup H of G is zero, then in particular $|Y^H| = 0$ for such a subgroup, and $Y = |Y^G| e_G^G$.

(3) If K is a subgroup of H, then

$$|(\mathrm{Res}_H^G e_H^G)^K| = |(e_H^G)^K| = \begin{cases} 1 & \text{if } K = H, \\ 0 & \text{if } K \neq H. \end{cases}$$

This shows that $\mathrm{Res}_H^G e_H^G = e_H^H$.

Consider next the element $\mathrm{Ind}_H^G e_H^H$. If X is any element of $\mathbb{Q}B(G)$, then by Frobenius identity

$$X.\mathrm{Ind}_H^G e_H^H = \mathrm{Ind}_H^G\big((\mathrm{Res}_H^G X).e_H^H\big) = \mathrm{Ind}_H^G\big(|X^H|.e_H^H\big) = |X^H|\mathrm{Ind}_H^G e_H^H.$$

It follows that there is a rational number r_H^G such that

$$\mathrm{Ind}_H^G e_H^H = r_H^G e_H^G.$$

By the Mackey formula, the restriction of the left hand side to H is equal to

$$\mathrm{Res}_H^G \mathrm{Ind}_H^G e_H^H = \sum_{x \in [H\backslash G/H]} \mathrm{Ind}_{H\cap^x H}^H {}^x\mathrm{Res}_{H^x\cap H}^H e_H^H = \sum_{x\in N_G(H)/H} {}^x e_H^H$$

since the restriction of e_H^H to $H^x \cap H$ is zero if $H^x \neq H$. Moreover if $y \in G$, then for any subgroup K of ${}^yH$

$$|({}^y e_H^H)^K| = |(e_H^H)^{K^y}| = \begin{cases} 1 & \text{if } K = {}^yH, \\ 0 & \text{otherwise.} \end{cases}$$

Thus ${}^y e_H^H = e_{{}^yH}^{{}^yH}$, and finally

$$\mathrm{Res}_H^G \mathrm{Ind}_H^G e_H^H = |N_G(H) : H| e_H^H = r_H^G e_H^H.$$

Hence $r_H^G = |N_G(H) : H|$, and

$$e_H^G = \frac{1}{|N_G(H) : H|} \mathrm{Ind}_H^G e_H^H$$

as was to be shown. □

THEOREM 3.3.2 (Gluck). *Let G be a finite group. If H is a subgroup of G, then*

$$e_H^G = \frac{1}{|N_G(H)|} \sum_{K\subseteq H} |K|\mu(K, H) G/K,$$

where μ is the Möbius function of the poset of subgroups of G, and G/K is the element $1 \otimes G/K \in \mathbb{Q}B(G)$.

PROOF. One can write

$$e_H^H = \sum_{K\subset H} r(K, H) H/K,$$

where $r(K,H)$ is a rational number. Since ${}^{y}e_H^H = e_{{}^yH}^{{}^yH}$ for $y \in G$, one can suppose $r({}^yK, {}^yH) = r(K,H)$ for $y \in G$. Taking induction from H to G gives

$$e_H^G = \frac{1}{|N_G(H):H|} \sum_{K \subseteq H} r(K,H)G/K. \tag{3.3.3}$$

Now the sum of the elements e_H^G, for $H \in [s_G]$, is equal to $\bullet = G/G$. Summing over all subgroups H of G instead of $H \in [s_G]$ gives

$$G/G = \sum_{H \subseteq G} \frac{|N_G(H)|}{|G|} \frac{1}{|N_G(H):H|} \sum_{K \subseteq H} r(K,H)G/K.$$

The coefficient of G/K in the right hand side is equal to the coefficient of G/K', for any conjugate K' of K in G, and it is equal to

$$\sum_{K \subseteq H \subseteq G} \frac{|H|}{|G|} r(K,H).$$

This must be equal to zero if $K \neq G$, and equal to 1 if $K = G$. Setting $r'(K,H) = \frac{|H|}{|K|} r(K,H)$ for $K \subseteq H$, this gives

$$\sum_{K \subseteq H \subseteq G} r'(K,H) = \begin{cases} |G:K| = 1 & \text{if } K = G, \\ 0 & \text{otherwise.} \end{cases}$$

This shows that $r'(K,H)$ is equal to $\mu(K,H)$, where μ is the Möbius function of the poset of subgroups of G. Thus $r(K,H) = \frac{|K|}{|H|}\mu(K,H)$, and Eq. (3.3.3) gives

$$e_H^G = \frac{1}{|N_G(H)|} \sum_{K \subseteq H} |K|\mu(K,H)G/K$$

as was to be shown. □

The formulae for primitive idempotents in $\mathbb{Q}B(G)$ lead to a natural question: when are these idempotents actually in $B(G)$? More generally, if π is a set of prime numbers, let $\mathbb{Z}_{(\pi)}$ denote the subring of $\mathbb{Q}$ of irreducible fractions with denominator prime to all the elements of π. In other words $\mathbb{Z}_{(\pi)}$ is the localization of $\mathbb{Z}$ with respect to the set $\mathbb{Z} - \bigcup_{p \in \pi} p\mathbb{Z}$. One may ask when the idempotents of $\mathbb{Q}B(G)$ are actually in $\mathbb{Z}_{(\pi)}B(G)$. The answer is as follows:

THEOREM 3.3.4 (Dress). *Let G be a finite group, and π be a set of primes. Let $\mathcal{F}$ be a family of subgroups of G, closed under conjugation in G. Let $[\mathcal{F}]$ denote the set $\mathcal{F} \cap [s_G]$. Then the following conditions are equivalent:*

(1) *The idempotent $\sum_{H \in [\mathcal{F}]} e_H^G$ lies in $\mathbb{Z}_{(\pi)}B(G)$.*

(2) *Let H and K be any subgroups of G such that H is a normal subgroup of K and the quotient K/H is cyclic of prime order $p \in \pi$. Then $H \in \mathcal{F}$ if and only if $K \in \mathcal{F}$.*

PROOF. Suppose first that (1) holds, i.e. that the idempotent $e = \sum_{H\in[\mathcal{F}]} e_H^G$ lies in $\mathbb{Z}_{(\pi)}B(G)$. This is equivalent to the existence of an integer m coprime to all the elements of π, such that $me \in B(G)$. Now if H and K are subgroups of G such that $H \trianglelefteq K$ and K/H is cyclic of prime order p, then for any element X of $B(G)$, one has $|X^H| \equiv |X^K|\ (p)$. It follows that

$$m|e^H| \equiv m|e^K|\ (p).$$

But $m|e^H|$ is equal to m if $H \in \mathcal{F}$, and it is equal to zero otherwise. Thus if $p \in \pi$, it follows that assertion (2) holds, because $m \not\equiv 0\ (p)$.

Conversely, suppose that (2) holds. Let m be an integer such that $me \in B(G)$. By Dress's Theorem 3.2.1 this is equivalent to require that for each subgroup K of G

$$\sum_{H\in[s_G]} n(K,H)m|e^H| = m\sum_{H\in[\mathcal{F}]} n(K,H) \equiv 0\ \big(|N_G(K):K|\big).$$

This can also be written as

$$m\sum_{H\in[\mathcal{F}]} \big|\{x \in N_G(K)/K \mid \langle x,K\rangle =_G H\}\big| \equiv 0\ \big(|N_G(K):K|\big),$$

i.e.

$$m\big|\{x \in N_G(K)/K \mid \langle x,K\rangle \in \mathcal{F}\}\big| \equiv 0\ \big(|N_G(K):K|\big).$$

There are two cases: if $K \in \mathcal{F}$, then if (2) holds, the condition $\langle x,K\rangle \in \mathcal{F}$ is equivalent to require that the prime factors of the order of x in $N_G(K)/K$ are all in π. Let $(N_G(K)/K)_\pi$ denote the set of such elements $x \in N_G(K)/K$. The previous congruence is

$$m\big|\big(N_G(K)/K\big)_\pi\big| \equiv 0\ \big(|N_G(K):K|\big). \tag{3.3.5}$$

Now if $K \notin \mathcal{F}$, the condition $\langle x,K\rangle \in \mathcal{F}$ is equivalent to require that there is a prime factor of the order of x which is not in π. In other words $x \in N_G(K)/K - (N_G(K)/K)_\pi$, and the congruence to check is the same.

Thus assertion (1) holds if and only if there exists an integer m coprime to all the elements of π, such that congruence (3.3.5) holds for any subgroup K of G. Such an integer exist if and only if for any subgroup K of G, the cardinality of $(N_G(K)/K)_\pi$ is a multiple of the π-part $|N_G(K):K|_\pi$ of $|N_G(K):K|$, i.e. the biggest divisor of $|N_G(K):K|$ having all its prime factors in π.

But for any finite group L, the set L_π it the set of elements of L of order dividing $|L|_\pi$ (if $l \in L_\pi$, then the subgroup $\langle l\rangle$ generated by l is a direct product of cyclic p-groups $\langle l_p\rangle$, for $p \in \pi$, and each l_p has order dividing $|G|_p$). Thus $|L_\pi|$ is a multiple of $|L|_\pi$, by a theorem of Frobenius (see Corollaire 1 of Théorème 23 of [42]). Thus (1) holds. □

COROLLARY 3.3.6. *Let G be a finite group, and π be a set of primes. If J is a π-perfect subgroup of G, set*

$$f_J^G = \sum_{\substack{H \in [s_G] \\ O^\pi(H) =_G J}} e_H^G.$$

Then the set of elements f_J^G, for π-perfect elements J of $[s_G]$, is the set of primitive idempotents of $\mathbb{Z}_{(\pi)}B(G)$.

In particular, the set of primitive idempotents of $B(G)$ is in one to one correspondence with the set of conjugacy classes of perfect subgroups of G.

The group G is solvable if and only if G/G is a primitive idempotent of $B(G)$.

PROOF. Let J be a π-perfect subgroup of G, and set

$$\mathcal{F} = \{H \subseteq G \mid O^\pi(H) =_G J\}.$$

Then clearly the family $\mathcal{F}$ is closed under conjugation in G, and satisfies condition (2) of Theorem 3.3.4. Thus f_J^G lies in $\mathbb{Z}_{(\pi)}B(G)$.

To prove that it is primitive, it suffices to show that the family $\mathcal{F}$ has no proper non-empty subfamily satisfying condition (2) of Theorem 3.3.4. But if $\mathcal{F}'$ is such a non-empty subfamily of $\mathcal{F}$, and if $H \in \mathcal{F}'$, then $O^\pi(H) \in \mathcal{F}'$ since the composition factors of $H/O^\pi(H)$ are cyclic groups of prime order belonging to π. Thus $J \in \mathcal{F}'$ since $\mathcal{F}'$ is closed by conjugation. Now if $H' \in \mathcal{F}$ the group $O^\pi(H')$ is in $\mathcal{F}'$, and $H' \in \mathcal{F}'$ by the same argument. Thus $\mathcal{F}' = \mathcal{F}$, and the idempotent f_J^G is primitive.

The last assertion of the corollary is the case where π is the set of all prime numbers. □

3.4. *Prime spectrum*

The prime spectrum of $B(G)$ has been determined by Dress ([24]):

THEOREM 3.4.1 (Dress). *Let G be a finite group, let p denote a prime number or zero, and H be a subgroup of G. Let*

$$I_{H,p}(G) = \{X \in B(G) \mid |X^H| \equiv 0\ (p)\}.$$

Then:

(1) *The set $I_{H,p}(G)$ is a prime ideal of $B(G)$.*
(2) *If I is a prime ideal of $B(G)$, then there is a subgroup H of G and an integer p equal to zero or a prime number, such that $I = I_{H,p}(G)$.*
(3) *If H and K are subgroups of G, and if p and q are prime numbers or zero, then $I_{H,p}(G) \subseteq I_{K,q}(G)$ if and only if one of the following holds:*
    (a) *One has $p = q = 0$, and the subgroups H and K are conjugate in G. In this case moreover $I_{H,p} = I_{K,q}$.*

(b) *One has $p = 0$ and $q > 0$, and the groups $O^q(H)$ and $O^q(K)$ are conjugate in G. In this case moreover $I_{H,p} \neq I_{K,q}$.*

(c) *One has $p = q > 0$, and the subgroups $O^p(H)$ and $O^p(K)$ are conjugate in G. In this case moreover $I_{H,p} = I_{K,q}$.*

PROOF. (1) Clearly $I_{H,p}$ is the kernel of the ring homomorphism from $B(G)$ to $\mathbb{Z}/p\mathbb{Z}$ mapping X to the class of $|X^H|$. Assertion (1) follows, since $\mathbb{Z}/p\mathbb{Z}$ is an integral domain.

(2) Conversely, if I is a prime ideal of $B(G)$, then the ring $R = B(G)/I$ is an integral domain. Let $\pi : B(G) \to R$ be the canonical projection. Since $R \neq \{0\}$, there is a subgroup H of G minimal subject to the condition $\pi(G/H) \neq 0$. Taking the image of equation (3.1.2) by π and using the minimality of H gives

$$\pi(G/H)\pi(G/K) = \sum_{\substack{x \in [G/K] \\ H^x \subseteq K}} \pi(G/H)$$

(since moreover $HxK = xK$ if $H^x \subseteq K$). Since R is an integral domain, and since $\pi(G/H) \neq 0$, it follows that

$$\pi(G/K) = \left|\left\{x \in G/K \mid H^x \subseteq K\right\}\right| 1_R = \left|(G/K)^H\right| 1_R.$$

It follows by linearity that $\pi(X) = |X^H| 1_R$ for any $X \in B(G)$. Let p denote the characteristic of R. Then p is equal to zero or a prime number. Clearly the kernel of π, equal to I by definition, is also equal to $I_{H,p}(G)$.

(3) Suppose that H and K are conjugate in G. Then for any $X \in B(G)$, one has $|X^H| = |X^K|$, and in particular $I_{H,0}(G) = I_{K,0}(G)$.

Now if p is a prime number, and if X is a finite G-set, since

$$X^H = \left(X^{O^p(H)}\right)^{H/O^p(H)}$$

and since $H/O^p(H)$ is a p-group, it follows that

$$\left|X^H\right| \equiv \left|X^{O^p(H)}\right| (p)$$

for any $X \in B(G)$. In particular $I_{H,0} \subseteq I_{O^p(H),p}$, and $I_{H,p}(G) = I_{K,p}(G)$ if $O^p(H)$ and $O^p(K)$ are conjugate in G. Thus $I_{H,p}(G) = I_{O^p(H),p}(G)$ for any $H \subseteq G$.

Conversely, suppose that $I_{H,p}(G) \subseteq I_{K,q}(G)$. There is a surjective ring homomorphism from $B(G)/I_{H,p}(G)$ to $B(G)/I_{K,q}(G)$. Since p is the characteristic of the ring $B(G)/I_{H,p}(G)$, there are three possible cases:

(a) First $p = q = 0$. Then the inclusion $I_{H,0}(G) \subseteq I_{K,0}(G)$ means that if $X \in B(G)$ is such that $|X^H| = 0$, then $|X^K| = 0$. Now $X = |N_G(K)| e_K^G$ is in $B(G)$ by Theorem 3.3.2, and $|X^K| = |N_G(K)| \neq 0$. Thus $|X^H| \neq 0$, hence H is conjugate to K in G. Clearly in this case $I_{H,0}(G) = I_{K,0}(G)$.

(b) The next possible case is $p = 0$ and $q > 0$. In this case if $X \in B(G)$ is such that $|X^H| = 0$, then $|X^K| \equiv 0\ (q)$. Consider the idempotent

$$f^G_{O^q(K)} = \sum_{\substack{L \in [s_G] \\ O^q(L) =_G O^q(K)}} e^G_L.$$

By Corollary 3.3.6, there is an integer m coprime to q such that $X = mf^G_{O^q(K)}$ is in $B(G)$. Now $|X^K| = m \not\equiv 0\ (q)$. Thus $|X^H| \neq 0$, and it follows that $O^q(H) =_G O^q(K)$. The inclusion $I_{H,0}(G) \subseteq I_{K,q}(G)$ is proper since the respective quotient rings have characteristic 0 and q.

(c) The last case is $p = q > 0$. Since $I_{H,0}(G) \subseteq I_{H,p}(G)$, it follows that $I_{H,0}(G) \subseteq I_{K,p}(G)$. Hence $O^p(H) =_G O^p(K)$ by the discussion of the previous case. And in this case $I_{H,p}(G) = I_{K,p}(G)$. □

3.5. *Application to induction theorems*

NOTATION 3.5.1. Let G be a finite group. Denote by $\mathcal{C}(G)$ the set of cyclic subgroups of G. Let $R_{\mathbb{C}}(G)$ be the ring of complex characters of G, and set $\mathbb{Q}R_{\mathbb{C}}(G) = \mathbb{Q} \otimes_{\mathbb{Z}} R_{\mathbb{C}}(G)$.

THEOREM 3.5.2 (Artin). *Let G be a finite group. Then*

$$\mathbb{Q}R_{\mathbb{C}}(G) = \sum_{H \in \mathcal{C}(G)} \mathrm{Ind}^G_H \mathbb{Q}R_{\mathbb{C}}(H).$$

In other words, any complex character of G is a linear combination with rational coefficients of characters induced from cyclic subgroups of G.

PROOF. There is a natural homomorphism from the Burnside ring $B(G)$ of G to the ring $R_{\mathbb{C}}(G)$, which maps a G-set X to the associated $\mathbb{C}G$-module $\mathbb{C}X$. This extends to a map of vector spaces $\mathbb{Q}B(G) \to \mathbb{Q}R_{\mathbb{C}}(G)$. Now the value of the character of the permutation module $\mathbb{C}X$ at the element s of G is the trace of s on $\mathbb{C}X$, which is equal to the number of fixed points $|X^s|$ of s on X.

Let H be a subgroup of G. Since $|(e^G_H)^s|$ is equal to 0 if H is not conjugate in G to the subgroup generated by s, it follows that the image of e^G_H in $\mathbb{Q}R_{\mathbb{C}}(G)$ is zero unless H is cyclic. And if H is cyclic, the image of e^G_H is a linear combination with rational coefficients of permutations characters $\mathrm{Ind}^G_K 1$, for subgroups K of H. Taking the image in $\mathbb{Q}R_{\mathbb{C}}(G)$ of the decomposition

$$G/G = \sum_{H \in [s_G]} e^G_H$$

shows that the trivial character is a linear combination with rational coefficients of such characters $\mathrm{Ind}_K^G 1$, which are induced from cyclic subgroups K of G. Then if χ is any character of G

$$\chi \,\mathrm{Ind}_K^G 1 = \mathrm{Ind}_K^G (\mathrm{Res}_K^G \chi)$$

and the theorem follows. □

NOTATION 3.5.3. Let p be a prime number, and $\mathcal{O}$ be a complete local Noetherian commutative ring with maximal ideal $\mathfrak{p}$ and residue field k of characteristic p. If G is a finite group, an *$\mathcal{O}G$-lattice* is a finitely generated $\mathcal{O}$-free $\mathcal{O}G$-module. Also denote by $\mathcal{O}$ the trivial $\mathcal{O}G$-lattice of $\mathcal{O}$-rank 1.

The *Green ring* $A_{\mathcal{O}}(G)$ is the Grothendieck group of the category of $\mathcal{O}G$- lattices, for the relations given by direct sum decompositions.

As an additive group, it is the quotient of the free Abelian group with basis the set of isomorphism classes of $\mathcal{O}G$-lattices, by the subgroup generated by elements $[M \oplus N] - [M] - [N]$, for $\mathcal{O}G$-lattices M and N, where $[M]$ denotes the class of M. Since the Krull–Schmidt theorem holds for $\mathcal{O}G$-lattices, the group $A_{\mathcal{O}}(G)$ is free, with basis the set of indecomposable $\mathcal{O}G$-lattices. The ring structure on $A_{\mathcal{O}}(G)$ is induced by tensor product (over $\mathcal{O}$) of $\mathcal{O}G$-lattices.

There is a natural ring homomorphism $\pi_{\mathcal{O}}$ from $B(G)$ to $A_{\mathcal{O}}(G)$, mapping the (class of the) finite G-set X to the (class of the) permutation lattice $\mathcal{O}X$. It is natural to look at the kernel $I_{\mathcal{O}}(G)$ of this morphism.

Since $\mathbb{Q}$ is flat over $\mathbb{Z}$, the sequence of Abelian groups

$$0 \to \mathbb{Q}I_{\mathcal{O}}(G) \to \mathbb{Q}B(G) \xrightarrow{\mathbb{Q}\pi_{\mathcal{O}}} \mathbb{Q}A_{\mathcal{O}}(G)$$

is exact. Since $\mathbb{Q}I_{\mathcal{O}}(G)$ is an ideal of the semi-simple algebra $\mathbb{Q}B(G)$, there exists a set $\mathcal{N}_{\mathcal{O}}(G)$ of subgroups of G, closed under conjugation, such that

$$\mathbb{Q}I_{\mathcal{O}}(G) = \sum_{H \in \mathcal{N}_{\mathcal{O}}(G) \cap [s_G]} \mathbb{Q}e_H^G.$$

Equivalently

$$I_{\mathcal{O}}(G) = \{X \in B(G) \mid \forall H \subseteq G,\ H \notin \mathcal{N}_{\mathcal{O}}(G) \Rightarrow |X^H| = 0\}.$$

Clearly the morphism $\pi_{\mathcal{O}}$ commutes with induction and restriction. Now assertion (3) of Lemma 3.3.1 shows that $e_H^G \in \mathbb{Q}I_{\mathcal{O}}(G)$ if and only if $e_H^H \in \mathbb{Q}I_{\mathcal{O}}(H)$. It follows that there exists a family of finite groups $\mathcal{N}_{\mathcal{O}}$ such that

$$\mathcal{N}_{\mathcal{O}}(G) = \{H \subseteq G \mid H \in \mathcal{N}_{\mathcal{O}}\}.$$

The main result of this section is a characterization of the family $\mathcal{N}_{\mathcal{O}}$. First a definition:

DEFINITION 3.5.4. Let p be a prime number. If H is a finite group, let $O_p(H)$ denote the largest normal p-subgroup of H. The group H is called *p-hypoelementary*, or *cyclic modulo p*, if the quotient group $H/O_p(H)$ is cyclic. The family of such finite groups is denoted by $\mathcal{Z}_p$. The set of p-hypoelementary subgroups of a finite group G is denoted by $\mathcal{Z}_p(G)$.

It turns out that the family $\mathcal{N}_{\mathcal{O}}$ depends only on p, and is equal to the complement of $\mathcal{Z}_p$, by the following theorem (see [17]):

THEOREM 3.5.5 (Conlon). *Let G be a finite group, and X, Y be finite G-sets. The following conditions are equivalent:*

(1) *The $\mathcal{O}G$-lattices $\mathcal{O}X$ and $\mathcal{O}Y$ are isomorphic.*
(2) *For any p-hypoelementary subgroup H of G, the sets X^H and Y^H have the same cardinality.*

PROOF. Recall that if M is an $\mathcal{O}G$-lattice, then a *vertex* of M is a subgroup P of G, minimal such that M is a direct summand of $\mathrm{Ind}_P^G\mathrm{Res}_P^G M$. The deep argument for Theorem 3.5.5 relies on the Green correspondence, which, for any p-subgroup P of G, is a bijection between the set of isomorphism classes of indecomposable $\mathcal{O}G$-lattices with vertex P and the set of isomorphism classes of indecomposable $\mathcal{O}N_G(P)$-lattices with vertex P (see [1, Ch. 3.12]).

In the case of permutation modules, or more generally of their direct summands (or *p-permutation modules*), there is a more elementary approach, due to Broué ([12]). It relies on the *Brauer construction*: if M is an $\mathcal{O}G$-module, and P is a subgroup of G, then the Brauer quotient $M[P]$ of M at P is defined by

$$M[P] = M^P \Big/ \left(\mathfrak{p}M^P + \sum_{Q \subset P} \mathrm{Tr}_Q^P\left(M^Q\right)\right),$$

where the sum runs over proper subgroups Q of P, and $\mathrm{Tr}_Q^P : M^Q \to M^P$ is the relative trace map or transfer map, defined by

$$\forall m \in M^Q,\ \mathrm{Tr}_Q^P(m) = \sum_{x \in [P/Q]} xm.$$

If S is a Sylow p-subgroup of P, and if $m \in M^P$, then $|P : S|$ is invertible in $\mathcal{O}$, and $m = \mathrm{Tr}_S^P(\frac{1}{|P:S|}m)$. Thus $M[P] = 0$ if P is not a p-subgroup of G. When P is a p-subgroup of G and $M = \mathcal{O}X$ is the permutation lattice associated to a G-set X, the image of the set X^P in $M[P]$ is a k-basis of $M[P]$.

In this case moreover, the value of the Brauer character of M at the element s of G is equal to the number of fixed points of s on X. Hence if $s \in N_G(P)$, the value of the Brauer character of the $kN_G(P)/P$-module $M[P]$ at the element $sP \in N_G(P)/P$ is equal to $|X^{H_{P,s}}|$, where $H_{P,s} = P\langle s\rangle$ is the subgroup of G generated by s and P. Note that $H_{P,s}$ is cyclic modulo p, and that conversely, if H is a p-hypoelementary subgroup of G, there is a p-subgroup P of G and an $s \in N_G(P)$ such that $H = H_{P,s}$.

It follows that if $\mathcal{O}X$ and $\mathcal{O}Y$ are isomorphic, then for any p-hypoelementary subgroup H of G, one has $|X^H| = |Y^H|$.

Conversely, if $|X^H| = |Y^H|$ for any p-hypoelementary subgroup H of G, then for any p-subgroup P of G, the modules $kX^P \simeq \mathcal{O}X[P]$ and $kY^P \simeq \mathcal{O}Y[P]$ have the same Brauer character. The next proposition shows that then $\mathcal{O}X$ and $\mathcal{O}Y$ are isomorphic, and this completes the proof of Theorem 3.5.5. □

PROPOSITION 3.5.6. *Let G be a finite group, and let M and N be p-permutation $\mathcal{O}G$-lattices. Then the following are equivalent:*

(1) *The $\mathcal{O}G$-lattices M and N are isomorphic.*
(2) *For any p-subgroup P of G, the $kN_G(P)/P$-modules $M[P]$ and $N[P]$ have the same Brauer character.*

PROOF. It is clear that (1) implies (2). The proof of the converse is by induction on the cardinality of the set

$$S(M) = \{P \subseteq G \mid M[P] \neq 0\}.$$

Note that if (2) holds, then $S(M) = S(N)$.

If $S(M) = \emptyset$, then $M[1] = M/\mathfrak{p}M = N/\mathfrak{p}N = 0$, hence $M = N = 0$. If $S(M) \neq \emptyset$, choose a maximal element P of $S(M)$. Write $M = M_P \oplus M'$ (resp. $N = N_P \oplus N'$), where all direct summands of M_P (resp. N_P) have non-zero Brauer quotient at P, and where $M'[P] = 0$ (resp. $N'[P] = 0$). Theorem 3.5.7 below shows that $M[P] = M_P[P]$ and $N[P] = N_P[P]$ are projective $kN_G(P)/P$-modules. By assumption, they have the same Brauer character, hence they are isomorphic, and then Theorem 3.5.7 shows that M_P and N_P are isomorphic. Now (2) holds for M' and N', and moreover $|S(M')| < |S(M)|$, since $S(M') \subseteq S(M)$ and $P \in S(M) - S(M')$. Hence $M' \simeq N'$ by induction hypothesis, and $M \simeq N$. □

THEOREM 3.5.7 (Broué). *Let G be a finite group, and P be a p-subgroup of G.*

(1) *Let M be an indecomposable p-permutation $\mathcal{O}G$-lattice with vertex P. Then for any subgroup Q of G, the module $M[Q]$ is non-zero if and only if Q is conjugate to a subgroup of P.*
(2) *The correspondence $M \mapsto M[P]$ induces a bijection between the set of isomorphism classes of indecomposable p-permutation $\mathcal{O}G$-lattices with vertex P and the set of isomorphism classes of indecomposable projective $kN_G(P)/P$-modules.*

PROOF. (1) Recall (Higman criterion) that if H is a subgroup of G and if M is an $\mathcal{O}G$-module, then M is a direct summand of $\mathrm{Ind}_H^G \mathrm{Res}_H^G M$ if and only if there exists an $\mathcal{O}H$-endomorphism ϕ of M such that $\mathrm{Id}_M = \mathrm{Tr}_H^G(\phi)$, where $\mathrm{Tr}_H^G(\phi) = \sum_{g \in [G/H]} g\phi g^{-1}$. If S is a Sylow p-subgroup of H, then $|H:S|$ is invertible in $\mathcal{O}$, and $\mathrm{Id}_M = \mathrm{Tr}_S^G(\phi/|H:S|)$. In particular, any vertex of M is a p-group.

Recall also that if P is a p-group, and Q is a subgroup of P, then the permutation $\mathcal{O}P$-lattice $\mathrm{Ind}_Q^P \mathcal{O}$ is indecomposable, since its reduction modulo $\mathfrak{p}$ is isomorphic to $\mathrm{Ind}_Q^P k$, which has simple socle k. Hence any p-permutation $\mathcal{O}P$-lattice is a permutation $\mathcal{O}P$-lattice.

Let M be an indecomposable p-permutation $\mathcal{O}G$-lattice. Then M is a direct summand of a permutation lattice $\mathcal{O}X$, for some finite G-set X. Let P be a p-subgroup of G such that M is a direct summand of $\mathrm{Ind}_P^G\mathrm{Res}_P^G M$. Then $\mathrm{Res}_P^G M$ is a direct summand of $\mathcal{O}\mathrm{Res}_P^G X$, hence it is a permutation $\mathcal{O}P$-lattice, and there is a set $\mathcal{S}$ of subgroups of P such that

$$\mathrm{Res}_P^G M \simeq \bigoplus_{Q\in\mathcal{S}} \mathrm{Ind}_Q^P \mathcal{O}.$$

Thus M is a direct summand of some lattice $\mathrm{Ind}_Q^G\mathcal{O}$, for $Q \in \mathcal{S}$. Since $\mathrm{Ind}_Q^P\mathcal{O}$ is a direct summand of $\mathrm{Res}_P^G M$, and since $\mathcal{O}$ is a direct summand of $\mathrm{Res}_Q^P\mathrm{Ind}_Q^P\mathcal{O}$, it follows that $\mathcal{O}$ is a direct summand of $\mathrm{Res}_Q^G M$, and that $\mathrm{Ind}_Q^G\mathcal{O}$ is a direct summand of $\mathrm{Ind}_Q^G\mathrm{Res}_Q^G M$.

Now if P is a vertex of M, then $Q = P \in \mathcal{S}$. Hence M is a direct summand of $\mathrm{Ind}_P^G\mathcal{O}$, and $\mathcal{O}$ is a direct summand of $\mathrm{Res}_P^G M$ (in other words the lattice M has *trivial source*). It follows that $M[Q] \neq 0$ for any subgroup of Q. Conversely, if $M[Q] \neq 0$ for some subgroup Q of G, then $(\mathrm{Ind}_P^G\mathcal{O})[Q] \neq 0$, thus Q has a fixed point on the set G/P, i.e. $Q \subseteq_G P$. Assertion (1) follows. It shows in particular that all the vertices of M are conjugate in G.

(2) Since $\mathrm{Ind}_P^G\mathcal{O} \simeq \mathrm{Ind}_{N_G(P)}^G \mathcal{O}N_G(P)/P$, and since the $\mathcal{O}N_G(P)/P$ is a direct sum of indecomposable projective $\mathcal{O}N_G(P)/P$-lattices, it follows from (1) that if M is an indecomposable p-permutation $\mathcal{O}G$-lattice with vertex P, then there is an indecomposable projective $\mathcal{O}N_G(P)/P$-lattice E such that M is a direct summand of $\mathrm{Ind}_{N_G(P)}^G E$. It is easy to check by the Mackey formula that

$$\big(\mathrm{Ind}_{N_G(P)}^G E\big)[P] \simeq \bigoplus_{\substack{x\in[G/N_G(P)]\\ P^x\subseteq N_G(P)}} E\big[P^x P/P\big] = E/\mathfrak{p}E$$

since $E[Q/P] = 0$ for any non-trivial subgroup Q/P of $N_G(P)/P$: indeed E is a direct summand of the permutation lattice $\mathcal{O}N_G(P)/P$ associated to the *free* $N_G(P)/P$-set $Y = N_G(P)/P$, and $Y^{Q/P} = \emptyset$ if $Q \neq P$.

Now $E/\mathfrak{p}E$ is an indecomposable projective $kN_G(P)/P$-module, having $M[P]$ as a non-zero direct summand. It follows that $M[P] \simeq E/\mathfrak{p}E$ is an indecomposable projective $kN_G(P)/P$-module, and the correspondence of assertion (2) is well defined.

Let E be an indecomposable projective $\mathcal{O}N_G(P)/P$-lattice. Write

$$\mathrm{Ind}_{N_G(P)}^G E \simeq L \oplus L',$$

where all the direct summands of L have vertex conjugate to P in G, and no direct summands of L' have vertex conjugate to P. Since L' is a direct summand of $\mathrm{Ind}_P^G\mathcal{O}$, all the indecomposable direct summands of L' have vertex strictly contained in P (up to G-conjugation). Hence $L'[P] = 0$, and $L[P] \simeq E/\mathfrak{p}E$ is indecomposable. It follows that L is indecomposable, and that it is the only indecomposable direct summand of $\mathrm{Ind}_{N_G(P)}^G E$ with vertex P, up to isomorphism.

Thus if M is an indecomposable p-permutation lattice with vertex P, and if E is the only indecomposable projective $\mathcal{O}N_G(P)/P$-lattice such that $M[P] \simeq E/\mathfrak{p}E$, then M is isomorphic to the only indecomposable direct summand of $\mathrm{Ind}_{N_G(P)}^G E$ with vertex P. This shows that the correspondence $M \mapsto M[P]$ of assertion (2) is injective.

Conversely, if $\overline{E}$ is an indecomposable projective $kN_G(P)/P$-module, then there is a projective $\mathcal{O}N_G(P)/P$-lattice E such that $\overline{E} \simeq E/\mathfrak{p}E$. In particular E is indecomposable. Now $\mathrm{Ind}_{N_G(P)}^G E$ has a unique indecomposable direct summand M with vertex P, and $M[P] \simeq \overline{E}$. This shows that the correspondence $M \mapsto M[P]$ of assertion (2) is surjective, and completes the proof of Theorem 3.5.7. □

COROLLARY 3.5.8. *Let G be a finite group. Then*

$$\mathbb{Q}A_{\mathcal{O}}(G) = \sum_{H \in \mathcal{Z}_p(G)} \mathrm{Ind}_H^G \mathbb{Q}A_{\mathcal{O}}(H).$$

Two $\mathcal{O}G$-lattices M and N are isomorphic if and only if for any p-hypoelementary subgroup H of G, the restrictions $\mathrm{Res}_H^G M$ and $\mathrm{Res}_H^G N$ are isomorphic.

PROOF. The image by $\mathbb{Q}\pi_{\mathcal{O}}(e_H^G)$ of the idempotent e_H^G of $\mathbb{Q}B(G)$ in the ring $\mathbb{Q}A_{\mathcal{O}}(G)$ is zero if H is not p-hypoelementary. Thus

$$\mathcal{O} = \mathbb{Q}\pi_{\mathcal{O}}(G/G) = \sum_{H \in \mathcal{Z}_p(G)} \mathbb{Q}\pi_{\mathcal{O}}(e_H^G).$$

Now e_H^G is a linear combination of elements G/K, for subgroups K of H. Since subgroups of p-hypoelementary groups are p-hypoelementary, it follows that there exists rational numbers r_K such that

$$\mathcal{O} = \sum_{K \in \mathcal{Z}_p(G)} r_K \mathrm{Ind}_K^G \mathcal{O}$$

since moreover $\mathbb{Q}\pi_{\mathcal{O}}(G/K)$ is the (class of) the permutation lattice $\mathrm{Ind}_K^G \mathcal{O}$. Tensoring this identity with M over $\mathcal{O}$, and using the Frobenius identity, it follows that

$$M = \sum_{K \in \mathcal{Z}_p(G)} r_K \mathrm{Ind}_K^G \mathrm{Res}_K^G M.$$

This proves both assertions of the corollary. □

REMARK 3.5.9. A different proof of Theorem 3.5.5 and Corollary 3.5.8 has been given by Dress ([25]). It is exposed in Curtis and Reiner ([18, Ch. 11.80D]).

DEFINITION 3.5.10. Let p and q be (non-necessarily distinct) prime numbers. A finite group H is called a *(p,q)-Dress group* if the group $O^q(H)$ is p-hypoelementary. A *(p,q)-Dress subgroup* H of a finite group G is a subgroup of G which is a *(p,q)-Dress group*. The set of (p,q)-Dress subgroups of a finite group G is denoted by $\mathcal{D}_{p,q}(G)$.

THEOREM 3.5.11 (Dress). *Let G be a finite group. Then*

$$A_{\mathcal{O}}(G) = \sum_{\substack{H \in \mathcal{D}_{p,q}(G) \\ \text{any } q}} \mathrm{Ind}_H^G A_{\mathcal{O}}(H).$$

PROOF. This follows from the expression of the idempotents of the Burnside ring $B(G)$: if q is a prime and J is a q-perfect subgroup of G, then the idempotent f_J^G of Corollary 3.3.6 is mapped to zero by $\mathbb{Z}_{(q)}\pi_{\mathcal{O}}$ if J is not p-hypoelementary.

And if J is p-hypoelementary, the idempotent f_J^G is a linear combination with coefficients in $\mathbb{Z}_{(q)}$ of elements G/K, where K runs through the (p,q)-Dress subgroups of G. This shows that there is an integer m_q coprime to q and integers n_K such that

$$m_q\mathcal{O} = \sum_{K \in \mathcal{D}_{p,q}(G)} n_K \mathrm{Ind}_K^G \mathcal{O} \quad \text{in } A_{\mathcal{O}}(G).$$

Setting

$$A'_{\mathcal{O}}(G) = \sum_{\substack{H \in \mathcal{D}_{p,q}(G) \\ \text{any } q}} \mathrm{Ind}_H^G A_{\mathcal{O}}(H)$$

it follows that the quotient $A_{\mathcal{O}}(G)/A'_{\mathcal{O}}(G)$ is a torsion group, with finite exponent coprime to q. Since this holds for any prime q, the groups $A_{\mathcal{O}}(G)$ and $A'_{\mathcal{O}}(G)$ are equal. □

THEOREM 3.5.12 (Brauer). *Let G be a finite group, and k be a field. Then*

$$R_k(G) = \sum_{\substack{H \in \mathcal{E}_q(G) \\ \text{any } q}} \mathrm{Ind}_H^G R_k(H),$$

where for a prime number q, the set $\mathcal{E}_q(G)$ is the set of q-hyperelementary subgroups of G, i.e. the set of subgroups H such that $O^q(H)$ is cyclic.

PROOF. By a similar argument, for any prime q, there is an integer m_q coprime to q, such that $m_q f_J^G$ is in $B(G)$ for any q-perfect subgroup J of G. The Brauer character of $m_q f_J^G$ is zero, unless J is cyclic, and in this case, this character is a linear combination of characters $\mathrm{Ind}_H^G 1$, for $H \in \mathcal{E}_q(G)$. □

3.6. *Further results and references*

The group of units of the Burnside ring has been studied by Matsuda ([35]), Matsuda and Miyata ([36]), and Yoshida ([54]).

Examples of non-isomorphic groups having isomorphic Burnside rings have been given by Thévenaz ([46]).

The Burnside ring of a compact Lie group has been defined and studied by tom Dieck ([19–21]) and Schwänzl ([40,41]).

General exposition of the properties of Burnside rings can be found in Benson ([1, Ch. 5.4]), Curtis and Reiner ([18, Ch. 11]), Karpilovsky ([28, Ch. 15]), tom Dieck ([21]).

4. Invariants

The Burnside ring is an analogue for finite G-sets of the ring $\mathbb{Z}$ for finite sets (and $\mathbb{Z}$ is actually isomorphic to the Burnside ring of the trivial group). To each finite set is associated its cardinality. Similarly, one can attach various elements in the Burnside ring, called *invariants*, to structured G-sets, such as G-posets or G-simplicial complexes.

This section is a self-contained algebraic exposition of the properties of those invariants. The original definitions and methods of Quillen ([38,39]) are used throughout, avoiding however the topological part of this material. Thus for example no use will be made of the geometric realization of a poset, and the accent will be put on acyclic posets rather than contractible ones.

In other words, in order to define and state properties of the invariants attached to finite G-posets in the Burnside ring, one can forget about the fundamental group of those posets, and consider only homology groups.

4.1. *Homology of posets*

Let $(X, \leqslant)$ be a partially ordered set (*poset* for short). As usual, if x, x' are in X, the notation $x < x'$ means $x \leqslant x'$ and $x \neq x'$. The notation $[x, x']_X$ (resp. $[x, x'[_X$, $]x, x']_X$, $]x, x'[_X$) stands for the set of elements $z \in X$ with $x \leqslant z \leqslant x'$ (resp. $x \leqslant z < x'$, $x < z \leqslant x'$, $x < z < x'$). The notation $[x, .[_X$ (resp. $]x, .[_X$, $]., x]_X$, $]., x[_X$) stands for the set of elements $z \in X$ with $x \leqslant z$ (resp. $x < z$, $z \leqslant x$, $z < x$).

If $n \in \mathbb{N}$, let $Sd_n(X)$ denote the set of chains $x_0 < \cdots < x_n$ of elements of X of cardinality $n+1$. The chain complex $C_*(X, \mathbb{Z})$ is the complex of $\mathbb{Z}$-modules defined as follows: for $n \in \mathbb{N}$, the module $C_n(X, \mathbb{Z})$ is the free $\mathbb{Z}$-module with basis $Sd_n(X)$. The differential $d_n : C_n(X, \mathbb{Z}) \to C_{n-1}(X, \mathbb{Z})$ is given by

$$d_n(x_0, \ldots, x_n) = \sum_{i=0}^{n} (-1)^i (x_0, \ldots, \hat{x}_i, \ldots, x_n),$$

where $(x_0, \ldots, \hat{x}_i, \ldots, x_n)$ denotes the chain $(x_0, \ldots, x_n) - \{x_i\}$.

The *reduced* chain complex $\tilde{C}_*(X, \mathbb{Z})$ is the augmented complex obtained by setting $C_{-1}(X, \mathbb{Z}) = \mathbb{Z}$, the augmentation map $d_0 : C_0(X, \mathbb{Z}) \to C_{-1}(X, \mathbb{Z})$ sending each $x_0 \in X$ to $1 \in \mathbb{Z}$.

The nth homology group of the complex $C_*(X, \mathbb{Z})$ (resp. $\tilde{C}_*(X, \mathbb{Z})$) is denoted by $H_n(X, \mathbb{Z})$, and called the nth *homology group* (resp. *reduced homology group*) of X. A poset X is *acyclic* if all its reduced homology groups are equal to zero.

More generally, if K is a ring, the nth homology group of X *with coefficients in* K is the nth homology group of the complex $K \otimes_{\mathbb{Z}} C_*(X, \mathbb{Z})$. When K is a field, one has $H_n(X, K) = K \otimes H_n(X, \mathbb{Z})$.

The *Euler–Poincaré characteristic* $\chi(X)$ of a finite poset X is defined by

$$\chi(X) = \sum_{n \geqslant 0} (-1)^n \operatorname{rank}_{\mathbb{Z}} C_n(X, \mathbb{Z}) = \sum_{n \geqslant 0} (-1)^n \left| Sd_n(X) \right|.$$

Similarly, the *reduced* Euler–Poincaré characteristic $\tilde{\chi}(X)$ is defined by

$$\tilde{\chi}(X) = \sum_{n \geqslant -1} (-1)^n \operatorname{rank}_{\mathbb{Z}} \tilde{C}_n(X, \mathbb{Z}) = \chi(X) - 1.$$

If K is a field, then setting $k_n = \dim_K \operatorname{Ker}(K \otimes_{\mathbb{Z}} d_n)$ for any $n \in \mathbb{N}$

$$\begin{aligned} \dim_K H_n(X, K) &= \dim_K K \otimes_{\mathbb{Z}} H_n(X, \mathbb{Z}) = k_n - \dim_K \operatorname{Im}(K \otimes d_{n+1}) \\ &= k_n + k_{n+1} - \dim_K K \otimes_{\mathbb{Z}} C_{n+1}(X, \mathbb{Z}). \end{aligned}$$

It follows that

$$\chi(X) = \sum_{n \geqslant 0} (-1)^n \dim_K H_n(X, K), \qquad \tilde{\chi}(X) = \sum_{n \geqslant -1} (-1)^n \dim_K \tilde{H}_n(X, K).$$

REMARK 4.1.1. In particular, if X is finite and acyclic, then $\tilde{\chi}(X) = 0$. Similarly, if X and Y are finite posets, and if there is an homotopy equivalence f_* from the complex $C_*(X, \mathbb{Z})$ to the complex $C_*(Y, \mathbb{Z})$, then $\chi(X) = \chi(Y)$, since f induces a group isomorphism from $H_n(X, \mathbb{Z})$ to $H_n(Y, \mathbb{Z})$, for any $n \in \mathbb{N}$.

If X and Y are posets, a *map of posets* $f : X \to Y$ is a map from X to Y such that $f(x) \leqslant f(x')$ whenever x and x' are elements of X such that $x \leqslant x'$. If f is such a map, there is an induced map of chain complexes $C_*(f, \mathbb{Z}) : C_*(X, \mathbb{Z}) \to C_*(Y, \mathbb{Z})$ defined for $n \in \mathbb{N}$ by

$$C_n(f, \mathbb{Z})(x_0, \ldots, x_n) = \begin{cases} \big(f(x_0), \ldots, f(x_n)\big) & \text{if } f(x_0) < \cdots < f(x_n), \\ 0 & \text{otherwise.} \end{cases}$$

One also defines a reduced map $\tilde{C}_*(f, \mathbb{Z}) : \tilde{C}_*(X, \mathbb{Z}) \to \tilde{C}_*(Y, \mathbb{Z})$ by $\tilde{C}_n(f, \mathbb{Z}) = C_n(f, \mathbb{Z})$ if $n \geqslant 0$, and $\tilde{C}_{-1}(f, \mathbb{Z}) = \mathrm{Id}_{\mathbb{Z}}$.

If f and g are maps of posets from X to Y, the notation $f \leqslant g$ means that $f(x) \leqslant g(x)$ for any $x \in X$. The maps f and g are said to be *comparable* if either $f \leqslant g$ or $g \leqslant f$.

LEMMA 4.1.2. *Let f and g be maps of posets from X to Y. If f and g are comparable, then the maps of complexes $C_*(f, \mathbb{Z})$ and $C_*(g, \mathbb{Z})$ are homotopic, as well as the maps $\tilde{C}_*(f, \mathbb{Z})$ and $\tilde{C}_*(g, \mathbb{Z})$.*

PROOF. Suppose for instance that $f \leqslant g$. Consider the map $h_n : C_n(X,\mathbb{Z}) \to C_{n+1}(Y,\mathbb{Z})$ defined for $n \in \mathbb{N}$ by

$$h_n(x_0,\ldots,x_n) = \sum_{i=0}^{n} (-1)^i \big(f(x_0),\ldots,f(x_i),g(x_i),\ldots,g(x_n)\big),$$

where the sequence $(f(x_0),\ldots,f(x_i),g(x_i),\ldots,g(x_n))$ is replaced by 0 if it is not strictly increasing. It is easy to check that

$$d_{n+1} \circ h_n + h_{n-1} \circ d_n = C_n(g,\mathbb{Z}) - C_n(f,\mathbb{Z})$$

thus the maps $C_n(f,\mathbb{Z})$ and $C_n(g,\mathbb{Z})$ are homotopic. A similar argument can be used for the augmented complexes, the map h_{-1} being the zero map. □

COROLLARY 4.1.3 (Quillen). (1) *Let X and Y be posets, and let $f : X \to Y$ and $g : Y \to X$ be maps of posets. If $g \circ f$ is comparable to Id_X and if $f \circ g$ is comparable to Id_Y, then the maps of complexes $\tilde{C}_*(f,\mathbb{Z})$ and $\tilde{C}_*(g,\mathbb{Z})$ are mutual inverse homotopy equivalences between $\tilde{C}_*(X,\mathbb{Z})$ and $\tilde{C}_*(Y,\mathbb{Z})$.*

(2) *If the poset X has a largest element, or a smallest element, the chain complex $\tilde{C}_*(X,\mathbb{Z})$ is contractible.*

PROOF. The first assertion is a direct consequence of the previous lemma. For the second one, denote by m the largest (or the smallest) element of X, and let $\bullet$ denote a poset of cardinality one. Apply assertion (1) to the unique map $f : X \to \bullet$ and to the map $g : \bullet \to X$ sending the unique element of $\bullet$ to m. The result follows, since the complex $\tilde{C}_*(\bullet,\mathbb{Z})$ is clearly contractible. □

4.2. *Invariants attached to finite G-posets*

The following definition of the Lefschetz invariants is due to Thévenaz ([45]):

DEFINITION 4.2.1. Let G be a finite group. A *G-poset* X is a G-set equipped with an order relation $\leqslant$ compatible with the G-action: if $x \leqslant x'$ are elements of X and if $g \in G$, then $gx \leqslant gx'$.

If X and Y are G-posets, a *map of G-posets* $f : X \to Y$ is a map such that $f(gx) = gf(x)$ if $g \in G$ and $x \in X$, and such that $f(x) \leqslant f(x')$ in Y, whenever $x \leqslant x'$ in X. If $y \in Y$, then

$$f^y = \{x \in X \mid f(x) \leqslant y\}, \qquad f_y = \{x \in X \mid f(x) \geqslant y\}.$$

Those sets are sub-G_y-posets of the restriction of X to the stabilizer G_y of y in G.

If $x \leqslant y$ are elements of X, the set $]x,y[_X$ is a $G_{x,y}$-poset, where $G_{x,y}$ is the stabilizer $G_{x,y}$ of the pair (x,y). Similarly, the sets $]x,.[_X$ and $].,x[_X$ are G_x-posets.

If X is a G-poset, then for $n \in \mathbb{N}$, the set $Sd_n(X)$ is a G-set. When X is finite, the *Lefschetz invariant* Λ_X of X is the element of $B(G)$ defined by

$$\Lambda_X = \sum_{n \geqslant 0} (-1)^n Sd_n(X).$$

The *reduced Lefschetz invariant* $\tilde{\Lambda}_X$ is the element of $B(G)$ defined by

$$\tilde{\Lambda}_X = \Lambda_X - G/G.$$

If $x < y$ are elements of X, the *Möbius invariant* $\mu_X(x, y)$ is defined as the Lefschetz invariant of the poset $]x, y[_X$. It is an element of the Burnside ring $B(G_{x,y})$. By convention, the Möbius invariant $\mu_X(x, x)$ is equal to G_x/G_x.

It follows from those definitions that $|\Lambda_X|$ is equal to the Euler–Poincaré characteristic of X. One can say more:

LEMMA 4.2.2. *Let G be a finite group.*
(1) *If X is a finite G-poset, then for any subgroup H of G*

$$(\Lambda_X)^H = \Lambda_{X^H}$$

in $B(N_G(H)/H)$. In particular $|(\Lambda_X)^H| = \chi(X^H)$.
(2) *If X and Y are finite G-posets, then $\Lambda_X = \Lambda_Y$ in $B(G)$ if and only if $\chi(X^H) = \chi(Y^H)$ for any subgroup H of G.*

PROOF. The first assertion is obvious, since $Sd_n(X)^H = Sd_n(X^H)$ for all $n \in \mathbb{N}$. The second one follows from Burnside's Theorem 2.3.2. □

DEFINITION 4.2.3. A G-poset X is called *G-acyclic* if the poset X^H is acyclic for any subgroup H of G.

The following is a direct consequence of this definition:

LEMMA 4.2.4. *Let G be a finite group, and X be a finite G-poset. If X is G-acyclic, then $\tilde{\Lambda}_X = 0$ in $B(G)$.*

PROOF. This follows from Burnside's theorem, since $|(\tilde{\Lambda}_X)^H| = \tilde{\chi}(X^H) = 0$ for any subgroup H of G. □

PROPOSITION 4.2.5. *Let G be a finite group.*
(1) *Let X and Y be finite G-posets, and let $f : X \to Y$ and $g : Y \to X$ be maps of G-posets. If $g \circ f$ is comparable to Id_X and if $f \circ g$ is comparable to Id_Y, then $\tilde{\Lambda}_X = \tilde{\Lambda}_Y$ in $B(G)$.*
(2) *If a G-poset X has a largest element, or a smallest element, then it is G-acyclic.*

PROOF. This is a direct consequence of Corollary 4.1.3 and Remark 4.1.1: for any subgroup H of G, the restrictions of f and g to the posets X^H and Y^H verify the hypotheses of Corollary 4.1.3. Hence $\tilde{\chi}(X^H) = \tilde{\chi}(Y^H)$, and the first assertion follows. For the second, note that the largest (resp. smallest) element of X is also a largest (resp. smallest) element of X^H, for any subgroup H of G. □

EXAMPLE 4.2.6. Let $f : X \to Y$ be a maps of finite G-posets. Denote by $X *_f Y$ the G-poset defined as follows: the underlying G-set is the disjoint union $X \sqcup Y$ of X and Y. The ordering is defined for z and z' in $X *_f Y$ by

$$z \leqslant z' \Leftrightarrow \begin{cases} z, z' \in X & \text{and } z \leqslant z' \text{ in } X, \\ z, z' \in Y & \text{and } z \leqslant z' \text{ in } Y, \\ z \in X, z' \in Y, & \text{and } f(z) \leqslant z' \text{ in } Y. \end{cases}$$

Let f denote the injection from Y to $X *_f Y$, and g denote the map from $X *_f Y$ to Y defined by

$$g(z) = \begin{cases} f(z) & \text{if } z \in X, \\ z & \text{if } z \in Y. \end{cases}$$

Then f and g are maps of G-posets, such that $g \circ f = \mathrm{Id}_Y$ and $\mathrm{Id}_{X *_f Y} \leqslant f \circ g$. It follows that $\Lambda_{X *_f Y} = \Lambda_Y$.

A consequence is the following relation between $\tilde{\Lambda}_X$ and $\tilde{\Lambda}_Y$:

PROPOSITION 4.2.7. *Let $f : X \to Y$ be a map of finite G-posets. Then in $B(G)$*

$$\tilde{\Lambda}_Y = \tilde{\Lambda}_X + \sum_{y \in [G \backslash Y]} \mathrm{Ind}_{G_y}^G \left(\tilde{\Lambda}_{f^y} \tilde{\Lambda}_{]y,.[_Y} \right),$$

$$\tilde{\Lambda}_Y = \tilde{\Lambda}_X + \sum_{y \in [G \backslash Y]} \mathrm{Ind}_{G_y}^G \left(\tilde{\Lambda}_{f_y} \tilde{\Lambda}_{].,y[_Y} \right).$$

In particular, if $\tilde{\Lambda}_{f^y} = 0$ for all $y \in Y$ (for instance if f^y is G_y-acyclic), then $\tilde{\Lambda}_X = \tilde{\Lambda}_Y$ in $B(G)$.

PROOF. Let $n \in \mathbb{N}$. The set $Sd_n(X *_f Y)$ is the disjoint union of $Sd_n(X)$ and of the set of sequences $z_0 < \cdots < z_n$ for which $z_n \in Y$. Such a sequence has a smallest element $y = z_i$ in Y, hence it can be written as

$$x_0 < \cdots < x_{i-1} < y < y_0 < \cdots < y_{n-i-1},$$

where $x_0 < \cdots < x_{i-1}$ is in $Sd_{i-1}(f^y)$ (for $i = 0$, the convention is that $Sd_{-1}(f^y)$ is a set of cardinality one), and $y_0 < \cdots < y_{n-i-1}$ is in $Sd_{n-i-1}(]y, .[_Y)$ (for $i = n$, the convention is that $Sd_{-1}(]y, .[_Y)$ has cardinality one).

Keeping track of the action of G, this leads to the following isomorphism of G-sets

$$\begin{aligned} &Sd_n(X *_f Y)\\ &= Sd_n(X) \sqcup \bigsqcup_{y\in[G\backslash Y]} \mathrm{Ind}_{G_y}^G\Bigl(\bigsqcup_{i=0}^{n}\bigl(Sd_{i-1}(f^y)\times Sd_{n-i-1}(]y,.[_Y)\bigr)\Bigr).\end{aligned}$$

Taking alternating sums gives the first equality of the proposition. The second one follows, by considering the map $f: X^{\mathrm{op}} \to Y^{\mathrm{op}}$ between the opposite posets of X and Y, since clearly $\tilde{\Lambda}_{X^{\mathrm{op}}} = \tilde{\Lambda}_X$ for any finite group G and any finite G-poset X. □

COROLLARY 4.2.8. *Let X be a finite G-poset.*

(1) *The reduced Lefschetz invariant of X is equal to*

$$\tilde{\Lambda}_X = -G/G - \sum_{x\in[G\backslash X]} \mathrm{Ind}_{G_x}^G \tilde{\Lambda}_{]x,.[_X}.$$

(2) *If $x \leqslant y$ in X, then*

$$\sum_{z\in[G_{x,y}\backslash[x,y]]} \mathrm{Ind}_{G_{x,y,z}}^{G_{x,y}} \mathrm{Res}_{G_{x,y,z}}^{G_{y,z}} \mu_X(z,y) = \begin{cases} 0 & \text{if } x<y,\\ G_x/G_x & \text{if } x=y.\end{cases}$$

(3) *If $f: X \to Y$ is a map of finite G-posets, then*

$$\Lambda_X = -\sum_{y\in[G\backslash Y]} \mathrm{Ind}_{G_y}^G \Lambda_{f^y} \tilde{\Lambda}_{]y,.[_Y} = -\sum_{y\in[G\backslash Y]} \mathrm{Ind}_{G_y}^G \Lambda_{f_y} \tilde{\Lambda}_{].,y[_Y}.$$

Assertion (2) is the reason for the name of the Möbius invariant.

PROOF. Assertion (1) follows from the previous proposition, applied to the inclusion $\emptyset \to X$, since $\tilde{\Lambda}_\emptyset = -G/G$. Assertion (2) follows from assertion (1), applied to the $G_{x,y}$-poset $[x, y[_X$, which has a smallest element x if $x < y$.

Assertion (3) follows from assertion (1) and Proposition 4.2.7:

$$\begin{aligned} \tilde{\Lambda}_Y &= \tilde{\Lambda}_X + \sum_{y\in[G\backslash Y]} \mathrm{Ind}_{G_y}^G\bigl((\Lambda_{f^y} - G_y/G_y)\tilde{\Lambda}_{]y,.[_Y}\bigr)\\ &= \Lambda_X - G/G + \sum_{y\in[G\backslash Y]} \mathrm{Ind}_{G_y}^G(\Lambda_{f^y}\tilde{\Lambda}_{]y,.[_Y}) - \sum_{y\in[G\backslash Y]} \mathrm{Ind}_{G_y}^G \tilde{\Lambda}_{]y,.[_Y}\\ &= \Lambda_X + \sum_{y\in[G\backslash Y]} \mathrm{Ind}_{G_y}^G(\Lambda_{f^y}\tilde{\Lambda}_{]y,.[_Y}) + \tilde{\Lambda}_Y.\end{aligned}$$

The second equality in assertion (3) is similar. □

COROLLARY 4.2.9. *Let G be a finite group and X be a finite poset. Denote by $X_\sharp$ (resp. $X^\sharp$) the set of elements $x \in X$ such that $\tilde{\Lambda}_{]\cdot,x[_X} \neq 0$ (resp. $\tilde{\Lambda}_{]x,\cdot[_X} \neq 0$) in $B(G_x)$. If Y is a sub-G-poset of X such that $X_\sharp \subseteq Y \subseteq X$ (resp. $X^\sharp \subseteq Y \subseteq X$), then $\tilde{\Lambda}_Y = \tilde{\Lambda}_X$ in $B(G)$.*

PROOF. By considering the opposite poset, it is enough to prove the assertion for $X_\sharp$, and this can be done by induction on the cardinality of X: if $X = \emptyset$, then $X = X_\sharp$ and there is nothing to prove. For the inductive step, consider the inclusion map $i : X_\sharp \to Y$. It is a map of G-posets. Moreover if $y \in X_\sharp$, then i^y has a largest element y, hence it is G_y-acyclic, and $\tilde{\Lambda}_{G_y} = 0$ in $B(G_y)$. Now if $y \notin X_\sharp$, then $\tilde{\Lambda}_{]\cdot,y[_X} = 0$ in $B(G_y)$ by definition of $X_\sharp$. Moreover in this case

$$\begin{aligned} i^y =]\cdot, y[_X \cap X_\sharp &= \{z \in X \mid z < y,\ \tilde{\Lambda}_{]\cdot,z[_X} \neq 0 \text{ in } B(G_z)\} \\ &\supseteq \{z \in X \mid z < y,\ \tilde{\Lambda}_{]\cdot,z[_X} \neq 0 \text{ in } B(G_z \cap G_y)\} = (]\cdot, y[_X)_\sharp. \end{aligned}$$

It follows that there are inclusions of G_y-posets

$$(]\cdot, y[_X)_\sharp \subseteq i^y \subseteq]\cdot, y[_X.$$

Moreover $|]\cdot, y[_X| < |X|$. By induction hypothesis, it follows that $\tilde{\Lambda}_{i^y} = \tilde{\Lambda}_{]\cdot,y[_X} = 0$. Now the corollary follows from Proposition 4.2.7. □

4.3. *Steinberg invariants*

DEFINITION 4.3.1. Let G be a finite group, and p be a prime number. The *Steinberg invariant* $St_p(G)$ of G at p is the reduced Lefschetz invariant of the poset $s_p(G)$ of non-trivial p-subgroups of G, on which G acts by conjugation.

The reason for this terminology is that if G is a finite simple Chevalley group in characteristic p, then the (virtual) permutation character associated to $St_p(G)$ is equal up to a sign to the Steinberg character of G.

PROPOSITION 4.3.2. *Let G be a finite group, and p be a prime number. Then $St_p(G) = 0$ in $B(G)$ if and only if G has a non-trivial normal p-subgroup.*

PROOF. If $St_p(G) = 0$, then in particular $\tilde{\chi}(s_p(G)^G) = 0$. Thus $s_p(G)^G$ is non-empty, and G has a non-trivial normal p-subgroup.

Conversely, suppose that $R \neq \mathbb{1}$ is a non-trivial normal p-subgroup of G. Let f be the map from $s_p(G)$ to $[R, \cdot[_{s_p(G)}$ defined by $f(Q) = Q.R$, and let g denote the inclusion map from $[R, \cdot[_{s_p(G)}$ to $s_p(G)$. Then f and g are maps of G-posets, and moreover $\mathrm{Id} \leqslant g \circ f$ and $f \circ g = \mathrm{Id}$. Since $[R, \cdot[_{s_p(G)}$ has a smallest element, it follows from Proposition 4.2.5 that $St_p(G) = 0$ in $B(G)$. □

REMARK 4.3.3. Quillen has conjectured that $s_p(G)$ is contractible if and only if $O_p(G) \neq \mathbb{1}$. The above proof shows actually that $s_p(G)$ is G-contractible if and only if $O_p(G) \neq \mathbb{1}$ (see Thévenaz and Webb [48]).

PROPOSITION 4.3.4. *Let G be a finite group, and p be a prime number. Let $a_p(G)$ denote the sub-G-poset of $s_p(G)$ consisting of non-trivial elementary Abelian p-subgroups of G, and $b_p(G)$ denote the sub-G-poset of $s_p(G)$ consisting of non-trivial p-subgroups P of G such that $P = O_p(N_G(P))$. Then*

$$St_p(G) = \tilde{\Lambda}_{a_p(G)} = \tilde{\Lambda}_{b_p(G)} \quad \textit{in } B(G).$$

PROOF. Let P be a non-trivial p-subgroup of G. Suppose P is not elementary Abelian, and denote by $\Phi(P)$ the Frattini subgroup of P. Let $f :]\cdot, P[_{s_p(G)} \to [\Phi(P), P[_{s_p(G)}$ be the map defined by $f(Q) = Q.\Phi(P)$, and let g be the inclusion map from $[\Phi(P), P[_{s_p(G)}$ to $]\cdot, P[_{s_p(G)}$. Then f and g are maps of $N_G(P)$-posets. Moreover $f \circ g = \mathrm{Id}$, and $\mathrm{Id} \leqslant g \circ f$. Now $[\Phi(P), P[_{s_p(G)}$ has a smallest element, thus $\tilde{\Lambda}_{]\cdot, P[_{s_p(G)}} = 0$ in $B(N_G(P))$ by Proposition 4.2.5. In other words

$$\big(s_p(G)\big)_\sharp \subseteq a_p(G) \subseteq s_p(G)$$

and Corollary 4.2.9 shows that $\tilde{\Lambda}_{s_p(G)} = \tilde{\Lambda}_{a_p(G)}$.

The other equality is similar: let P be a non-trivial p-subgroup of G. Set $R = O_p(N_G(P)/P)$, and suppose $R \neq P$. Let $f :]P, \cdot[_{s_p(G)} \to s_p(N_G(P)/P)$ defined by $f(Q) = N_Q(P)/P$, and let $g : s_p(N_G(P)/P) \to]P, \cdot[_{s_p(G)}$ defined by $g(Q/P) = Q$. Then $f \circ g = \mathrm{Id}$, and $g \circ f \leqslant \mathrm{Id}$. By Proposition 4.2.5, it follows that

$$\tilde{\Lambda}_{]P,\cdot[_{s_p(G)}} = \mathrm{Inf}_{N_G(P)/P}^{N_G(P)} St_p\big(N_G(P)/P\big)$$

and this is zero since $O_p(N_G(P)/P) = R/P \neq \mathbb{1}$. Hence $\tilde{\Lambda}_{]P,\cdot[_{s_p(G)}} = 0$ in $B(N_G(P)/P)$. In other words

$$\big(s_p(G)\big)^\sharp \subseteq b_p(G) \subseteq s_p(G)$$

and the proposition follows from Corollary 4.2.9. $\square$

NOTATION 4.3.5. For the remainder of this section, the group G will be a finite group, and $\underline{\mathcal{F}}$ will denote a family of subgroups of G such that:

(1) $\mathbb{1} \in \underline{\mathcal{F}}$.
(2) $\underline{\mathcal{F}}$ is closed under conjugation.
(3) $\underline{\mathcal{F}}$ is closed under products, i.e. if P and Q are elements of $\underline{\mathcal{F}}$ such that P normalizes Q, then $P.Q \in \underline{\mathcal{F}}$.

If H is a subgroup of G, then $\underline{\mathcal{F}}(H)$ is the set of subgroups of H which are in $\underline{\mathcal{F}}$. It is a family of subgroups of H having the above three properties.

Denote by $\mathcal{F}$ the family $\underline{\mathcal{F}}$ with the trivial group removed. Then $\underline{\mathcal{F}}$ and $\mathcal{F}$ are ordered by inclusion of subgroups, and they are G-posets. When H is a subgroup of G, denote similarly by $\mathcal{F}(H)$ the set of non-trivial subgroups of H which are in $\mathcal{F}$.

DEFINITION 4.3.6. The *Steinberg invariant* $St_{\mathcal{F}}(G)$ of G with respect to $\mathcal{F}$ is the reduced Lefschetz invariant of the G-poset $\mathcal{F}$.

Thus if $\mathcal{F} = s_p(G)$, then $St_{\mathcal{F}}(G) = St_p(G)$.

LEMMA 4.3.7. *Let G and $\mathcal{F}$ as in* 4.3.5. *If $P \in \mathcal{F}$, then $St_{\mathcal{F}}(G)^P = 0$ in $B(N_G(P)/P)$.*

PROOF. Let a denote the inclusion map from $[P, .[_{\mathcal{F}}$ to $\mathcal{F}^P$, and b denote the map from $\mathcal{F}^P$ to $[P, .[_{\mathcal{F}}$ defined by $b(P) = FP$. Then $b \circ a = \mathrm{Id}_{[P,.[_{\mathcal{F}}}$, and $\mathrm{Id}_{\mathcal{F}^P} \leqslant a \circ b$. Thus $St_{\mathcal{F}}(G)^P = \tilde{\Lambda}_{\mathcal{F}^P} = \tilde{\Lambda}_{[P,.[_{\mathcal{F}}} = 0$ since $[P, .[_{\mathcal{F}}$ has a smallest element P. □

THEOREM 4.3.8 (Bouc). *Let G and $\mathcal{F}$ as in* 4.3.5. *If $X \in B(G)$, set*

$$St_{\mathcal{F}}(G, X) = \sum_{P \in [G \backslash \underline{\mathcal{F}}]} \mathrm{Ind}_{N_G(P)}^{G}\big(\mu_{\underline{\mathcal{F}}}(\mathrm{II}, P)\, \mathrm{Inf}_{N_G(P)/P}^{N_G(P)} X^P\big).$$

Then the map $X \mapsto St_{\mathcal{F}}(G, X)$ is an idempotent group endomorphism of $B(G)$, and its image is the set of elements X of $B(G)$ such that $X^P = 0$ in $B(N_G(P)/P)$ for all $P \in \mathcal{F}$.

PROOF. Clearly if $X^P = 0$ for $P \in \mathcal{F}$, then $St_{\mathcal{F}}(G, X) = X$. Hence the only thing to check is that if $P \in \mathcal{F}$, then $St_{\mathcal{F}}(G, X)^P = 0$.

By linearity, one can suppose that X is a finite G-set, viewed as a G-poset for the discrete ordering. Let Z denote the subposet of $X \times \mathcal{F}$ consisting of pairs (x, P) such that $P \subseteq G_x$. Let $a : Z \to X$ defined by $a((x, P)) = x$. Then a is a map of G-posets, and a^x is isomorphic to $\mathcal{F}(G_x)$.

Now obviously $\Lambda_X = X$ for any discrete G-poset X. Moreover all the elements of X are maximal in X, thus $\tilde{\Lambda}_{]x,.[_X} = -G_x/G_x$ for all $x \in X$. It follows from Proposition 4.2.7 that

$$\Lambda_X = X = \Lambda_Z - \sum_{x \in [G \backslash X]} \mathrm{Ind}_{G_x}^{G} St_{\mathcal{F}(G_x)}(G_x). \tag{4.3.9}$$

Now let $b : Z \to \mathcal{F}$ be the map defined by $b((x, P)) = P$. Then b is a map of G-posets, and for $Q \in \mathcal{F}$

$$g_Q = \big\{(x, P) \in X \times \mathcal{F} \mid Q \subseteq P \subseteq G_x\big\}.$$

The maps $c : q_Q \to X^Q$ and $d : X^Q \to g_Q$ defined by $c((x, P)) = x$ and $d(x) = (x, Q)$ are maps of $N_G(Q)$-posets, such that $d \circ c \leqslant \mathrm{Id}_{g_Q}$ and $c \circ d = \mathrm{Id}_{X^Q}$. Thus by Corollary 4.2.8

$$\Lambda_Z = - \sum_{Q \in [G \backslash \mathcal{F}]} \mathrm{Ind}_{N_G(Q)}^{G}\big(\mu_{\underline{\mathcal{F}}}(\mathrm{II}, Q) X^Q\big).$$

Hence Eq. (4.3.9) gives

$$X = -\sum_{Q\in[G\backslash\mathcal{F}]} \mathrm{Ind}_{N_G(Q)}^G\big(\mu_{\underline{\mathcal{F}}}(\mathbb{1}, Q)X^Q\big) - \sum_{x\in[G\backslash X]} \mathrm{Ind}_{G_x}^G St_{\mathcal{F}(G_x)}(G_x)$$

and finally

$$St_{\mathcal{F}}(G, X) = -\sum_{x\in[G\backslash X]} \mathrm{Ind}_{G_x}^G St_{\mathcal{F}(G_x)}(G_x).$$

Now for any $x \in X$ and any $P \in \mathcal{F}$, if $P \subseteq G_x$, then $St_{\mathcal{F}(G_x)}(G_x)^P = 0$ by Lemma 4.37. Theorem 4.3.8 follows from assertion (3) of Proposition 2.2.1. □

REMARK 4.3.10. Another proof of Theorem 4.3.8 can be found in [5], where decompositions of $B(G)$ associated to $\mathcal{F}$ are constructed.

COROLLARY 4.3.11. *Let G and $\underline{\mathcal{F}}$ as in 4.3.5, let p be a prime number, and let $\mathcal{O}$ be a complete local Noetherian commutative ring with residue field of characteristic p. Suppose that $\mathcal{F}$ contains the set $s_p(G)$ of non-trivial p-subgroups of G. Then for any finite G-poset X the image of $St_{\mathcal{F}}(G, X)$ in $A_{\mathcal{O}}(G)$ is a linear combination of projective $\mathcal{O}G$-lattices.*

PROOF. Indeed, in this case, if Y is an element of $B(G)$ such that $Y^P = 0$ for $P \in \mathcal{F}$, then in particular $Y^P = 0$ for any non-trivial p-subgroup P of G. It follows that $|Y^H| = 0$ whenever H is a subgroup of G with $O_p(H) \neq \mathbb{1}$.

Now $Y = \sum_H |Y^H| e_H^G$ in $\mathbb{Q}B(G)$, and by Theorem 3.5.5, the idempotent e_H^G maps to zero in $\mathbb{Q}A_{\mathcal{O}}(G)$ if H is not p-hypoelementary. If H is hypoelementary and if $O_p(H) = \mathbb{1}$, then H is a cyclic p'-group, and the idempotent e_H^G is a linear combination of elements G/K, for subgroups K of H.

Those remarks show that the image of Y in $A_{\mathcal{O}}(G)$ is a linear combination of permutation lattices $\mathrm{Ind}_K^G \mathcal{O}$, for cyclic p'-subgroups K of G. The corollary follows, since those lattices are $\mathcal{O}G$-projective. □

REMARK 4.3.12. If $X = G/G$, then $St_{\mathcal{F}}(G, X) = -St_{\mathcal{F}}(G)$. Hence the image of $St_{\mathcal{F}}(G)$ in $A_{\mathcal{O}}(G)$ is a linear combination of projective $\mathcal{O}G$-lattices. This has been used by Webb ([50]) to compute the cohomology of an $\mathcal{O}G$-module in terms of the cohomology of the stabilizers of chains of elements of $\mathcal{F}$.

Later, Webb ([51]) showed a stronger result: the chain complex $\tilde{C}_*(\mathcal{F}, \mathbb{Z}_p)$ is the direct sum of a complex of projective $\mathbb{Z}_pG$-lattices, and of a split acyclic augmented subcomplex.

COROLLARY 4.3.13. *Let G and $\underline{\mathcal{F}}$ as in 4.3.5, and suppose moreover that $\underline{\mathcal{F}}$ is closed under taking subgroups.*

(1) *If $P \in \underline{\mathcal{F}}$ and $X \in B(G)$, there exists an integer $m_{P,X}$ such that*

$$\mathrm{Res}_P^G St_{\mathcal{F}}(G, X) = m_{P,X}\, P/\mathbb{1}.$$

(2) *Denote by* $|G|_{\mathcal{F}}$ *the l.c.m of the orders of elements of* $\mathcal{F}$. *Then for any finite* G-*poset* X

$$\tilde{\chi}(X) + \sum_{P \in \mathcal{F}} \mu_{\underline{\mathcal{F}}}(\mathbb{1}, P)\tilde{\chi}(X^P) \equiv 0 \left(|G|_{\mathcal{F}}\right)$$

and in particular

$$\tilde{\chi}(\mathcal{F}) \equiv 0 \left(|G|_{\mathcal{F}}\right).$$

PROOF. Let Y be an element of $B(G)$ such that $Y^P = 0$ for all $P \in \mathcal{F}$. Consider the restriction of Y to an element Q of $\mathcal{F}$. Then

$$\mathrm{Res}_Q^G Y = \sum_{R \in [s_Q]} |Y^R| e_R^Q.$$

Since $\underline{\mathcal{F}}$ is closed under taking subgroups, the only non-zero term in this sum is obtained for $R = \mathbb{1}$. Hence

$$\mathrm{Res}_Q^G Y = |Y| e_{\mathbb{1}}^Q = \frac{|Y|}{|Q|} Q/\mathbb{1}.$$

This shows that $|Y| \equiv 0$ $(|Q|)$, and that $\mathrm{Res}_Q^G Y$ is an integer multiple of $Q/\mathbb{1}$. The first assertion follows, for $Y = St_{\mathcal{F}}(G, X)$. Taking cardinalities gives the first congruence. The second one follows from Lemma 4.3.7 in the case $X = \mathcal{F}$. □

REMARK 4.3.14. When $\mathcal{F}$ is the poset $s_p(G)$ of non-trivial p-subgroups of G, the second congruence is $\tilde{\chi}(s_p(G)) \equiv 0$ $(|G|_p)$, and it is due to Brown ([13]). A proof of this congruence using the idempotents of the Burnside ring was given by Gluck ([26]) and Yoshida ([53]). Similar congruences have been stated by Thévenaz ([44]) and Brown and Thévenaz ([14]).

The computation of the Steinberg invariants of the symmetric groups, or more generally of wreath products of a finite group with symmetric groups, can be found in [6]. It requires the use of a ring of formal power series with coefficients in Burnside rings.

5. The Mackey and Green functor structure

The notion of a Mackey functor is a formal generalization of the properties of induction, restriction, and conjugation exposed in Section 2.2. The notion of a Green functor keeps track moreover of the ring structure. Both have several equivalent definitions, that are quickly recalled below. The Burnside functor is a universal object in this framework also.

5.1. *Mackey functors and subgroups*

Let G be a finite group, and R be a ring. Let R-**Mod** denote the category of R-modules. The first definition of Mackey functors is due to Green ([27]):

DEFINITION 5.1.1 (Green). A *Mackey functor* M for the group G over R (or with values in R-**Mod**) consists of the following data:

- For each subgroup H of G, an R-module $M(H)$.
- Whenever $H \subseteq K$ are subgroups of G with $H \subseteq K$, a map of R-modules $t_H^K : M(H) \to M(K)$, called transfer or induction, and a map of R-modules $r_H^K : M(K) \to M(H)$ called restriction.
- For each subgroup H of G and each element $x \in G$, a map of R-modules $c_{x,H} : M(H) \to M({}^xH)$, called conjugation.

Those maps are subject to four types of conditions:

(1) (Triviality conditions) For any subgroup H of G, and any $h \in H$, the maps t_H^H, r_H^H, and $c_{h,H}$ are equal to the identity map of $M(H)$.

(2) (Transitivity conditions) If $H \subseteq K \subseteq L$ are subgroups of G, then $t_K^L \circ t_H^K = t_H^L$ and $r_H^K \circ r_K^L = r_H^L$. If $x, y \in G$, then $c_{y,{}^xH} \circ c_{x,H} = c_{yx,H}$.

(3) (Compatibility conditions) If $H \subseteq K$ are subgroups of G and if x is an element of G, then $c_{x,K} \circ t_H^K = t_{{}^xH}^{{}^xK} \circ c_{x,H}$ and $c_{x,H} \circ r_H^K = r_{{}^xH}^{{}^xK} \circ c_{x,K}$.

(4) (Mackey axiom) If $H \subseteq K \supseteq L$ are subgroups of G, then

$$r_H^K \circ t_L^K = \sum_{x \in [H \backslash K / L]} t_{H \cap {}^xL}^H \circ c_{x,H^x \cap L} \circ r_{H^x \cap L}^L.$$

If M and N are Mackey functors for G over R, then a *morphism of Mackey functors* $f : M \to N$ is a collection of morphisms of R-modules $f_H : M(H) \to N(H)$, which commute with the maps t_H^K, r_H^K and $c_{x,H}$.

The category of Mackey functors for G over R is denoted by $Mack_R(G)$.

REMARK 5.1.2. One can check easily that if the above conditions (1), (2) and (3) hold, then the right-hand side of the equality in the Mackey axiom (4) does not depend on the choice of the set of representatives $[H \backslash K / L]$.

REMARK 5.1.3. If M is a Mackey functor for G over R, and H is a subgroup of G, then the maps $c_{x,H}$ for $x \in N_G(H)$ define a structure of $RN_G(H)/H$-module on $M(H)$. In particular $M(\mathbb{1})$ is an RG-module.

EXAMPLE 5.1.4. The Burnside Mackey functor B is the Mackey functor with values in $\mathbb{Z}$-**Mod** whose value at H is the Burnside ring $B(H)$. If $H \subseteq K$ are subgroups of G, then $t_H^K = \mathrm{Ind}_H^K$ and $r_H^K = \mathrm{Res}_H^K$. The conjugation maps $c_{x,H}$ are defined by $c_{x,H}(Z) = {}^xZ$ for $Z \in B(H)$ and $x \in G$.

More generally, if R is a ring, the Burnside functor RB is defined by "tensoring everything with R", i.e. setting $RB(H) = R \otimes_{\mathbb{Z}} B(H)$, and extending the maps t_H^K, r_H^K and $c_{x,H}$ in the obvious way.

Another example of Mackey functor is the Green ring functor $A_{\mathcal{O}}$, whose value at H is the Green ring $A_{\mathcal{O}}(H)$ of $\mathcal{O}H$-lattices defined in Section 3.5. The transfer is given by induction of lattices, and the operations of restriction and conjugation are the obvious ones. It is clear that the morphism $\pi_{\mathcal{O}}$ actually defines a morphism of Mackey functors from the Burnside functor to the Green ring functor.

5.2. *Mackey functors and G-sets*

If M is a Mackey functor for G over R, if H and K are subgroups of G, and if $x \in G$ is such that $H^x \subseteq K$, then there are maps of R-modules

$$a_x = t_{H^x}^K \circ c_{x^{-1},H} : M(H) \to M(K), \qquad b_x = c_{x,H^x} \circ r_{H^x}^K : M(K) \to M(H).$$

Moreover, if $k \in K$, then since M is a Mackey functor

$$\begin{aligned} a_{xk} &= t_{H^{xk}}^K \circ c_{(xk)^{-1},H} = c_{k^{-1}x^{-1},K} \circ t_H^{{}^xK} = c_{k^{-1},K} \circ c_{x^{-1},{}^xK} \circ t_H^{{}^xK} \\ &= c_{x^{-1},{}^xK} \circ t_H^{{}^xK} = a_x. \end{aligned}$$

Similarly $b_{xk} = b_x$. Hence a_x and b_x only depend on the class xK. The crucial observation is that the set of classes xK such that $H^x \subseteq K$ is in one to one correspondence with the set of G-sets homomorphisms from G/H to G/K. This leads to the second definition of Mackey functors, due to Dress ([25]).

Recall that a *bivariant* functor M from a category $\mathcal{C}$ to a category $\mathcal{D}$ is a pair of functors $M = (M_*, M^*)$, where M_* is a functor from $\mathcal{C}$ to $\mathcal{D}$ and M^* is a functor from $\mathcal{C}$ to $\mathcal{D}^{\mathrm{op}}$ (or a cofunctor from $\mathcal{C}$ to $\mathcal{D}$), which coincide on objects, i.e. $M_*(C) = M^*(C)$ for all objects C of $\mathcal{C}$. This common value is simply denoted by $M(C)$.

DEFINITION 5.2.1 (Dress). A *Mackey functor* M for the group G with values in the category R-**Mod** of R-modules is a bivariant functor $M = (M_*, M^*)$ from the category of finite G-sets to R-**Mod**, with the following two properties:

(1) Let X and Y be any finite G-sets, and let i_X (resp. i_Y) denote the canonical injection from X (resp. Y) into $X \sqcup Y$. Then the morphisms

$$\big(M_*(i_X), M_*(i_Y)\big) : M(X) \oplus M(Y) \to M(X \sqcup Y),$$

$$\begin{pmatrix} M^*(i_X) \\ M^*(i_Y) \end{pmatrix} : M(X \sqcup Y) \to M(X) \oplus M(Y)$$

are mutually inverse isomorphisms.

(2) Let

$$\begin{tikzcd} X \arrow[r, "a"] \arrow[d, "b"'] & Y \arrow[d, "c"] \\ Z \arrow[r, "d"'] & T \end{tikzcd}$$

be any Cartesian (i.e. pull-back) square of finite G-sets. Then

$$M_*(b) \circ M^*(a) = M^*(d) \circ M_*(c).$$

A *morphism of Mackey functors* $f : M \to N$ is a natural transformation of bivariant functors, i.e. a collection of morphisms of R-modules $f_X : M(X) \to N(X)$, for finite G-sets X, which commute with the maps $M_*(a)$ and $M^*(a)$ for any morphism of finite G-sets $a : X \to Y$.

If M is a Mackey functor for the first definition, then M yields a Mackey functor $\widehat{M}$ for this definition by choosing a set of representatives $[G\backslash X]$, for each finite G-set X, and then setting

$$\widehat{M}(X) = \bigoplus_{x \in [G\backslash X]} M(G_x).$$

Let $f : X \to Y$ be a morphism of G-sets. If $x \in [G\backslash X]$ and $y \in [G\backslash Y]$ are such that $f(x) \in Gy$, then there exists $g \in G$ such that $f(x) = gy$. In this case $G_x \subseteq {}^g G_y$, and one can define two maps

$$\alpha_{y,x} = t_{{}^g G_x}^{G_y} \circ c_{g, G_x} : M(G_x) \to M(G_y),$$

$$\beta_{x,y} = c_{g^{-1}, {}^g G_x} \circ r_{{}^g G_x}^{G_y} : M(G_y) \to M(G_x).$$

Those maps depend only on x and y, and do not depend on the chosen element g. Define moreover $\alpha_{y,x} = \beta_{x,y} = 0$ for $x \in [G\backslash X]$ and $y \in [G\backslash Y]$ if $f(x) \notin Gy$.

Then the map $\widehat{M}_*(f)$ is defined by the block matrix $(\alpha_{y,x})_{y \in [G\backslash Y], x \in [G\backslash X]}$, and the map $\widehat{M}^*(f)$ is defined by the block matrix $(\beta_{x,y})_{x \in [G\backslash X], y \in [G\backslash Y]}$. One can check that $\widehat{M}$ is a Mackey functor for the second definition.

Conversely, a Mackey functor $\widehat{M}$ for the second definition yields a Mackey functor M for the first definition, by setting

$$M(H) = \widehat{M}(G/H), \qquad t_H^K = \widehat{M}_*(\pi_H^K),$$
$$r_H^K = \widehat{M}^*(\pi_H^K), \qquad c_{x,H} = \widehat{M}_*(\gamma_{x,H}),$$

where the morphism π_H^K is the canonical projection $G/H \to G/K$ for $H \subseteq K$, and $\gamma_{x,H}$ is the isomorphism $yH \mapsto yx^{-1}{}^xH$ from G/H to $G/{}^xH$, for $x \in G$.

EXAMPLE 5.2.2. One can show (see for instance [9, Ch. 2.4]) that the Mackey functor associated to the Burnside functor B (still denoted by B) can be described as follows: if X is a finite G-set, let $G\text{-}\mathbf{set}{\downarrow}_X$ denote the category the objects of which are the finite G-sets over X, i.e. the pairs (Y, f), where Y is a finite G-set and $f : Y \to X$ is a map of G-sets. A morphism in $G\text{-}\mathbf{set}{\downarrow}_X$ from (Y, f) to (Z, g) is a map of G-sets $h : Y \to Z$ such that $g \circ h = f$. The composition of morphisms is the composition of maps.

Then $B(X)$ is the Grothendieck group of the category G-$\mathbf{set}\!\downarrow_X$, for the relations given by decomposition into disjoint union. If $\phi : X \to X'$ is a morphism of G-sets, then the map $B_*(\phi)$ sends (Y, f) to $(Y, \phi \circ f)$, and the map $M^*(\phi)$ sends (Z, g) to the G-set over X defined by the pull-back square

$$\begin{array}{ccc} Y & \longrightarrow & Z \\ \big\downarrow & & \big\downarrow{\scriptstyle g} \\ X & \underset{\phi}{\longrightarrow} & X' \end{array}$$

5.3. *Mackey functors as modules*

A third definition of Mackey functors is due to Thévenaz and Webb ([49]), who define the following algebra:

DEFINITION 5.3.1 (Thévenaz and Webb). Let G be a finite group, and R be a commutative ring. The *Mackey algebra* $\mu_R(G)$ is the (associative, unital) R-algebra with generators t_H^K, r_H^K, and $c_{x,H}$, for subgroups $H \subseteq K$ of G, and $x \in G$, subject to the following relations:

- If H is a subgroup of G, and if $h \in H$, then $t_H^H = r_H^H = c_{h,H}$. Moreover $\sum_{H \subseteq G} t_H^H = 1_{\mu_R(G)}$, and if $H \neq K$ are subgroups of G, then $t_H^H t_K^K = 0$.
- If $H \subseteq K \subseteq L$ are subgroups of G, then $t_K^L \circ t_H^K = t_H^L$ and $r_H^K \circ r_K^L = r_H^L$. If $x, y \in G$, then $c_{y,{}^xH} \circ c_{x,H} = c_{yx,H}$.
- If $H \subseteq K$ are subgroups of G and $x \in G$, then $c_{x,K} \circ t_H^K = t_{{}^xH}^{{}^xK} \circ c_{x,H}$ and $c_{x,H} \circ r_H^K = r_{{}^xH}^{{}^xK} \circ c_{x,K}$.
- If $H \subseteq K \supseteq L$ are subgroups of G, then

$$r_H^K \circ t_L^K = \sum_{x \in [H \backslash K / L]} t_{H \cap {}^xL}^H \circ c_{x, H^x \cap L} \circ r_{H^x \cap L}^L.$$

A *Mackey functor* for G over R is a just a $\mu_R(G)$-module, and a *morphism of Mackey functors* is a morphism of $\mu_R(G)$-modules.

If M is a Mackey functor over R for the first definition, set

$$\widetilde{M} = \bigoplus_{H \subseteq G} M(H).$$

This is endowed with an obvious structure of $\mu_R(G)$-module: for instance, the action of the generator t_H^K of $\mu_R(G)$ is zero on the component $M(H')$ of $\widetilde{M}$, if $H' \neq H$, and it is equal to the map $t_H^K : M(H) \to M(K)$ on the component $M(H)$.

Conversely, being given a $\mu_R(G)$-module $\widetilde{M}$, one recovers a Mackey functor for the first definition by setting $M(H) = t_H^H \widetilde{M}$. The transfer maps, restriction maps, and conjugation maps for M are obtained by multiplication by the generators of $\mu_R(G)$ with the same name: for instance, the relations of the Mackey algebra show that $t_H^K M(H) \subseteq M(K)$.

5.4. *Green functors*

Roughly speaking, a *Green functor* for the finite group G over the commutative ring R is a "Mackey functor with a compatible ring structure". More precisely, there are two equivalent definitions of Green functors. The first one is due to Green ([27]):

DEFINITION 5.4.1. A *Green functor* A for G over R is a Mackey functor for G over R, such that for any subgroup H of G, the R-module $A(H)$ has the structure of an R-algebra (associative, with unit), with the following properties:

(1) If K is a subgroup of G and $x \in G$, then the map $c_{x,H}$ is a morphism of rings (with unit) from $A(K)$ to $A({}^xK)$.
(2) If $H \subseteq K$ are subgroups of G, then the map r_H^K is a morphism of rings (with unit) from $A(K)$ to $A(H)$.
(3) (Frobenius identities) Under the same conditions, if $a \in A(K)$ and $b \in A(H)$, then

$$a.t_H^K(b) = t_H^K\big(r_H^K(a).b\big), \qquad t_H^K(b).a = t_H^K\big(b.r_H^K(a)\big).$$

If A and A' are Green functors for G over R, a *morphism of Green functors* $f : A \to A'$ is a morphism of Mackey functors such that for each subgroup H of G, the map $f_H : A(H) \to A'(H)$ is a ring homomorphism. The morphism is *unitary* if moreover all the maps f_H are *unitary*, or equivalently, if f_G is unitary.

EXAMPLE 5.4.2. The Burnside Mackey functor RB is a Green functor, if the product on $RB(H)$ for $H \subseteq G$ is defined by linearity from the direct product of H-sets. The Frobenius identities follow from Proposition 2.2.1.

The second definition of Green functors is in terms of G-sets, and is detailed in [9, Ch. 2.2]:

DEFINITION 5.4.3. A *Green functor* A for G over R is a Mackey functor for G over R, endowed for any G-sets X and Y with R-bilinear maps $A(X) \times A(Y) \to A(X \times Y)$ with the following properties:

- (Bifunctoriality) If $f : X \to X'$ and $g : Y \to Y'$ are morphisms of G-sets, then the squares

$$\begin{array}{ccc} A(X) \times A(Y) & \xrightarrow{\times} & A(X \times Y) \\ {\scriptstyle A_*(f)\times A_*(g)}\downarrow & & \downarrow{\scriptstyle A_*(f\times g)} \\ A(X') \times A(Y') & \xrightarrow[\times]{} & A(X' \times Y') \end{array}$$

$$\begin{array}{ccc} A(X) \times A(Y) & \xrightarrow{\times} & A(X \times Y) \\ {\scriptstyle A^*(f)\times A^*(g)}\uparrow & & \uparrow{\scriptstyle A^*(f\times g)} \\ A(X') \times A(Y') & \xrightarrow[\times]{} & A(X' \times Y') \end{array}$$

are commutative.

- (Associativity) If X, Y and Z are G-sets, then the square

$$\begin{array}{ccc} A(X)\times A(Y)\times A(Z) & \xrightarrow{\mathrm{Id}_{A(X)}\times(\times)} & A(X)\times A(Y\times Z) \\ {\scriptstyle (\times)\times \mathrm{Id}_{A(Z)}}\downarrow & & \downarrow{\scriptstyle \times} \\ A(X\times Y)\times A(Z) & \xrightarrow[\times]{} & A(X\times Y\times Z) \end{array}$$

  is commutative, up to identifications $(X \times Y) \times Z \simeq X \times Y \times Z \simeq X \times (Y \times Z)$.
- (Unitarity) If $\bullet$ denotes the trivial G-set G/G, there exists an element $\varepsilon_A \in A(\bullet)$, called the unit of A, such that for any G-set X and for any $a \in A(X)$

$$A_*(p_X)(a \times \varepsilon_A) = a = A_*(q_X)(\varepsilon_A \times a)$$

  denoting by p_X (resp. q_X) the (bijective) projection from $X \times \bullet$ (resp. from $\bullet \times X$) to X.

If A and A' are Green functors for G over R, a morphism of Green functors $f : A \to A'$ is a morphism of Mackey functors such that for any G-sets X and Y, the square

$$\begin{array}{ccc} A(X)\times A(Y) & \xrightarrow{\times} & A(X\times Y) \\ {\scriptstyle f_X\times f_Y}\downarrow & & \downarrow{\scriptstyle f_{X\times Y}} \\ A'(X)\times A'(Y) & \xrightarrow[\times]{} & A'(X\times Y) \end{array}$$

is commutative. The morphism f is unitary if moreover $f_\bullet(\varepsilon_A) = \varepsilon_{A'}$.

EXAMPLE 5.4.4. If X and Y are finite G-sets, then the map $B(X) \times B(Y) \to B(X \times Y)$ is defined by linearity from the map sending the finite G-set U over X and the finite G-set V over Y to the Cartesian product $U \times V$, which is a G-set over $X \times Y$. The element $\varepsilon_B \in B(\bullet)$ is the image of the trivial G-set $\bullet = G/G$.

REMARK 5.4.5. There is an obvious notion of module over a Green functor: a (left) module M over the Green functor A is a Mackey functor M endowed for any G-sets X and Y with bilinear maps $A(X) \times M(Y) \to M(X \times Y)$ which are bifunctorial, associative, and unitary. With this definition, a Mackey functor for G over R is just a module over the Green functor RB. This can be viewed as a generalization of the identification of Abelian groups with $\mathbb{Z}$-modules.

5.5. *Induction, restriction, inflation*

The second definition of Mackey functors leads to a natural notion of induction, restriction, and inflation for Mackey functors:

DEFINITION 5.5.1. Let G be a finite group, and R be a commutative ring.

- If H is a subgroup of G, and M is a Mackey functor for G over R, then the *restriction* $\mathrm{Res}_H^G M$ is the Mackey functor for H over R obtained by composition of the functor $M : G\text{-}\mathbf{set} \to R\text{-}\mathbf{Mod}$ with the induction functor $\mathrm{Ind}_H^G : H\text{-}\mathbf{set} \to G\text{-}\mathbf{set}$. If A is a Green functor for G over R, then $\mathrm{Res}_H^G A$ has a natural structure of Green functor for H over R.
- If H is a subgroup of G, and N is a Mackey functor for H over R, then the *induced* Mackey functor $\mathrm{Ind}_H^G M$ is the Mackey functor for G over R obtained by composition of the functor $M : H\text{-}\mathbf{set} \to R\text{-}\mathbf{Mod}$ with the restriction functor $\mathrm{Res}_H^G : G\text{-}\mathbf{set} \to H\text{-}\mathbf{set}$. If A is a Green functor for H over R, then $\mathrm{Ind}_H^G A$ has a natural structure of Green functor for H over R.
- If K is a normal subgroup of G, and N is a Mackey functor for G/K, then the inflated Mackey functor $\mathrm{Inf}_{G/K}^G N$ is the Mackey functor for G over R obtained by composition of the functor $M : G/K\text{-}\mathbf{set} \to R\text{-}\mathbf{Mod}$ with the fixed points functor $G\text{-}\mathbf{set} \to G/K\text{-}\mathbf{set}$. If A is a Green functor for G/K over R, then $\mathrm{Inf}_{G/K}^G A$ has a natural structure of Green functor for G over R.

REMARK 5.5.2. If M is a Mackey functor for G, and if $K \subseteq H$ are subgroups of G, then

$$\big(\mathrm{Res}_H^G M\big)(K) = \big(\mathrm{Res}_H^G M\big)(H/K) = M\big(\mathrm{Ind}_H^G H/K\big) = M(G/K) = M(K)$$

so the above definition of $\mathrm{Res}_H^G M$ coincides with the naive one. In particular, the restriction of the Burnside functor for G to the subgroup H is isomorphic to the Burnside functor for H.

REMARK 5.5.3. The constructions of Definition 5.5.1 are examples of functors between categories of Mackey functors, obtained by composition with functors between the corresponding categories of G-sets. A uniform description of this kind of functors is given in [7], using the formalism of *bisets*. In [9, Ch. 8], it is shown that those constructions also apply to Green functors.

Another fundamental example of this kind of functors is the following:

DEFINITION 5.5.4 (Dress). If X is a finite G-set, and if M is is a Mackey functor for G over R, then the Mackey functor M_X is the Mackey functor for G over R obtained by composition of the functor $M : G\text{-}\mathbf{set} \to R\text{-}\mathbf{Mod}$ with the endofunctor $Y \mapsto Y \times X$ of G-**set**. If A is a Green functor for G over R, then A_X has a natural structure of Green functor for G over R.

The endofunctor $M \mapsto M_X$ on $Mack_R(G)$ is denoted by $\mathcal{I}_X$.

REMARK 5.5.5. In the case of a transitive G-set $X = G/H$, the isomorphism

$$G/H \times Y \simeq \mathrm{Ind}_H^G \mathrm{Res}_H^G Y$$

of Proposition 2.2.1 leads to an isomorphism of Mackey functors

$$M_{G/H} \simeq \mathrm{Ind}_H^G \mathrm{Res}_H^G M.$$

Induction and restriction are mutual left and right adjoint:

PROPOSITION 5.5.6 (Thévenaz and Webb). *Let G be a finite group, let H be a subgroup of G, and R be a commutative ring.*

(1) *The functors*

$$\mathrm{Ind}_H^G : Mack_R(H) \to Mack_R(G) \quad \textit{and} \quad \mathrm{Res}_H^G : Mack_R(G) \to Mack_R(H)$$

are mutual left and right adjoint.

(2) *For any finite G-set X, the endofunctor $\mathcal{I}_X$ is self adjoint.*

PROOF. For assertion (1), see [47, Proposition 4.2]. Assertion (2) follows trivially. □

5.6. *The Burnside functor as projective Mackey functor*

The third definition of Mackey functors shows that the category $Mack_R(G)$ is an Abelian category, with enough projective objects. A sequence of Mackey functors

$$0 \to L \to M \to N \to 0$$

is exact if and only if for each subgroup H of G, the sequence

$$0 \to L(H) \to M(H) \to N(H) \to 0$$

is an exact sequence of R-modules.

The Burnside functor RB provides the first example of a projective Mackey functor:

PROPOSITION 5.6.1. *Let G be a finite group, and R be a commutative ring. Set $\bullet = G/G$.*

(1) *If M is a Mackey functor for G over R, then the map*

$$\Theta_M : f \in \mathrm{Hom}_{Mack_R(G)}(RB, M) \mapsto f_\bullet(\bullet) \in M(\bullet) = M(G)$$

is an isomorphism of R-modules.

(2) *The Burnside functor RB is a projective object in $Mack_R(G)$.*

(3) *The map Θ_{RB} is an isomorphism of rings (with unit)*

$$\Theta_{RB} : \mathrm{End}_{Mack_R(G)}(RB) \overset{\simeq}{\to} RB(G).$$

PROOF. Let X be a finite G-set, let (Y, a) be a G-set over X, and denote also by (Y, a) its image in $RB(X) = R \otimes_{\mathbb{Z}} B(X)$. Then clearly $(Y, a) = RB_*(a)(Y, \mathrm{Id}_Y)$. Denote by $p_Y : Y \to \bullet$ be the unique morphism of G-sets. Then the square

$$\begin{array}{ccc} Y & \xrightarrow{p_Y} & \bullet \\ {\scriptstyle \mathrm{Id}_Y}\downarrow & & \downarrow{\scriptstyle \mathrm{Id}_\bullet} \\ Y & \xrightarrow[p_Y]{} & \bullet \end{array}$$

is Cartesian. Hence denoting by ε_{RB} the element $(\bullet, \mathrm{Id}_\bullet)$ of $RB(\bullet)$, it follows that

$$(Y,a) = RB_*(a)RB^*(p_Y)(\varepsilon_{RB}).$$

Now a morphism $f: RB \to M$ is entirely determined by the element $u = f_\bullet(\varepsilon_{RB})$ of $M(\bullet) = M(G)$, since the map f_X must verify

$$f_X\big((Y,a)\big) = M_*(a)M^*(p_Y)(u).$$

Conversely, if $u \in M(\bullet)$ is given, this equality defines a map of R-modules f_X from $RB(X)$ to $M(X)$. If $g: X \to X'$ is a morphism of G-sets, then obviously

$$M_*(g)f_X\big((Y,a)\big) = M_*(ga)M^*(p_Y)(u) = f_{X'}\big((Y,ga)\big) = f_{X'}RB_*(g)\big((Y,a)\big).$$

Similarly, if the square

$$\begin{tikzcd} Y \arrow[r, "h"] \arrow[d, "a"'] & Y' \arrow[d, "a'"] \\ X \arrow[r, "g"'] & X' \end{tikzcd}$$

is Cartesian, then

$$\begin{aligned} M^*(g)f_{X'}\big((Y',a')\big) &= M^*(g)M_*(a')M^*(p_{Y'})(u) = M_*(a)M^*(h)M^*(p_{Y'})(u) \\ &= M_*(a)M^*(p_Y)(u) = f_X\big((Y,a)\big) \\ &= f_X RB^*\big((Y',a')\big). \end{aligned}$$

This proves that f is a morphism of Mackey functors, and assertion (1) follows. Assertion (2) is now clear, since the evaluation functor $M \mapsto M(G)$ from $Mack_R(G)$ to R-**Mod** is exact.

Now let f and g be two endomorphisms of RB, and suppose that $f_\bullet(\bullet)$ (resp. $g_\bullet(\bullet)$) is equal to (U, p_U) (resp. (V, p_V)) for some G-set U (resp. V). Then

$$\begin{aligned} (g \circ f)_\bullet(\bullet) &= g_\bullet\big((U,p_U)\big) = RB_*(p_U)RB^*(p_U)\big((V,p_V)\big) \\ &= RB_*(p_U)\big((U \times V, p_1)\big) = (U \times V, p_{U\times V}), \end{aligned}$$

where $p_1: U \times V \to U$ is the first projection, since the square

$$\begin{tikzcd} U \times V \arrow[r] \arrow[d, "p_1"'] & V \arrow[d, "p_V"] \\ U \arrow[r, "p_U"'] & \bullet \end{tikzcd}$$

is Cartesian.

By linearity, this shows that Θ_{RB} is a ring homomorphism. Clearly Θ_{RB} maps the identity endomorphism to the trivial G-set $\bullet$, which is the unit of $RB(G)$. Assertion (3) follows. □

COROLLARY 5.6.2. *Let G be a finite group and R be a commutative ring.*

(1) *If H is a subgroup of G, then $\mathrm{Ind}_H^G RB$ is a projective Mackey functor for G over R.*

(2) *If X is a finite G-set, then RB_X is a projective Mackey functor for G over R.*

PROOF. The first assertion is a special case of the second one, since

$$\mathrm{Ind}_H^G RB \simeq \mathrm{Ind}_H^G \mathrm{Res}_H^G B \simeq RB_{G/H}.$$

Now if M is any Mackey functor for G over R, there are isomorphisms of R-modules

$$\mathrm{Hom}_{Mack_R(G)}(RB_X, M) \simeq \mathrm{Hom}_{Mack_R(G)}(RB, M_X) \simeq M_X(\bullet) \simeq M(X)$$

and the second assertion follows, since evaluation at X is an exact functor from $Mack_R(G)$ to R-**Mod**. □

The following proposition states a link between the Burnside ring and the Mackey algebra:

PROPOSITION 5.6.3 (Thévenaz and Webb). *Let G be a finite group, and R be a commutative ring. Set $\Omega_G = \bigsqcup_{H \subseteq G} G/H$.*

(1) *There is an isomorphism of Mackey functors*

$$\mu_R(G) \simeq RB_{\Omega_G} \simeq \bigoplus_{H \subseteq G} \mathrm{Ind}_H^G \mathrm{Res}_H^G RB.$$

(2) *Any Mackey functor is a quotient of a direct sum of induced Burnside Mackey functors.*

PROOF. By adjunction, there are isomorphisms of R-modules

$$\mathrm{Hom}_{Mack_R(G)}(RB_{\Omega_G}, M) \simeq \mathrm{Hom}_{Mack_R(G)}(RB, M_{\Omega_G}) \simeq M_{\Omega_G}(G) = M(\Omega_G)$$

and this can be explicitely described as follows. Let $f \in \mathrm{Hom}_{Mack_R(G)}(RB_{\Omega_G}, M)$. Then in particular $f_{\Omega_G} \in \mathrm{Hom}_R(RB(\Omega_G^2), M(\Omega_G))$. Let Δ_{Ω_G} denote the diagonal inclusion of Ω_G in Ω_G^2. Then $\delta = (\Omega_G, \Delta_{\Omega_G}) \in RB(\Omega_G^2)$. Set

$$\theta_M(f) = f_{\Omega_G}(\delta) \in M(\Omega_G).$$

If X is a finite G-set, then any element in $RB_{\Omega_G}(X) = RB(X \times \Omega_G)$ is a linear combination of finite G-sets over $X \times \Omega_G$. If (Y, α) is such a G-set, and if α_1 (resp. α_2) denotes the composition of α with the projection from $X \times \Omega_G$ to X (resp. to Ω_G), then clearly in $RB_{\Omega_G}(X)$

$$(Y, \alpha) = (RB_{\Omega_G})_*(\alpha_1) \circ (RB_{\Omega_G})^*(\alpha_2)(\delta).$$

It follows that one must have

$$f_X\big((Y, \alpha)\big) = M_*(\alpha_1) \circ M^*(\alpha_2)\big(f_{\Omega_G}(\delta)\big).$$

Thus f is determined by the element $\theta_M(f)$ of $M(\Omega_G)$. Conversely, if $m \in M(\Omega_G)$ is given, then the formula

$$f_X\big((Y, \alpha)\big) = M_*(\alpha_1) \circ M^*(\alpha_2)(m)$$

defines a morphism from $RB(X \times \Omega_G)$ to $M(X)$, that is easily seen to induce a morphism of Mackey functors.

These maps between $\mathrm{Hom}_{Mack_R(G)}(RB_{\Omega_G}, M)$ and $M(\Omega_G)$ are inverse to each other, and it follows that the map

$$\theta_M : \mathrm{Hom}_{Mack_R(G)}(RB_{\Omega_G}, M) \to M(\Omega_G)$$

is an isomorphism of R-modules.

Moreover if $\phi : M \to N$ is a morphism of Mackey functors, then

$$\theta_N(\phi \circ f) = (\phi \circ f)_{\Omega_G}(\delta) = \phi_{\Omega_G} \circ f_{\Omega_G}(\delta) = \phi_{\Omega_G} \circ \theta_M(f).$$

In other words, the functors $M \mapsto \mathrm{Hom}_{Mack_R(G)}(RB_{\Omega_G}, M)$ and $M \mapsto M(\Omega_G)$ from $Mack_R(G)$ to R-**Mod** are isomorphic. But from the point of view of $\mu_R(G)$-modules, there is an isomorphism of R-modules

$$\mathrm{Hom}_{\mu_R(G)}\big(\mu_R(G), \widetilde{M}\big) \simeq \widetilde{M} \simeq M(\Omega_G)$$

and this is also functorial in M. It follows that there is an isomorphism of functors from $\mu_R(G)$-**Mod** to R-**Mod**

$$\mathrm{Hom}_{\mu_R(G)}\big(\mu_R(G), -\big) \simeq \mathrm{Hom}_{\mu_R(G)}\big(\widetilde{RB_{\Omega_G}}, -\big).$$

By Yoneda's Lemma, this shows that there is an isomorphism of $\mu_R(G)$-modules

$$\mu_R(G) \simeq \widetilde{RB_{\Omega_G}} = RB\big(\Omega_G^2\big).$$

This proves assertion (1), since $RB_{\Omega_G} \simeq \bigoplus_{H \subseteq G} \mathrm{Ind}_H^G \mathrm{Res}_H^G RB$. Now assertion (2) is clear, since the restriction of a Burnside functor is a Burnside functor. □

The isomorphism $RB(\Omega_G^2) \simeq \mu_R(G)$ can be detailed as follows: let M be the Mackey functor defined by $M(H) = t_H^H \mu_R(G) \subseteq \mu_R(G)$. Then by construction $M(\Omega_G) = \tilde{M} =$

$\mu_R(G)$. The unit of $\mu_R(G)$ is the element $\sum_H t_H^H$, and it corresponds by Yoneda's Lemma to the morphism from $RB(\Omega_G^2)$ to $M(\Omega_G)$ mapping the G-set (Y, α) over Ω_G^2 to $M_*(\alpha_1)M^*(\alpha_2)(\sum_H t_H^H)$. Using the isomorphism of G-sets

$$RB(\Omega_G^2) \simeq \bigoplus_{H,K \subseteq G} RB(G/H \times G/K) \tag{5.6.4}$$

it follows that the map sending the G-set $(G/L, \pi_{H,x,K})$ over $G/H \times G/K$ defined for $x \in G$ and $L \subseteq H \cap {}^xK$ by

$$\pi_{H,x,K}(gL) = \big(gH, g^x K\big) \quad \text{for } g \in G$$

to the element

$$M_*\big((\pi_{H,x,K})_1\big)M^*\big((\pi_{H,x,K})_2\big)\bigg(\sum_H t_H^H\bigg) = t_L^H c_{x,L^x} r_{L^x}^K$$

of $\mu_R(G)$, is an isomorphism of R-modules. Thus (see [49, Propositions (3.2) and (3.3)], or [49, Ch. 4.4]):

PROPOSITION 5.6.5 (Thévenaz and Webb). *Let G be a finite group and R be a commutative ring. Then the algebra $\mu_R(G)$ is a finitely generated free R-module, with basis the set of elements $t_L^H c_{x,L^x} r_{L^x}^K$, for subgroups H and K of G, for $x \in [H\backslash G/K]$, and a subgroup L of $H \cap {}^xK$ up to $H \cap {}^xK$-conjugation.*

PROOF. Indeed the G-sets $(G/L, \pi_{H,x,K})$, for subgroups H and K of G, for $x \in [H\backslash G/K]$, and $L \subseteq H \cap {}^xK$ up to $H \cap {}^xK$-conjugation, form an R-basis of $RB(\Omega_G^2)$, up to identifications induced by isomorphism (5.6.4). $\square$

REMARK 5.6.6. It follows in particular that the rank over R of $\mu_R(G)$ does not depend on R. In other words, the algebra $\mu_R(G)$ is isomorphic to $R \otimes_{\mathbb{Z}} \mu_{\mathbb{Z}}(G)$. It is sometimes convenient to define $\mu_R(G) = R \otimes_{\mathbb{Z}} \mu_{\mathbb{Z}}(G)$ for any ring R (not necessarily commutative). The case $R = \mu_S(G)$ for a commutative ring S is of interest (see [9, Ch. 1.2]). Not that if R is not commutative, then $\mu_R(G)$ is not strictly speaking an R-algebra.

REMARK 5.6.7. The previous remarks show that the R-linear map defined by

$$t_L^H c_{x,L^x} r_{L^x}^K \in \mu_R(G) \mapsto (G/L, \pi_{H,x,K}) \in RB\big(\Omega_G^2\big)$$

is an isomorphism of R-modules. Of course, this is *not* an algebra isomorphism, since $\mu_R(G)$ is not commutative in general. One can show (see [9, Ch. 4.5.1]) that the multiplication law it induces on $RB(\Omega_G^2)$ is given by

$$(V, W) \in B\big(\Omega_G^2\big)^2 \mapsto V \circ_{\Omega_G} W = B_*\begin{pmatrix} xyz \\ xz \end{pmatrix} B^*\begin{pmatrix} xyz \\ xyyz \end{pmatrix}(V \times W),$$

where $\left(\begin{smallmatrix}xyz\\xz\end{smallmatrix}\right)$ is the map $(x, y, z) \in \Omega_G^3 \mapsto (x, z) \in \Omega_G^2$, and $\left(\begin{smallmatrix}xyz\\xyyz\end{smallmatrix}\right)$ is the map $(x, y, z) \in \Omega_G^3 \mapsto (x, y, y, z) \in \Omega_G^4$, and $V \times W$ is the product for the Green functor structure, defined in Example 5.4.4.

In the case where V and W are G-sets over Ω_G^2, it is easy to see that $V \circ_{\Omega_G} W$ is the G-set obtained by pullback

$$\begin{tikzcd}
& & V \circ_{\Omega_G} W \arrow[dl] \arrow[dr] & & \\
& V \arrow[dl] \arrow[dr] & & W \arrow[dl] \arrow[dr] & \\
\Omega_G & & \Omega_G & & \Omega_G
\end{tikzcd}$$

This has been used by Lindner ([31]) to give an alternative definition of Mackey functors. It is also the starting point of tom Dieck ([22, Ch. IV.8]) and Lewis ([29]) to define Mackey functors for compact Lie groups.

5.7. *The Burnside functor as a universal Green functor*

The following proposition shows that the Burnside functor is an initial object in the category of Green functors:

PROPOSITION 5.7.1. *Let G be a finite group, and R be a commutative ring. If A is a Green functor for G over R, then there is a unique unitary morphism of Green functor from RB to A.*

PROOF. By Proposition 5.6.1, there is a unique morphism f of Mackey functors from RB to A such that $f_\bullet(\bullet) = \varepsilon_A$. It is easy to check that f is a morphism of Green functors (see [9, Proposition 2.4.4] for details). $\square$

REMARK 5.7.2. For any Green functor A for G over R, there is a natural algebra structure on $A(\Omega_G^2)$, defined as in Remark 5.6.7 (see [9, Ch. 4]). Moreover (see [9, Ch. 12.2]), one can build a Green functor ζ_A for G over R such that $\zeta_A(\bullet) = Z(A(\Omega_G^2))$: the evaluation of ζ_A at the finite G-set X is the set of natural transformations from the identity endofunctor $\mathcal{I} = \mathcal{I}_\bullet$ of $Mack_R(G)$ to the endofunctor $\mathcal{I}_X$. The previous proposition shows that for any Green functor A, there is a ring homomorphism from $RB(G)$ to $Z(A(\Omega_G^2))$.

Hence there should exist a ring homomorphism from $RB(G)$ to the center of the Mackey algebra. But the center of a ring S is isomorphic to the endomorphism algebra of the identity functor on S-**Mod**. Thus, for any Mackey functor M, there should be a morphism from $RB(G)$ to the endomorphism algebra of M, with functorial properties in M.

It can be defined as follows: if X is a finite G-set, then there is a natural transformation $a_X : M \to M_X$ defined for the finite G-set Z by

$$a_{X,Z} = M^*\begin{pmatrix} zx \\ z \end{pmatrix} : M(Z) \to M(Z \times X),$$

where $\binom{zx}{z}$ denotes the projection map $Z \times X \to Z$. Similarly, there is a natural transformation $b_X : M_X \to M$ defined by

$$b_{X,Z} : M_*\begin{pmatrix} zx \\ z \end{pmatrix} : M(Z \times X) \to M(Z).$$

This defines an endomorphism $z(X)_M$ of M by $z(X)_{M,Z} = b_{X,Z} \circ a_{X,Z}$. It is easy to check that for finite G-sets X and Y

$$z(X \sqcup Y)_M = z(X)_M + z(Y)_M, \qquad z(X)_M \circ z(Y)_M = z(X \times Y)_M.$$

This leads by linearity to an R-algebra homomorphism

$$X \in RB(G) \mapsto z(X)_M \in \mathrm{End}_{Mack_R(G)}(M)$$

for any Mackey functor M for G over R.

Moreover, if $f : M \to N$ is a morphism of Mackey functors, then the square

$$\begin{array}{ccc} M & \xrightarrow{z(X)_M} & M \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ N & \xrightarrow[z(X)_N]{} & N \end{array}$$

is commutative.

The ring homomorphism from $RB(G)$ to $Z(\mu_R(G))$ can be described concretely as follows: if H is a subgroup of G, then the relations of the Mackey algebra show that there is a (non-unitary) ring homomorphism β_H from $RB(H)$ to $\mu_R(G)$ defined by $\beta_H(H/K) = t_K^H r_K^H$. With this notation

PROPOSITION 5.7.3. *Let G be a finite group, and R be a commutative ring. Then the map $z : RB(G) \to \mu_R(G)$ defined by*

$$z(X) = \sum_{H \subseteq G} \beta_H\big(\mathrm{Res}_H^G X\big)$$

is an injective unitary ring homomorphism from $RB(G)$ to $Z(\mu_R(G))$.

PROOF. By inspection, or see [49, Proposition 9.2]. □

Taking the image by z of a set of primitive idempotents of $RB(G)$ will give a decomposition of the unit of $\mu_R(G)$ into a sum of orthogonal central idempotents. This in turn will give a decomposition of the category of Mackey functors as a direct sum of Abelian categories.

NOTATION 5.7.4. If G is a finite group and R is a commutative ring, denote by $\pi = \pi_R(G)$ be the set of prime factors of $|G|$ which are not invertible in R. If J is a π-perfect subgroup of G, denote by $Mack_R(G, J)$ the subcategory of $Mack_R(G)$ on which the idempotent $z(f_J^G)$ acts trivially. If M is a Mackey functor for G over R, denote by $f_J^G \times M$ the direct summand of M on which $z(f_J^G)$ acts trivially.

THEOREM 5.7.5 (Thévenaz and Webb). *Let G be a finite group, and R be a commutative ring. Set $\pi = \pi_R(G)$.*

(1) *If J is a π-perfect subgroup of G, then the direct summand $f_J^G \times B$ of B has a natural structure of Green functor. Moreover, there are isomorphisms of Green functors*

$$f_J^G \times B \simeq \mathrm{Ind}_{N_G(H)}^G \mathrm{Inf}_{N_G(H)/H}^G \left(f_{\mathbb{1}}^{N_G(H)/H} \times B\right),$$

$$RB \simeq \bigoplus_{\substack{J \in [s_G] \\ J = O^{\pi}(J)}} f_J^G \times RB.$$

(2) *The functor*

$$M \mapsto \mathrm{Ind}_{N_G(J)}^G \mathrm{Inf}_{N_G(J)/J}^{N_G(J)} M$$

is an equivalence of categories from $Mack_R(N_G(J)/J, \mathbb{1})$ to the category $Mack_R(G, J)$.

PROOF. See [49, Section 10], (or [9, Proposition 2.1.11] for a generalization to arbitrary Green functors). □

COROLLARY 5.7.6. *Let G be a finite group, and π be a set of prime numbers. If J is a π-perfect subgroup of G, then there is a ring isomorphism*

$$f_J^G \mathbb{Z}_{(\pi)} B(G) \simeq f_{\mathbb{1}}^{N_G(J)/J} \mathbb{Z}_{(\pi)} B\big(N_G(J)/J\big).$$

PROOF. By evaluation at the trivial G-set of the first isomorphism of Green functors in Theorem 5.7.5, in the case $R = \mathbb{Z}_{(\pi)}$. Of course, there is also an elementary proof, using the ring homomorphism $X \mapsto X^J$ from $B(G)$ to $B(N_G(J)/J)$. □

REMARK 5.7.7. By definition, the category $Mack_R(G, \mathbb{I})$ is the category of modules over the algebra $\mu_R(G, \mathbb{I}) = z(f_{\mathbb{I}}^G)\mu_R(G)$. One can show that $\mu_R(G, \mathbb{I})$ is equal to the R-submodule of $\mu_R(G)$ generated by the elements $t_P^H c_{x,P^x} r_{P^x}^K$, where H and K are subgroups of G, where $x \in G$, and P is a π-solvable subgroup of $H \cap {}^x K$.

One can also show (see [10, Lemme 2.2], for the case $|\pi| = 1$) that this algebra is Morita-equivalent to its subalgebra $\mu_{R,\mathbb{I}}(G)$ generated by the elements $t_P^H c_{x,P^x} r_{P^x}^K$, where $x \in G$ and H and K are π-solvable subgroups of G. The $\mu_{R,\mathbb{I}}(G)$-modules can be viewed as "Mackey functors defined only on π-solvable subgroups".

Recall (see [51, Section 3.2]) that a Mackey functor M is said to be *projective relative to a finite G-set X* if the morphism $b_X : M_X \to M$ is split surjective. If $\mathcal{X}$ is a set of subgroups of G, the functor M is said to be projective relative to $\mathcal{X}$ if it is projective relative to the G-set $X = \bigsqcup_{H \in \mathcal{X}} G/H$. With these definitions, one can show that the Mackey functors in $Mack_R(G, \mathbb{I})$ are exactly the Mackey functors which are projective relative to π-solvable subgroups of G, for $\pi = \pi_R(G)$.

REMARK 5.7.8. If R is a field k of characteristic p, then the Krull–Schmidt theorem holds for finitely generated $\mu_k(G)$-modules. If M is an indecomposable projective Mackey functor in *Mack*$(G, \mathbb{I})$, then the previous remarks show that there is a p-subgroup P such that M is a direct summand of $\mathrm{Ind}_P^G B$. In particular, the kG-module $M(\mathbb{I})$ is a direct summand of $(\mathrm{Ind}_P^G B)(\mathbb{I})$, which is easily seen to be isomorphic to $\mathrm{Ind}_P^G k$. In particular, the module $M(\mathbb{I})$ is a p-permutation module.

This argument can be refined, and leads to the following nice interpretation of the indecomposable projective objects in $Mack_k(G, \mathbb{I})$, in terms of p-permutation modules:

THEOREM 5.7.9 (Thévenaz and Webb). *Let G be a finite group, and k be a field of characteristic $p > 0$. Then evaluation at the trivial subgroup $M \mapsto M(\mathbb{I})$ induces a one to one correspondence between the set of isomorphism classes of indecomposable projective objects in $Mack_k(G, \mathbb{I})$ and the set of isomorphism classes of indecomposable p-permutation kG-modules.*

PROOF. See [49, Theorem 12.7]. □

6. The Burnside ring as a biset-functor

6.1. *Bisets*

In the previous section, the Burnside functor was defined on the subgroups of a given finite group G. The three operations relating the Burnside rings of subgroups of G were induction, restriction, and conjugation.

But there are other natural operations on the Burnside ring, which have not yet been considered: if N is a normal subgroup of the group G, then the *inflation functor* $\mathrm{Inf}_{G/N}^G : G/N$-**set** $\to G$-**set** introduced in Section 2.2 induces a ring homomorphism

$\mathrm{Inf}_{G/N}^{G} : B(G/N) \to B(G)$, and the *deflation functor* $\mathrm{Def}_{G/N}^{G} : G\text{-}\mathbf{set} \to G/N\text{-}\mathbf{set}$ induces a group homomorphism $\mathrm{Def}_{G/N}^{G} : B(G) \to B(G/N)$.

The common point about all those operations is the following: in all cases, there are two finite groups G and H, and a functor F from G-**set** to H-**set**, preserving disjoint unions, i.e. such that $F(X \sqcup Y) \simeq F(X) \sqcup F(Y)$. I all cases moreover there exists a finite set U with a left H-action and a right G-action, such that

$$F(X) \simeq U \times_G X,$$

where $U \times_G X$ is the quotient of $U \times X$ by the right action of G given for $(u, x) \in U \times X$ and $g \in G$ by

$$(u, x)g = \big(ug, g^{-1}x\big).$$

The set $U \times X$ is an H-set for the action defined by

$$h(u, x) = (hu, x)$$

for $(u, x) \in U \times X$ and $h \in H$. If the actions of G and H on U commute, i.e. if $(hu)g = h(ug)$ for all $u \in U$, $h \in H$, and $g \in G$, then this action passes down to a well defined left action of H on $U \times_G X$. This leads to the following definition:

DEFINITION 6.1.1. Let G and H be groups. An *H-set-G*, or a *biset* for short, is a set with a left $(H \times G^{\mathrm{op}})$-action, i.e. a set U with a left H-action and a right G-action which commute.

If K is another group, and V is a K-set-H, then the product $V \times_H U$ is the quotient of the product $V \times U$ by the right action of H given by $(v, u)h = (vh, h^{-1}u)$ for $v \in V$, $u \in U$, and $h \in H$. The class of (v, u) in $V \times_H U$ is denoted by $(v,_H u)$.

The set $V \times_H U$ is a K-set-G for the action given by

$$k(v,_H u)g = (kv,_H ug)$$

for $k \in K$, $g \in G$, $u \in U$, and $v \in V$.

EXAMPLE 6.1.2. For the induction functor Ind_H^G from a subgroup H of G to G, the set U is the set G, with its natural left G-action and right H-action by multiplication.

For the restriction functor Res_H^G, the set U is the set G, for its left H-action and right G-action.

For the inflation functor $\mathrm{Inf}_{G/N}^{G}$, the set U is the set G/N, for its left G-action given by projection and multiplication, and its right G/N-action given by multiplication.

For the deflation functor $\mathrm{Def}_{G/N}^{G}$, the set U is the set G/N, for its left G/N-action and right G-action.

There is still another operation, involving the case of conjugation. It is associated to a group isomorphism $f : G \to H$. Any G-set X can be viewed as an H-set on which $h \in H$ acts as $f^{-1}(h) \in G$. Here the set U is the set G, with right G-action by multiplication, and left H-action of $h \in H$ by left multiplication by $f^{-1}(h)$. This operation is denoted by Iso_G^H (without reference to f, which is generally clear from the context).

NOTATION 6.1.3. Let G and H be finite groups, and L be a subgroup of $H \times G$. Set

$$\begin{aligned} p_1(L) &= \{h \in H \mid \exists g \in G,\ (h,g) \in L\},\\ p_2(L) &= \{g \in G \mid \exists h \in H,\ (h,g) \in L\},\\ k_1(L) &= \{h \in H \mid (h,1) \in L\}, \qquad k_2(L) = \{g \in G \mid (1,g) \in L\}. \end{aligned}$$

Then $k_i(L)$ is a normal subgroup of $p_i(L)$, for $i = 1, 2$. The quotient group is denoted by $q_i(L)$, for $i = 1, 2$. There is a canonical group isomorphism $q_1(L) \to q_2(L)$.

The quotient set $(H \times G)/L$ is viewed as an H-set-G by

$$h.(x, y)L.g = (hx, g^{-1}y)L \quad \forall h, x \in H,\ \forall y, g \in G.$$

This biset is transitive, i.e. it is a transitive $(H \times G^{\mathrm{op}})$-set. Conversely, any transitive H-set-G is isomorphic to a biset $(H \times G)/L$, for some subgroup L of $H \times G$. Any H-set-G is a disjoint union of transitive bisets.

If G, H, and K are finite groups, if L is a subgroup of $H \times G$ and M is a subgroup of $K \times H$, set

$$M * L = \{(k, g) \in K \times G \mid \exists h \in H,\ (k, h) \in M \text{ and } (h, g) \in L\}.$$

It is a subgroup of $K \times G$.

PROPOSITION 6.1.4. (1) *Let G and H be finite groups, and L be a subgroup of $H \times G$. Then there is an isomorphism of bisets*

$$(H \times G)/L \simeq H \times_{p_1(L)} q_1(L) \times_{q_1(L)} q_1(L) \times_{q_2(L)} q_2(L) \times_{p_2(L)} G.$$

(2) *Let G, H, and K be finite groups, let L be a subgroup of $H \times G$ and M be a subgroup of $K \times H$. Then there is an isomorphism of K-sets-G*

$$((K \times H)/M) \times_H ((H \times G)/L) \simeq \bigsqcup_{x \in [p_2(M)\backslash H/p_1(L)]} (K \times G)/(M * {}^{(x,1)}L).$$

PROOF. See [8, Lemme 3 and Proposition 1]. □

REMARK 6.1.5. Assertion (1) shows that any transitive biset $(H \times G)/L$ is a product of bisets associated to a restriction $\mathrm{Res}^G_{p_2(L)}$, a deflation $\mathrm{Def}^{p_2(L)}_{q_2(L)}$, an isomorphism $\mathrm{Iso}^{q_1(L)}_{q_2(L)}$, an inflation $\mathrm{Inf}^{p_1(L)}_{q_1(L)}$, and an induction $\mathrm{Ind}^H_{p_1(L)}$. Assertion (2) can be viewed as a biset version of the Mackey formula (see Proposition 2.2.1).

6.2. *Bisets and functors*

If G and H are finite groups, then any finite H-set-G U induces a functor $I_U : G\text{-}\mathbf{set} \to H\text{-}\mathbf{set}$ defined for a finite G-set X by $I_U(X) = U \times_G X$, and for a morphism of finite G-sets $f : X \to Y$ by

$$I_U(f)\big((u,_G x)\big) = \big(u,_G f(x)\big) : I_U(X) \to I_U(Y).$$

If K is another finite group, if V is a K-set-H, and if X is a finite G-set, there is a natural isomorphism of K-sets

$$V \times_H (U \times_G X) \to (V \times_H U) \times_G X$$

mapping $(v,_H (u,_G x))$ to $((v,_H u),_G x)$. This induces an isomorphism of functors between $I_V \circ I_U$ and $I_{V \times_H U}$.

The functor I_U preserves disjoint unions, hence it induces a morphism $\mathbb{B}(U)$ from $B(G)$ to $B(H)$ (this morphism should not be confused with the value $B(U)$ of the Burnside functor at the $(H \times G^{\mathrm{op}})$-set U!). If U' is an H-set-G, and if U' is isomorphic to U (as bisets), then the functors I_U and $I_{U'}$ are clearly isomorphic. Hence $\mathbb{B}(U) = \mathbb{B}(U')$.

With the same notation, one has $\mathbb{B}(V) \circ \mathbb{B}(U) = \mathbb{B}(V \times_H U)$. Moreover, if the biset U is a disjoint union $U_1 \sqcup U_2$ of two H-sets-G U_1 and U_2, then the functor I_U is clearly isomorphic to the disjoint union of the functors I_{U_1} and I_{U_2}. It follows that $\mathbb{B}(U) = \mathbb{B}(U_1) + \mathbb{B}(U_2)$.

Those remarks show that there is a well defined morphism from the Burnside group $B(H \times G^{\mathrm{op}})$ to the group $\mathrm{Hom}_{\mathbb{Z}}\big(B(G), B(H)\big)$ of group homomorphisms from $B(G)$ to $B(H)$.

If $G = H$, then viewing G as a G-set-G by left and right multiplication, it is clear that the functor I_G is isomorphic to the identity functor. Hence $\mathbb{B}(G)$ is the identity map of $B(G)$. The biset G will be called the *identity biset* for G.

There are many other situations where similar transformations can be associated to bisets: for instance, a finite H-set-G U induces a morphism $R(U) : R_{\mathbb{Q}}(G) \to R_{\mathbb{Q}}(H)$ between the corresponding Grothendieck groups of rational representations, defined for a finite dimensional $\mathbb{Q}G$-module M by

$$R(U)(M) = \mathbb{Q}U \otimes_{\mathbb{Q}G} M,$$

where $\mathbb{Q}U$ is the permutation bimodule with basis U. If K is a finite group, and if V is a finite K-set-H, then $R(V) \circ R(U) = R(V \times_H U)$, since there is an isomorphism of bimodules $\mathbb{Q}V \otimes_{\mathbb{Q}H} \mathbb{Q}U \simeq \mathbb{Q}(V \times_H U)$.

More generally, if k is a field, one can define a morphism $R(U) : R_k(G) \to R_k(H)$ by $R(U)(M) = kU \otimes_{kG} M$ for a finitely generated kG-module M. However, this map is well defined only if tensoring with kU preserves exact sequences, i.e. if kU is projective as right kG-module. If k has characteristic p, this is equivalent to requiring that for each $u \in U$, the stabilizer in G of u is a p'-group.

This shows that in some natural situations, not all the bisets are allowed, and leads to the following definition:

DEFINITION 6.2.1. Let $\mathcal{P}$ and $\mathcal{Q}$ be non-empty classes of finite groups, closed under isomorphisms and taking sections and extensions. This means that if H isomorphic to an element of $\mathcal{P}$, then H is in $\mathcal{P}$, that if H is a subgroup or a quotient of an element of $\mathcal{P}$, then H is in $\mathcal{P}$, and conversely that if N is a normal subgroup of H such that N and H/N are in $\mathcal{P}$, then $H \in \mathcal{P}$.

If G and H are finite groups, then an H-set-G U is said to be *$\mathcal{P}$-free-$\mathcal{Q}$* if for any $u \in U$, the stabilizer of u in H is in $\mathcal{P}$, and the stabilizer of u in G is in $\mathcal{Q}$.

One denotes by $B_{\mathcal{P},\mathcal{Q}}(H, G)$ the subgroup of $B(H \times G^{\mathrm{op}})$ generated by the H-sets-G which are $\mathcal{P}$-free-$\mathcal{Q}$. If R is a commutative ring, then as usual $RB_{\mathcal{P},\mathcal{Q}}(H, G)$ denotes $R \otimes_{\mathbb{Z}} B_{\mathcal{P},\mathcal{Q}}(H, G)$.

LEMMA 6.2.2. *Let G, H, and K be finite groups. Let U be an H-set-G, and V be a K-set-H. If U and V are $\mathcal{P}$-free-$\mathcal{Q}$, then so is $V \times_H U$.*

PROOF. See [8, Lemme 4]. □

It follows that if G, H, and K are finite groups, there is a bilinear product

$$\times_H : B_{\mathcal{P},\mathcal{Q}}(K, H) \times B_{\mathcal{P},\mathcal{Q}}(H, G) \to B_{\mathcal{P},\mathcal{Q}}(K, G)$$

extending the product $(X, Y) \mapsto X \times_H Y$. It should be noted moreover that since $\mathcal{P}$ and $\mathcal{Q}$ are non-empty and closed under taking quotients, they both contain the trivial group. Since the identity bisets are left and right free, they are always $\mathcal{P}$-free-$\mathcal{Q}$.

DEFINITION 6.2.3. Denote by $\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ the category defined as follows:
- The objects of $\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ are the finite groups.
- If G and H are finite groups, then

$$\mathrm{Hom}_{\mathcal{C}_{\mathcal{P},\mathcal{Q}}}(G, H) = B_{\mathcal{P},\mathcal{Q}}(H, G).$$

- The composition of morphisms in $\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ is defined for finite groups G, H, and K by

$$v \circ u = v \times_H u \quad \forall u \in B_{\mathcal{P},\mathcal{Q}}(H, G),\ \forall v \in B_{\mathcal{P},\mathcal{Q}}(K, H).$$

The category $\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ is preadditive (see [33, Ch. 1.8]), i.e. the sets of morphisms in $\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ are Abelian groups, and moreover the composition of morphisms is left and right distributive with respect to addition of morphisms.

More generally, if R is a commutative ring, then the category $R\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ is the category obtained from $\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ by tensoring with R: the objects are the finite groups, but

$$\mathrm{Hom}_{R\mathcal{C}_{\mathcal{P},\mathcal{Q}}}(G, H) = R \otimes_{\mathbb{Z}} \mathrm{Hom}_{\mathcal{C}_{\mathcal{P},\mathcal{Q}}}(G, H).$$

Denote by $\mathcal{F}_{\mathcal{P},\mathcal{Q}}$ the category the objects of which are additive functors from $\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ to $\mathbb{Z}$-**Mod**, morphisms are natural transformations of functors, and composition of morphisms is composition of natural transformations. More generally, denote by $R\mathcal{F}_{\mathcal{P},\mathcal{Q}}$ the category of R-linear functors from $R\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ to R-**Mod**.

The category $R\mathcal{F}_{\mathcal{P},\mathcal{Q}}$ is Abelian. A sequence

$$0 \to L \to M \to N \to 0$$

of objects and morphisms in $R\mathcal{F}_{\mathcal{P},\mathcal{Q}}$ is exact if and only if its evaluation at any finite group is exact.

It is now clear that the correspondence sending a finite group G to its Burnside group $B(G)$, and a finite $\mathcal{P}$-free-$\mathcal{Q}$ H-set-G U, for finite groups G and H, to the map $\mathbb{B}(U)$, extends to an additive functor from $\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ to $\mathbb{Z}$-**Mod**. More generally, the correspondence sending a finite group G to $RB(G)$ and a biset U to the map $R\mathbb{B}(U): RB(G) \to RB(H)$ induced by $\mathbb{B}(U)$, defines an object of the category $R\mathcal{F}_{\mathcal{P},\mathcal{Q}}$. It is called the *Burnside biset-functor* with coefficients in R.

Similarly, the correspondence mapping a finite group G to $R_{\mathbb{Q}}(G)$ and a finite $\mathcal{P}$-free-$\mathcal{Q}$ H-set-G U to the map $R(U): R_{\mathbb{Q}}(G) \to R_{\mathbb{Q}}(H)$, extends to an object $R_{\mathbb{Q}}$ of $\mathcal{F}_{\mathcal{P},\mathcal{Q}}$.

For any finite group G, the set $\mathrm{Hom}_{R\mathcal{C}_{\mathcal{P},\mathcal{Q}}}(\mathbb{1}, G)$ identifies with the R-submodule $RB_{\mathcal{P}}(G)$ of $RB(G)$ generated by finite $\mathcal{P}$-free G-sets, or, equivalently, by sets G/P, for $P \in \mathcal{P}$. This correspondence $G \mapsto RB_{\mathcal{P}}(G)$ is clearly a subfunctor of the biset functor RB.

The analogue of Proposition 5.6.1 is the following:

PROPOSITION 6.2.4. *Let M be any object of $R\mathcal{F}_{\mathcal{P},\mathcal{Q}}$. Then the map*

$$f \in \mathrm{Hom}_{R\mathcal{F}_{\mathcal{P},\mathcal{Q}}}(RB_{\mathcal{P}}, M) \mapsto f_{\mathbb{1}}(\bullet) \in M(\mathbb{1})$$

is an isomorphism of R-modules. In particular, the functor $RB_{\mathcal{P}}$ is a projective object in $R\mathcal{F}_{\mathcal{P},\mathcal{Q}}$, and $\mathrm{End}_{R\mathcal{F}_{\mathcal{P},\mathcal{Q}}}(RB_{\mathcal{P}}) \simeq R$.

PROOF. This is essentially Yoneda's Lemma: let G be any finite group, and X be any finite $\mathcal{P}$-free G-set. Then X is also a G-set-$\mathbb{1}$, for the right (trivial!) action of the trivial group, and moreover the G-set X is isomorphic to $X \times_{\mathbb{1}} \bullet$, where $\bullet$ is the trivial set for the trivial group.

Hence if $f: RB_{\mathcal{P}} \to M$ is a natural transformation, the map $f_G: RB_{\mathcal{P}}(G) \to M(G)$ is the map of R-modules sending the G-set X to $M(X)(f_{\mathbb{1}}(\bullet))$. Conversely, if $m \in M(\mathbb{1})$ is given, one can define a map f_G from $RB_{\mathcal{P}}(G)$ to $M(G)$ by setting

$$f_G(X) = M(X)(m)$$

for a finite $\mathcal{P}$-free G-set X, and extending f_G by linearity. It is easy to check that this defines a natural transformation from $RB_{\mathcal{P}}$ to M. □

When the ring R is a field, and $\mathcal{P}$ and $\mathcal{Q}$ are both equal to the family $\mathcal{A}ll$ of all finite groups, one can say a little more (see [8, Lemme 1 and Proposition 8]):

THEOREM 6.2.5 (Bouc). *Let K be a field. Then the functor KB is a projective object in $K\mathcal{F}_{\mathcal{A}ll,\mathcal{A}ll}$, with a unique maximal proper subfunctor J, defined for a finite group G by*

$$J(G) = \{X \in KB(G) \mid \forall Y \in KB(G),\ |G\backslash(Y \times X)| = 0_K\}.$$

If K has characteristic 0, the quotient functor KB/J is isomorphic to the functor $KR_{\mathbb{Q}}$. It is a simple object in $K\mathcal{F}_{All,All}$.

PROOF. If L is a subfunctor of KB, then $L(\mathbb{1})$ is a K-subspace of $KB(\mathbb{1}) = K$. Thus either $L(\mathbb{1}) = K$. In this case $\bullet \in L(\mathbb{1})$, thus for any finite group G and any finite G-set X, the G-set $X = X \times_{\mathbb{1}} \bullet = L(X)(\bullet)$ is in $L(G)$. Hence $L = KB$ in this case.

The other case is $L(1) = 0$, and then for any finite group G, any $X \in L(G)$, and any morphism Y from G to $\mathbb{1}$ in $K\mathcal{C}_{All,All}$, the image $L(Y)(X)$ has to be zero. But clearly $\mathrm{Hom}_{K\mathcal{C}_{All,All}}(G, \mathbb{1}) \simeq KB(G)$, and with this identification, the element $L(Y)(X)$ of $B(\mathbb{1})$ identifies with $|G\backslash(Y \times X)|$. Thus $L \subseteq J$, and moreover J is clearly a subfunctor of KB.

This shows in particular that the dimension over K of $KB(G)/J(G)$ is equal to the rank of the bilinear form on $KB(G)$ with values in K defined by

$$(X, Y) \in \left(KB(G)\right)^2 \mapsto \langle X, Y\rangle = \left|G\backslash(Y \times X)\right|.$$

Now if $K \supseteq \mathbb{Q}$, the K-basis $\{e_H^G\}_{H\in[s_G]}$ of $KB(G)$ is orthogonal for this bilinear form, and moreover

$$\begin{aligned}\left\langle e_H^G, e_H^G\right\rangle &= \left|G\backslash e_H^G\right| = \frac{1}{|N_G(H)|}\sum_{K\subseteq H}|K|\mu(K, H)\\ &= \frac{1}{|N_G(H)|}\sum_{x\in H}\sum_{\langle x\rangle\subseteq K\subseteq H}\mu(K, H) = \frac{\phi_1(H)}{|N_G(H)|},\end{aligned}$$

where $\phi_1(H)$ is the number of elements $x \in H$ such that $\langle x\rangle = H$. This is non-zero if and only if H is cyclic.

Hence the dimension of $KB(G)/J(G)$ is equal to the number $c(G)$ of conjugacy classes of cyclic subgroups of G.

Now the natural morphism $KB(G) \to KR_{\mathbb{Q}}(G)$ mapping a G-set to its permutation module induces clearly a natural transformation of functors

$$\chi : KB \to KR_{\mathbb{Q}} \tag{6.2.6}$$

since for any finite group H and any H-set-G U, there is an isomorphism of $\mathbb{Q}G$-modules $\mathbb{Q}(U \times_G X) \simeq \mathbb{Q}U \otimes_{\mathbb{Q}G} \mathbb{Q}X$.

Since χ is non-zero, the kernel of χ is contained in J. Hence the image of $KB(G)$ in $KR_{\mathbb{Q}}(G)$ has dimension at least equal to $c(G)$. But it is well known (see [42, Ch. 13, Théorème 29, Corollaire 1]) that the dimension of $KR_{\mathbb{Q}}(G)$ is equal to $c(G)$. The theorem follows. □

REMARK 6.2.7 (see [8, Proposition 2]). The set of isomorphism classes of simple objects in $R\mathcal{F}_{\mathcal{P},\mathcal{Q}}$ is in one to one correspondence with the set of isomorphism classes of pairs (H, V), where H is a finite group, and V is a simple $R\mathrm{Out}(H)$-module, where $\mathrm{Out}(H)$ is the group of outer automorphisms of H. The simple functor associated to such a pair (H, V) is denoted by $S_{H,V}$.

For example, when K is a field of characteristic zero, and $\mathcal{P} = \mathcal{Q} = \mathcal{A}ll$, the functor $KR_{\mathbb{Q}}$ of Proposition 6.2.5 is the functor $S_{\mathbb{1},K}$. In this case, the subfunctors of KB are studied in [8]. In particular, the composition factors (i.e. the simple sections) of KB are simple functors $S_{H,K}$ associated to the trivial $K\mathrm{Out}(H)$-module, for a class of finite group called *b-groups* ([8, Proposition 10]).

REMARK 6.2.8. In the case where $\mathcal{P}$ and $\mathcal{Q}$ are reduced to the trivial group, the objects of $R\mathcal{F}_{\mathcal{P},\mathcal{Q}}$ are called *global Mackey functors*. They have been studied by Webb ([52]). The reason for this denomination is that the operations associated to bisets which are both left and right free only involve induction, restriction and group isomorphisms.

In this paper, Webb also considers the *inflation functors*, which correspond to the case where $\mathcal{Q}$ is reduced to the trivial group, and $\mathcal{P}$ is the class of all finite groups.

REMARK 6.2.9. It is also of interest to consider subcategories of the above categories $R\mathcal{C}_{\mathcal{P},\mathcal{Q}}$. For example, let p be a prime number, and $\mathcal{D}_p$ denote the full subcategory of $\mathbb{Q}\mathcal{C}_{\mathcal{A}ll,\mathcal{A}ll}$ consisting of finite p-groups. Denote by $\mathbb{Q}B$ and $\mathbb{Q}R_{\mathbb{Q}}$ the restriction to $\mathcal{D}_p$ of the functors $\mathbb{Q}B$ and $\mathbb{Q}R_{\mathbb{Q}}$ on $\mathbb{Q}\mathcal{C}_{\mathcal{A}ll,\mathcal{A}ll}$.

It has been shown recently (Bouc and Thévenaz [11]) that the (restriction to $\mathcal{D}_p$ of the) kernel of the morphism χ defined in (6.2.6) above is isomorphic to the *torsion-free Dade functor* $\mathbb{Q}D$, which value at a p-group is equal to $\mathbb{Q} \otimes_{\mathbb{Z}} D(P)$, where $D(P)$ is the Dade group of endo-permutation kP-modules, for a field k of characteristic p. This functor $\mathbb{Q}D$ is simple, isomorphic to (the restriction to $\mathcal{D}_p$ of) $S_{E_2,K}$, where E_2 is an elementary Abelian p-group of order p^2.

6.3. *Double Burnside rings*

The endomorphism ring of a finite group G in the category $R\mathcal{C}_{\mathcal{P},\mathcal{Q}}$ is called the *double Burnside ring* of G with coefficients in R, for the classes $\mathcal{P}$ and $\mathcal{Q}$.

An essential tool to study those rings is the following lemma:

LEMMA 6.3.1. *Let G be a finite group. If X is a finite G-set, denote by $\tilde{X}$ the set $G \times X$, with its G-set-G structure defined by*

$$g(a,x)g' = (gag', gx) \quad \forall g, a, g' \in G \text{ and } x \in X.$$

Then the correspondence $X \mapsto \tilde{X}$ induces a morphism of rings (with unit) from $B(G)$ to $B_{\mathcal{P},\mathcal{Q}}(G,G)$.

PROOF. See [8, Lemme 13]. □

In particular, this lemma gives a way to carry the idempotents of the Burnside ring to the double Burnside ring, giving information on the projective and simple modules for those rings. This was used intensively in [8].

REMARK 6.3.2. If K is a field of characteristic 0, then the Burnside algebra $KB(G)$ is semi-simple for any finite group G. It is natural to ask if the *double* Burnside algebra $KB_{\mathcal{P},\mathcal{Q}}(G,G)$ is. The answer is no in general. One can show for instance that if $\mathcal{P}=\mathcal{Q}=\mathcal{A}ll$ is the class of all finite groups, then the algebra $KB_{\mathcal{P},\mathcal{Q}}(G,G)$ is semi-simple if and only if the group G is *cyclic*.

6.4. *Stable maps between classifying spaces*

Double Burnside rings have been studied intensively in the case where $\mathcal{P}=\{1\!\!1\}$ is reduced to the trivial group, and $\mathcal{Q}=\mathcal{A}ll$ is the class of all finite groups, because they are an essential tool to describe the *stable splitting* of the classifying spaces of finite groups. The origin of this theory is the Segal conjecture, proved by Carlsson ([16]), which states an isomorphism between the stable cohomotopy groups of the classifying space BG of a finite group G and the completion of the Burnside ring at the augmentation ideal.

Recall first some notation and definitions (see [2, Ch. 2.8]):

NOTATION 6.4.1. If X is a pointed CW-complex, denote by SX its (reduced) suspension, and by ΩX its loop space. If X and Y are pointed CW-complexes, denote by $[X;Y]$ the set of homotopy classes of pointed continuous maps from X to Y. If $m\in\mathbb{N}$, there is a map

$$\Omega^m S^m X=\left[S^m;S^m X\right]\to\left[S^{m+1};S^{m+1}X\right]=\Omega^{m+1}S^{m+1}X$$

defined by suspension, and one can set

$$\Omega^\infty S^\infty X=\varinjlim_{m\to\infty}\Omega^m S^m X.$$

If X and Y are pointed CW-complexes, then the set $\{X;Y\}$ of stable maps from X to Y is defined by

$$\{X;Y\}=\left[X;\Omega^\infty S^\infty Y\right].$$

It is an Abelian group. If X, Y, and Z are pointed CW-complexes, there is a bilinear map

$$\{Y;Z\}\times\{X;Y\}\to\{X;Z\}.$$

The *stable cohomotopy groups* $\pi_s^r(X)$ of a pointed CW-complex X are defined for $r\in\mathbb{Z}$ by

$$\pi_s^r(X)=\left\{X_+;S^r\right\}=\left[X_+;\varinjlim_{n\to\infty}\Omega^n S^{n+r}\right],$$

where X_+ is the space X with a disjoint basepoint. Note that this last expression makes sense even for $r<0$.

The other definition required in the Segal conjecture is the notion of completion. Observe that another consequence of Lemma 6.3.1 is that if G and H are finite groups, then the natural structure of $B_{\mathcal{P},\mathcal{Q}}(H,H)$-module-$B_{\mathcal{P},\mathcal{Q}}(G,G)$ on $B_{\mathcal{P},\mathcal{Q}}(H,G)$ gives by restriction a structure of $B(H)$-module-$B(G)$.

It is easy to check (see [8, Lemme 15]) that if X is an H-set-G and Y is a G-set, then $X \times_G \tilde{Y}$ identifies with the Cartesian product $X \times Y$, for the double action given by

$$h(x,y)g = \left(hxg, g^{-1}y\right) \quad \forall h \in H,\ \forall g \in G,\ \forall x \in X,\ \forall y \in Y.$$

NOTATION 6.4.2. Let G be a finite group. Denote by I_G the prime ideal $I_{\mathbb{I},0}(G)$ of $B(G)$ defined in Section 3.4. It is called the *augmentation ideal*. The completion of $B(G)$ at the ideal I_G is the inverse limit

$$B(G)\hat{} = \varprojlim_{n\in\mathbb{N}} B(G)/I_G^n.$$

More generally, if H and G are finite groups, the *completion* $B_{\{\mathbb{I}\},\mathcal{A}ll}(H,G)\hat{}$ of $B_{\{\mathbb{I}\},\mathcal{A}ll}(H,G)$ is defined as the inverse limit

$$B_{\{\mathbb{I}\},\mathcal{A}ll}(H,G)\hat{} = \varprojlim_{n\in\mathbb{N}} B_{\{\mathbb{I}\},\mathcal{A}ll}(H,G)/B_{\{\mathbb{I}\},\mathcal{A}ll}(H,G)I_G^n,$$

where $B_{\{\mathbb{I}\},\mathcal{A}ll}(H,G)I_G^n$ denotes $B_{\{\mathbb{I}\},\mathcal{A}ll}(H,G) \times_G \widetilde{I_G^n}$.

If Y is a pointed CW-complex, then the functors $X \mapsto \{X; S^r Y\}$ form a *generalized cohomology theory* (see [2, Ch. 2.5]). In particular, if H is a subgroup of a finite group G, there is a *transfer map* from $\{BH_+; Y\}$ to $\{BG_+; Y\}$. Taking $Y = BH_+$, the image of the identity of $\{BH_+; BH_+\}$ is an element Tr_H^G of $\{BG_+; BH_+\}$, also called transfer. This element can be composed with the stable map from BH_+ to S^0 obtained by identifying BH to a point. This gives an element τ_H of $\{BG_+, S^0\} = \pi_s^0(BG)$.

The precise statement of Segal's conjecture is now the following:

THEOREM 6.4.3 (Carlsson). *Let G be a finite group. Then the map $G/H \mapsto \tau_H$ defined above induces a group isomorphism*

$$B(G)\hat{} \simeq \pi_s^0(BG).$$

Furthermore $\pi_s^r(BG) = 0$ for $r > 0$.

This has been generalized by Lewis, May and McClure ([30]): suppose H and G are finite groups. If Q is a subgroup of G and $\phi: Q \to H$ is a group homomorphism, then the stable map $\mathrm{Tr}_Q^G \in \{BG_+; BQ_+\}$ can be composed with the element of $\{BQ_+; BH_+\}$ deduced from ϕ, to get an element $\tau_{Q,\phi} \in \{BG_+; BH_+\}$.

This element depends only on the conjugacy class of the subgroup $\Delta_\phi(Q)$ of $H \times G$ consisting of the elements $(\phi(l), l)$, for $l \in Q$. Moreover, the biset $(H \times G)/\Delta_\phi(Q)$ is transitive, and free on the left. Conversely, any left-free transitive H-set-G is isomorphic to a biset $(H \times G)/\Delta_\phi(Q)$, for some subgroup Q of G and some morphism ϕ from Q to H.

This construction gives all the stable maps from BG_+ to BH_+. More precisely:

THEOREM 6.4.4 (Lewis–May–McClure). *Let G and H be finite groups. The above correspondence sending the left-free transitive H-set-G $(H \times G)/\Delta_\phi(Q)$ to $\tau_{Q,\phi} \in \{BG_+; BH_+\}$ induces an isomorphism*

$$B_{\{\mathbb{1}\},\mathcal{A}ll}(H,G)\hat{} \to \{BG_+; BH_+\}.$$

This theorem was at the origin of the study of *stable splittings* of the classifying spaces of finite groups. It provides indeed an algebraic translation from the stable endomorphisms of BG_+ for a finite group G, in terms of the completion of the double Burnside ring $B_{\{\mathbb{1}\},\mathcal{A}ll}(G,G)$. For details see the survey paper of Benson ([3]), or the articles of Benson and Feschbach ([4]), Martino and Priddy ([34]), Priddy ([37]), Webb ([52]).

Those two deep theorems on stable homotopy raise some natural questions on the algebraic side. For example, being given three finite groups G, H, and K, the composition of stable maps

$$\{BH_+; BK_+\} \times \{BG_+; BH_+\} \to \{BG_+; BK_+\}$$

shows that there is a map

$$B_{\{\mathbb{1}\},\mathcal{A}ll}(K,H)\hat{}\times B_{\{\mathbb{1}\},\mathcal{A}ll}(H,G)\hat{} \to B_{\{\mathbb{1}\},\mathcal{A}ll}(K,G)\hat{}$$

induced by the product $(X,Y) \mapsto X \times_H Y$. Without Theorem 6.4.4, the existence of such a map is not obvious a priori, and seems to require the following lemma:

LEMMA 6.4.5. *Let G and H be finite groups. Then for any integer $m \in \mathbb{N}$, there exists an integer $n \in \mathbb{N}$ such that*

$$I_H^n B_{\{\mathbb{1}\},\mathcal{A}ll}(H,G) \subseteq B_{\{\mathbb{1}\},\mathcal{A}ll}(H,G) I_G^m,$$

where $I_H^n B_{\{\mathbb{1}\},\mathcal{A}ll}(H,G)$ denotes $\widetilde{I_H^n} \times_H B_{\{\mathbb{1}\},\mathcal{A}ll}(H,G)$.

PROOF. Let X be a finite H-set, and Y be a finite H-set-G. Then it is easy to see that the biset $Z = \tilde{X} \times_H Y$ identifies with the Cartesian product $X \times Y$, for the double action given by

$$h(x,y)g = (hx, hyg) \quad \forall h \in H,\ \forall g \in G,\ \forall x \in X,\ \forall y \in Y.$$

In particular for any subgroup L of $H \times G$, the fixed points of L on Z, i.e. the set of $z \in Z$ such that $az = zb$ whenever $(a,b) \in L$, identifies with $X^{p_1(L)} \times Y^L$. It follows that for any $X \in B(H)$ and any $Y \in B(H \times G^{\mathrm{op}})$

$$\left|(\tilde{X} \times_H Y)^L\right| = \left|X^{p_1(L)}\right| \left|Y^L\right|.$$

In particular if $X \in I_G^n$, for an integer $n \geqslant 1$, then X is a sum of terms $X_1 \times \cdots \times X_n$, with $X_1, \ldots, X_n \in I_G$. Hence $|X| = 0$, and moreover $|X^P| \equiv 0\ (p^n)$ for any prime p and any p-subgroup P of H, since $|X_i^P| \equiv |X_i| = 0\ (p)$, for $i = 1, \ldots, n$. Thus

$$I_H^n B_{\{\mathbb{1}\},\mathcal{A}ll}(H, G) \subseteq J_n(H, G),$$

where $J_n(H, G)$ is the subset of $B(H, G) = B_{\{\mathbb{1}\},\mathcal{A}ll}(H, G)$ defined by

$$\begin{aligned} J_n(H, G) = \{Z \mid |Z| = 0,\ \forall p \text{ prime},\\ \forall L \subseteq H \times G,\ p_1(L)\ p\text{-group} \Rightarrow \left|Z^L\right| \equiv 0\,(p^n)\}. \end{aligned}$$

For each prime number p, set $m_p = |H|_{p'}|G|_{p'}$. Then for any p-perfect subgroup K of G, the idempotent

$$f_{p,K}^G = \sum_{\substack{M \in [s_G] \\ O^p(M) =_G K}} e_M^G$$

of $\mathbb{Q}B(G)$ lies in $\mathbb{Z}_{(p)}B(G)$, and $m_p f_K^G \in B(G)$. Moreover $|e_M^G| = 0$ unless $M = \mathbb{1}$. Thus $m_p f_{p,K}^G \in I_G$ unless $K = \mathbb{1}$.

It follows that for any integer n, the element $m_p^n f_{p,K}^G = (m_p f_{p,K}^G)^n$ is in I_G^n if $K \neq \mathbb{1}$. Now if $Z \in B(H, G)$

$$m_p^n Z = m_p^n Z f_{p,\mathbb{1}}^G + \sum_{\substack{K \in [s_G] \\ O^p(K) = K}} m_p^n Z f_{p,K}^G \equiv m_p^n Z f_{p,\mathbb{1}}^G \quad \left(B(H, G) I_G^n\right).$$

Moreover, the integers m_p, for p dividing $|H||G|$, are relatively prime, and there exist integers $a_{p,n}$, for p prime (equal to 0 for p not dividing $|H||G|$), such that $\sum_p a_{p,n} m_p^n = 1$. It follows that

$$Z \equiv \sum_p a_{p,n} Y_p \quad \left(B(H, G) I_G^n\right),$$

where $Y_p = m_p^n Z f_{p,\mathbb{1}}^G$.

Let L be a subgroup of $H \times G$, and M be a subgroup of G. Then in $B(H \times G)$

$$e_L^{H \times G} \times e_M^G = \left|\left(e_M^G\right)^L\right| e_L^{H \times G} = \left|\left(e_M^G\right)^{p_2(L)}\right| e_L^{H \times G} = \begin{cases} e_L^{H \times G} & \text{if } p_2(L) =_G M, \\ 0 & \text{otherwise.} \end{cases}$$

Now for each prime p, the element Y_p is equal to

$$Y_p = Y_p f_{p,\mathbb{1}}^G = \sum_{L \in [s_{H \times G}]} \left|Y_p^L\right| e_L^{H \times G}$$

and $e_L^{H\times G} f_{p,\mathbb{I}}^G = 0$ unless $p_2(L)$ is a p-group. Thus

$$Y_p = \sum_{\substack{L\in[s_{H\times G}]\\ O^p(p_2(L))=\mathbb{I}}} m_p^n |Z^L| e_L^{H\times G}.$$

Moreover, for each subgroup L of $H\times G$, one has $|L| = |k_1(L)||p_2(L)|$, and if $k_1(L)\neq \mathbb{I}$, then $|Z^L| = 0$ since Z is a linear combination of left-free bisets. It follows that if $|Z^L|\neq 0$, and if $p_2(L)$ is a p-group, then L is a p-group, equal to $\Delta_\phi(Q)$ for some p-subgroup Q of G, and some morphism $\phi: Q\to H$. Hence $p_1(L) = \phi(Q)$ is a p-group. Thus

$$Y_p = \sum_L m_p^n |Z^L| e_L^{H\times G},$$

where L runs through those p-subgroups $\Delta_\phi(Q)$, up to conjugation by $H\times G$. Finally, for such a subgroup $L = \Delta_\phi(Q)$

$$\begin{aligned} m_p^n |Z^L| e_L^{H\times G} &= |H|_{p'}^n |G|_{p'}^n |Z^L| e_L^{H\times G} e_Q^G \\ &= |H|_{p'}^{n-1} |G|_{p'}^{n-m-1} \frac{|Z^L|}{|H|_p |G|_p^{m+1}} \big(|H||G| e_L^{H\times G}\big)\big(|G|^m e_Q^G\big). \end{aligned}$$

This is zero if $Q = \mathbb{I}$ and $Z\in J_n(H,G)$, because $|Z^L| = |Z| = 0$ in this case. And if $Q\neq \mathbb{I}$, the element $(|H||G|e_L^{H\times G})(|G|^m e_Q^G)$ is in $B(H,G)I_G^m$, and if n is big enough, the quotient $\frac{|Z^L|}{|H|_p|G|_p^{m+1}}$ is an integer for $Z\in J_n(H,G)$, since $p_1(L) = \phi(Q)$ is a p-group. This completes the proof of the lemma. □

PROPOSITION 6.4.6. *Let G, H and K be finite groups. Then the product*

$$(X,Y)\in B_{\{\mathbb{I}\},\mathcal{All}}(K,H)\times B_{\{\mathbb{I}\},\mathcal{All}}(H,G)\mapsto X\times_H Y\in B_{\{\mathbb{I}\},\mathcal{All}}(K,G)$$

induces a well defined product

$$B_{\{\mathbb{I}\},\mathcal{All}}(K,H)\hat{}\times B_{\{\mathbb{I}\},\mathcal{All}}(H,G)\hat{}\to B_{\{\mathbb{I}\},\mathcal{All}}(K,G)\hat{}.$$

PROOF. This is clear, since by the previous lemma, the product

$$(X,Y)\in B_{\{\mathbb{I}\},\mathcal{All}}(K,H)\times B_{\{\mathbb{I}\},\mathcal{All}}(H,G)\mapsto X\times_H Y\in B_{\{\mathbb{I}\},\mathcal{All}}(K,G)$$

is continuous for the I_H-adic and I_G-adic topologies. □

PROPOSITION 6.4.7. *Let G be a finite group. Then the correspondence mapping a subgroup H of G to $B(H)\hat{}$ has a natural structure of Green functor.*

PROOF. Indeed, for any finite group H, there is an isomorphism

$$B(H)\hat{} \simeq B_{\{\mathbb{I}\},\mathcal{A}ll}(\mathbb{I}, H)\hat{}.$$

Thus for two finite groups H and K, there are maps

$$B(H)\hat{} \times B_{\{\mathbb{I}\},\mathcal{A}ll}(H, K) \to B(K)\hat{}.$$

Now if $H \subseteq K$ are subgroups of G, this gives a transfer map from $B(H)\hat{}$ to $B(K)\hat{}$, using K as a left-free H-set-K, and a restriction map from $B(K)\hat{}$ to $B(H)\hat{}$, using K as a left-free K-set-H. Similarly, for $x \in G$, there is a conjugation map from $B(H)\hat{}$ to $B({}^x H)\hat{}$, induced by the biset H with its obvious structure of H-set-${}^x H$. The axioms of Mackey functors follow from the properties of product of bisets.

In other words, the problem to define a Mackey functor structure on $B\hat{}$ is that the correspondence $H \mapsto I_H^n$ is not clearly a sub-Mackey functor of B. However if H is a subgroup of G, and n a positive integer, let

$$J_n(H) = \{X \in B(H) \mid |X| = 0,\ \forall p \text{ prime},\ \forall P \subseteq H,\ P\ p\text{-group}, \\ |X^P| \equiv 0\ (p^n)\}.$$

Then $J_n(H)$ is an ideal of $B(H)$, and one can show as in Lemma 6.4.5 that the topology defined by the ideals $J_n(H)$ is equivalent to the topology defined by the ideals I_H^n. It is easy to see moreover that the correspondence $H \mapsto J_n(H)$ is a sub-Mackey functor of B. Hence the completion $B(H)\hat{}$ is isomorphic to

$$B(H)\hat{} \simeq \varprojlim_{n\to\infty} B(H)/J_n(H)$$

and in this form, it has a natural structure of Green functor. □

REMARK 6.4.8. For further results on completion of Green functors, see Luca ([32]).

References

[1] D. Benson, *Representations and Cohomology I*, Cambridge Stud. Adv. Math., Vol. 30, Cambridge University Press (1991).

[2] D. Benson, *Representations and Cohomology II*, Cambridge Stud. Adv. Math., Vol. 31, Cambridge University Press (1991).

[3] D. Benson, *Stably splitting BG*, Bull. Amer. Math. Soc. **33** (1996), 189–198.

[4] D. Benson and M. Feschbach, *Stable splittings of classifying spaces of finite groups*, Topology **31** (1992), 157–176.

[5] S. Bouc, *Projecteurs dans l'anneau de Burnside, projecteurs dans l'anneau de Green, et modules de Steinberg généralisés*, J. Algebra **139** (2) (1991), 395–445.

[6] S. Bouc, *Exponentielle et modules de Steinberg*, J. Algebra **150** (1) (1992), 118–157.

[7] S. Bouc, *Construction de foncteurs entre catégories de G-ensembles*, J. Algebra **183** (0239) (1996), 737–825.

[8] S. Bouc, *Foncteurs d'ensembles munis d'une double action*, J. Algebra, **183** (0238) (1996), 664–736.
[9] S. Bouc, *Green-Functors and G-Sets*, Lecture Notes in Math., Vol. 1671, Springer, Berlin (October 1997).
[10] S. Bouc, *Résolutions de foncteurs de Mackey*, Group Representations: Cohomology, Group Actions and Topology, AMS Proc. Sympos. Pure Math., Vol. 63 (January 1998), 31–83.
[11] S. Bouc and J. Thévenaz, *The group of endo-permutation modules*, Invent. Math. (1998), to appear.
[12] M. Broué, *On Scott modules and p-permutation modules: An approach through the Brauer morphism*, Proc. Amer. Math. Soc. **93** (3) (1985), 401–408.
[13] K. Brown, *Euler characteristics of groups: The p-fractional part*, Invent. Math. **29** (1975), 1–5.
[14] K. Brown and J. Thévenaz, *A generalization of Sylow's third theorem*, J. Algebra **115** (1988), 414–430.
[15] W. Burnside, *Theory of Groups of Finite Order*, 2nd edn., Cambridge University Press (1911).
[16] G. Carlsson, *Equivariant stable homotopy and Segal's Burnside ring conjecture*, Ann. of Math. **120** (1984), 189–224.
[17] S.B. Conlon, *Decompositions induced from the Burnside algebra*, J. Algebra **10** (1968), 102–122.
[18] C. Curtis and I. Reiner, *Methods of Representation Theory with Applications to Finite Groups and Orders*, Wiley Classics Library, Vol. II, Wiley (1990).
[19] T. tom Dieck, *The Burnside ring of a compact Lie group*, Math. Ann. **215** (1975), 235–250.
[20] T. tom Dieck, *Idempotent elements in the Burnside ring*, J. Pure Appl. Algebra **10** (1977), 239–247.
[21] T. tom Dieck, *Transformation Groups and Representation Theory*, Lecture Notes in Math., Vol. 766, Springer, Berlin (1979).
[22] T. tom Dieck, *Transformation Groups*, Studies in Math., Vol. 8, de Gruyter (1987).
[23] A. Dress, *A characterization of solvable groups*, Math. Z. **110** (1969), 213–217.
[24] A. Dress, *Notes on the theory of representations of finite groups*, Bielefeld Notes (1971).
[25] A. Dress, *Contributions to the Theory of Induced Representations*, Lecture Notes in Math., Vol. 342, Springer, Berlin (1973), 183–240.
[26] D. Gluck, *Idempotent formula for the Burnside ring with applications to the p-subgroup simplicial complex*, Illinois J. Math. **25** (1981), 63–67.
[27] J. Green, *Axiomatic representation theory for finite groups*, J. Pure Appl. Algebra **1** (1971), 41–77.
[28] G. Karpilovsky, *Group Representations 4*, Math. Studies, Vol. 182, North-Holland (1995).
[29] L.G. Lewis, *The category of Mackey functors for a compact Lie group*, Group representations: Cohomology, Group Actions and Topology, AMS Proc. Symposia in Pure Math., Vol. 63 (January 1998).
[30] L.G. Lewis, J.P. May and J.E. McClure, *Classifying G-spaces and the Segal conjecture*, Current Trends in Algebraic Topology, Vol. 2, CMS Conference Proc. (1982), 165–179.
[31] H. Lindner, *A remark on Mackey functors*, Manuscripta Math. **18** (1976), 273–278.
[32] F. Luca, *The defect of completion of a Green functor*, Group Representations: Cohomology, Group Actions and Topology, AMS Proc. Symposia in Pure Math., Vol. 63 (January 1998), 369–381.
[33] S. Mac Lane, *Categories for the Working Mathematician*, Graduate Texts in Math., Vol. 5, Springer (1971).
[34] J. Martino and S.B. Priddy, *The complete stable splitting for the classifying space of a finite group*, Topology **31** (1992), 143–156.
[35] T. Matsuda, *On the unit group of Burnside rings*, Japan. J. Math. **8** (1) (1982), 71–93.
[36] T. Matsuda and T. Miyata, *On the unit groups of the Burnside rings of finite groups*, J. Math. Soc. Japan **35** (1983), 345–354.
[37] S.B. Priddy, *On characterizing summands in the classifying space of a finite group I*, Amer. J. Math. **112** (1990), 737–748.
[38] D. Quillen, *Higher algebraic K-theory*, Lecture Notes in Math., Vol. 341, Springer, Berlin (1973), 85–147.
[39] D. Quillen, *Homotopy properties of the poset of non-trivial p-subgroups*, Adv. in Math. **28** (2) (1978), 101–128.
[40] R. Schwänzl, *Koeffiziente in Burnside rings*, Arch. Math. **29** (1977), 621–622.
[41] R. Schwänzl, *On the spectrum of the Burnside ring*, J. Pure Appl. Algebra **15** (1979), 181–185.
[42] J.-P. Serre, *Représentations Linéaires des Groupes Finis*, 2nd ed., Collection Méthodes, Hermann (1971).
[43] L. Solomon, *The Burnside algebra of a finite group*, J. Combin. Theory **2** (1967), 603–615.
[44] J. Thévenaz, *Idempotents de l'anneau de Burnside et caractéristique d'Euler*, Séminaire Claude Chevalley sur les groupes finis, Vol. 25, Tome 3, Publications mathématiques de l'Université Paris 7 (1985), 207–217.
[45] J. Thévenaz, *Permutation representation arising from simplicial complexes*, J. Combin. Theory Ser. A **46** (1987), 122–155.

[46] J. Thévenaz, *Isomorphic Burnside rings*, Comm. Algebra **16** (9) (1988), 1945–1947.
[47] J. Thévenaz and P. Webb, *Simple Mackey functors*, Proc. of the 2nd International Group Theory Conference Bressanone 1989, Rend. Circ. Mat. Palermo, Serie II, Vol. 23 (1990), 299–319.
[48] J. Thévenaz and P. Webb, *Homotopy equivalences of posets with a group action*, J. Combin. Theory **56** (2) (1991), 173–181.
[49] J. Thévenaz and P. Webb, *The structure of Mackey functors*, Trans. Amer. Math. Soc. **347** (6) (1995), 1865–1961.
[50] P. Webb, *A local method in group cohomology*, Comment. Math. Helv. **62** (1987), 135–167.
[51] P. Webb, *A split exact sequence for Mackey functors*, Comment. Math. Helv. **66** (1991), 34–69.
[52] P. Webb, *Two classification of simple Mackey functors with application to group cohomology and the decomposition of classifying spaces*, J. Pure Appl. Algebra **88** (1993), 265–304.
[53] T. Yoshida, *Idempotents in Burnside rings and Dress induction theorem*, J. Algebra **80** (1983), 90–105.
[54] T. Yoshida, *On the unit groups of Burnside rings*, J. Math. Soc. Japan **42** (1) (1990), 31–64.

A Guide to Mackey Functors

Peter Webb

School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA
E-mail: webb@math.umn.edu

Contents

HANDBOOK OF ALGEBRA, VOL. 2
Edited by M. Hazewinkel

This Page Intentionally Left Blank

1. Introduction

A Mackey functor is an algebraic structure possessing operations which behave like the induction, restriction and conjugation mappings in group representation theory. Operations such as these appear in quite a variety of diverse contexts – for example group cohomology, the algebraic K-theory of group rings, and algebraic number theory – and it is their widespread occurrence which motivates the study of such operations in abstract.

The axioms for a Mackey functor which we will use were first formulated by Dress [24], [25] and by Green [30]. They follow on from earlier ideas of Lam on Frobenius functors [36], described in [19]. Another structure which appeared early on is Bredon's notion of a coefficient system [15].

A major preoccupation in studying Mackey functors is to compute their values, be it in the context of specific examples such as computing the cohomology or character ring of a finite group, or in a more general setting. It is important to develop techniques to do this, and if some method of calculation can be formulated within the general context of Mackey functors then we have the possibility to apply it to every specific instance without developing it each time from scratch.

One argument which generalizes to Mackey functors in this way is the method of stable elements which appears in the book of Cartan and Eilenberg [16], and which provides a way of computing the p-torsion subgroup of the cohomology of a finite group as a specifically identified subset of the cohomology of a Sylow p-subgroup. An ingredient in the general form of this calculation is the notion of relative projectivity of a Mackey functor, similar in spirit to the notion of relative projectivity of group representations. We will see in Section 3 how the method of stable elements can be formulated for all Mackey functors.

Induction theorems are another kind of result which are among the most important methods of computation. They have a very well-developed and well-known theory, especially in the context of group representations. Such theorems may also be formulated in the general setting, and in Section 6 we present an important induction theorem due to Dress. Work of great refinement obtaining explicit forms of induction theorems has been done by Boltje [5–7], but this goes beyond what we can describe here.

In order to present these applications we develop the technical machinery which they necessitate, and we do this in an order which to a large extent reflects chronology. As the theory of Mackey functors became more elaborate it became apparent that they are algebraic structures in their own right with a theory which fits into the framework of representations of algebras. They may, in fact, be identified with the representations of a certain algebra – called the Mackey algebra – and there are simple Mackey functors, projective and injective Mackey functors, resolutions of Mackey functors, and so on. We describe this theory in outline in Section 5.

A new notion of Mackey functor began to appear, namely that of a globally-defined Mackey functor, an early instance of which appeared in the work of Symonds [52]. These are structures which have a definition on all finite groups (whereas the original Mackey functors are only defined on the subgroups of some fixed group) and we present in Section 8 the context for these structures envisaged by Bouc [10]. We describe three uses for these functors: a method of computing group cohomology in Section 9, an approach to the stable decomposition of classifying spaces BG in Section 10, and a framework in which Dade's group of endopermutation modules plays a fundamental role in Section 11.

There is no full account of Mackey functors in text book form, and with this in mind I have tried to be comprehensive in my treatment. In this I have failed, and on top of everything the proofs that are given are often sketchy or left to the reader who must either work them out as an exercise or consult the literature. This guide to Mackey functors is deliberately concise. The omissions which seem most regrettable are these: the definition of a Mackey functor on compact Lie groups and other more general classes of groups (see [38,21]); the theory of Green functors (see [59,12]); and the theory of Brauer quotients (see [59]).

Finally, I wish to thank Serge Bouc for his comments of this exposition.

2. The definitions of a Mackey functor

There are several ways of giving the definition of a Mackey functor, but they all amount to the same thing. We present two definitions here, the first in terms of many axioms and the second in terms of bivariant functors on the category of finite G-sets. They may also be defined as functors on a specially-constructed category, an approach which is due to Lindner [40].

The most accessible definition of a Mackey functor for a finite group G is expressed in terms of axiomatic relations. We fix a commutative ring R with a 1 and let R-mod denote the category of R-modules. A *Mackey functor* for G over R is a function

$$M : \{\text{subgroups of } G\} \to R\text{-mod}$$

with morphisms

$$\begin{aligned} I_K^H &: M(K) \to M(H), \\ R_K^H &: M(H) \to M(K), \\ c_g &: M(H) \to M({}^gH) \end{aligned}$$

for all subgroups H and K of G with $K \leqslant H$ and for all g in G, such that

(0) $I_H^H, R_H^H, c_h : M(H) \to M(H)$ are the identity morphisms for all subgroups H and $h \in H$,

$$\left.\begin{aligned} &(1)\ R_J^K R_K^H = R_J^H \\ &(2)\ I_K^H I_J^K = I_J^H \end{aligned}\right\} \text{ for all subgroups } J \leqslant K \leqslant H,$$

(3) $c_g c_h = c_{gh}$ for all $g, h \in G$,

$$\left.\begin{aligned} &(4)\ R_{{}^gK}^{{}^gH} c_g = c_g R_K^H \\ &(5)\ I_{{}^gK}^{{}^gH} c_g = c_g I_K^H \end{aligned}\right\} \text{ for all subgroups } K \leqslant H \text{ and } g \in G,$$

(6) $R_J^H I_K^H = \sum_{x \in [J\backslash H/K]} I_{J\cap {}^xK}^J c_x R_{J^x \cap K}^K$ for all subgroups $J, K \leqslant H$.

We use the letters I, R and c because these operations are reminiscent of induction, restriction and conjugation of characters. We should properly write $c_{g,H}$ instead of c_g, since our notation does not distinguish between the conjugation morphisms with the same element g but different domain $M(H)$. The most elaborate of these axioms is (6), which is called the *Mackey decomposition formula*, and is responsible for the name of these functors. It is familiar from representation theory and cohomology. In this axiom we are using the notation $[J\backslash H/K]$ to denote a set of representatives in G for the double cosets $J\backslash H/K$. We write ${}^xH = xHx^{-1}$ and $H^x = x^{-1}Hx$.

Mackey functors form a category denoted $\mathrm{Mack}_R(G)$ in which the morphisms are natural transformations of Mackey functors; that is, a morphism $\eta : M \to N$ is a family of R-module homomorphisms $\eta_H : M(H) \to N(H)$ commuting with all operations I, R and c. This category is Abelian, the reason being that R-mod is Abelian, and in fact we may define kernels, cokernels, subfunctors, quotient functors and so forth pointwise using the fact that they exist in R-mod. We may speak of the intersection of subfunctors of a Mackey functor, defined pointwise, and it is again a subfunctor. If we are given for each subgroup $H \leqslant G$ a subset $N(H) \subseteq M(H)$ we may speak of the subfunctor $\langle N \rangle$ generated by N: it is the intersection of the subfunctors containing N.

We will encounter also the notion of a *Green functor*, which is a Mackey functor M with an extra multiplicative structure. Specifically, for each subgroup $H \leqslant G$, $M(H)$ should be an associative R-algebra with identity so that

(7) the R^H_K and c_g are always unitary R-algebra homomorphisms, and
(8) for all subgroups $K \leqslant H$, $a \in M(K)$ and $b \in M(H)$ we have
$I^H_K(a \cdot R^H_K(b)) = I^H_K(a) \cdot b$ and $I^H_K(R^H_K(b) \cdot a) = b \cdot I^H_K(a)$.

Axiom (8) is called the Frobenius axiom. Green functors have in some ways a tighter structure than Mackey functors, but we will not describe their theory in detail here. Fuller accounts of recent theory may be found in [58,59] and [12].

We mention now some examples of Mackey functors that immediately come to mind. We may take $M(G)$ to be

- $G_0(kG)$: the Grothendieck group of the category of finitely generated kG-modules. In characteristic zero this may be identified as the group of characters of kG-modules, and in characteristic p as the group of Brauer characters.
- $A(G)$: the Green ring of finitely generated kG-modules [19, Section 81].
- $H^n(G, U)$, $H_n(G, U)$, $\hat{H}^n(G, U)$: the cohomology, homology and Tate cohomology of G in some dimension n with coefficients in the $\mathbb{Z}G$-module U.
- $B(G)$: the Burnside ring of G. We may identify this as the free Abelian group with the isomorphism types of transitive G-sets as a basis.
- $K_n(\mathbb{Z}G)$: the algebraic K-theory of $\mathbb{Z}G$, and other related groups such as the Whitehead group.
- $\mathrm{Cl}(\mathcal{O}(F^G))$: the class group of the ring of integers of the fixed field F^G where G is a group of automorphisms of a number field F (see [35,50,8]).

For some more examples see [59, Section 53].

In the first instance these examples are only Mackey functors over the ground ring $\mathbb{Z}$. If we have some other ground ring R in mind we may always form a Mackey functor $R \otimes_{\mathbb{Z}} M$ whose values are $R \otimes_{\mathbb{Z}} M(H)$ for each subgroup H of G. Some examples may already be naturally defined over a ring R other than $\mathbb{Z}$, for example Tate cohomology. Since $|G|$ annihilates Tate cohomology, this example gives a Mackey functor over $\mathbb{Z}/|G|\mathbb{Z}$; if the module U happens to be defined over some further ring R then $M(G) = H^n(G, U)$ is also a Mackey functor over R.

It is important to have available a different definition of Mackey functors, less dependent on a large number of axioms. It is phrased in terms of the category G-set whose objects are the finite left G-sets, and whose morphisms are the G-equivariant mappings. We will be especially interested in the space of left cosets G/H for each subgroup H of G: each G-set is isomorphic to a disjoint union of these. We may now define a Mackey functor over R to be a pair of functors $M = (M_*, M^*)$ from G-set to R-mod so that M_* is covariant, M^* is contravariant, $M_*(\Omega) = M^*(\Omega)$ for all finite G-sets Ω, and such that the following axioms are satisfied:

(1) for every pullback diagram of G-sets

$$\begin{array}{ccc} \Omega_1 & \xrightarrow{\alpha} & \Omega_2 \\ \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} \\ \Omega_3 & \xrightarrow{\delta} & \Omega_4 \end{array}$$

we have $M^*(\delta)M_*(\gamma) = M_*(\beta)M^*(\alpha)$, and

(2) for every pair of finite G-sets Ω and Ψ, applying M_* to $\Omega \to \Omega \sqcup \Psi \leftarrow \Psi$ gives the component maps in a morphism $M(\Omega) \oplus M(\Psi) \to M(\Omega \sqcup \Psi)$, which we require to be an isomorphism.

The definition we have just given is a special case of the definition given by Dress [25], who phrased it in terms of more general categories than G-set and R-mod, given by certain axioms.

To some extent the definition of Mackey functors in terms of G-sets is a question of notation: according to the first definition we would write $M(H)$ for the value of M at the subgroup H, whereas with the G-set definition we would write $M(G/H)$. In this account we will sometimes use both of these notations and switch from one to the other without special comment. It should be clear by looking at whether the argument is a subgroup or a G-set which notation we are using.

To make the connection between the two definitions on morphisms, we first identify two particular morphisms of G-sets. When H is a subgroup of K there is a morphism $\pi_H^K : G/H \to G/K$ specified by $\pi_H^K(xH) = xK$. When $g \in G$ and H is a subgroup of G there is also a morphism $c_g : G/H \to G/{}^gH$ specified by $c_g(xH) = xg^{-1}\,{}^gH$. It is the case that any morphism between coset spaces is a composite of these two types of morphisms (and it is an instructive exercise to prove it).

We now identify the operation I_H^K of a Mackey functor given according to the first definition with the morphism $M_*(\pi_H^K)$, and R_H^K with $M^*(\pi_H^K)$. The operation c_g of the first definition is identified with $M_*(c_g)$, which is necessarily equal to $M^*(c_{g^{-1}})$ in the presence of the axioms.

We have to check that the axioms of the first definition imply the axioms of the second, and vice-versa. Most of this is routine, and the most sophisticated aspect is the reformulation of the Mackey formula as the axiom on pullbacks. The key here is the following result.

LEMMA 2.1. *Whenever H and K are subgroups of J, itself a subgroup of G, there is a pullback diagram of G-sets*

$$\begin{array}{ccc} \Omega & \longrightarrow & G/K \\ \downarrow & & \downarrow \\ G/H & \longrightarrow & G/J \end{array}$$

where $\Omega = \bigsqcup_{x \in [H\backslash J/K]} G/(H \cap {}^x K)$. With this identification of Ω the map $\Omega \to G/H$ has components $\pi^H_{H \cap {}^x K}$, and $\Omega \to G/K$ has components $\pi^K_{H^x \cap K} c_{x^{-1}}$.

PROOF. We first observe that

$$\begin{array}{ccc} J/H \times J/K & \longrightarrow & J/K \\ \downarrow & & \downarrow \\ J/H & \longrightarrow & J/J \end{array}$$

is a pullback diagram and that $J/H \times J/K \cong \bigsqcup_{x \in [H\backslash J/K]} J/(H \cap {}^x K)$. Now apply induction of G-sets from J to G (defined in the next section) to this diagram. □

In view of this it is immediate that the pullback axiom of the second definition implies the Mackey decomposition formula of the first. Conversely, the Mackey decomposition formula implies the pullback axiom for pullbacks of this form, and this is in fact sufficient to imply the axiom for all pullbacks.

3. The computation of Mackey functors using relative projectivity

The basic notion which permits the computation of a Mackey functor along the lines of the Cartan–Eilenberg stable elements method is that of relative projectivity, which is formally similar to relative projectivity in the context of representation theory. It may be expressed most intuitively in terms of induction and restriction.

We define *induction* and *restriction* of Mackey functors in terms of induction and restriction of G-sets. If Ω is a G-set and H a subgroup of G then $\Omega{\downarrow}^G_H$ denotes the set Ω regarded as an H-set by restriction of the action. If Ψ is an H-set we define a G-set $\Psi{\uparrow}^G_H = G \times_H \Psi$, namely the equivalence classes in $G \times \Psi$ of the equivalence relation $(gh, \psi) \sim (g, h\psi)$ whenever $g \in G$, $h \in H$ and $\psi \in \Psi$. Another way to describe this is that it is the set of orbits under the action of H on $G \times \Psi$ given by $h(g, \psi) = (gh^{-1}, h\psi)$ where $g \in G, h \in H$ and $\psi \in \Psi$. The action of G on $G \times_H \Psi$ comes from the left multiplication of G on G. We may check that ${\uparrow}^G_H$ is left adjoint to ${\downarrow}^G_H$ (but it is not right adjoint in general). We now define restriction and induction of Mackey functors by

$$N{\uparrow}^G_H(\Omega) = N(\Omega{\downarrow}^G_H),$$
$$M{\downarrow}^G_H(\Psi) = M(\Psi{\uparrow}^G_H).$$

Restriction of Mackey functors is what we would expect: regarding M as being defined on subgroups of G, if $K \leqslant H$ then $M{\downarrow}^G_H(K) = M(K)$. Induction is more complicated, and for subgroups H, K of G there is a formula

$$N{\uparrow}^G_H(K) = \bigoplus_{g \in [H\backslash G/K]} N(H \cap {}^gK).$$

Induction and restriction of Mackey functors satisfy relationships inherited from the corresponding operations for G-sets and most of them are what we would expect; for example there is a Mackey decomposition formula for $M{\uparrow}^G_K{\downarrow}^G_J$. The property which is perhaps surprising is that induction of Mackey functors is both left and right adjoint to restriction. A formal consequence of this, which we mention now and will use later, is that both induction and restriction are exact functors, and they send injective and projective Mackey functors (that is, injective and projective objects in the category $\mathrm{Mack}_R(G)$) to objects of the same type.

The notions of projectivity and injectivity are, however, distinct from those of relative projectivity and relative injectivity, which we now define. By applying M_* to the natural map of G-sets $\Omega{\downarrow}^G_H{\uparrow}^G_H \to \Omega$ we obtain a morphism of Mackey functors $M{\downarrow}^G_H{\uparrow}^G_H \to M$ specified by $M{\downarrow}^G_H{\uparrow}^G_H(\Omega) = M(\Omega{\downarrow}^G_H{\uparrow}^G_H) \to M(\Omega)$. If $\mathcal{X}$ is a set of subgroups of G we may form the morphism

$$\bigoplus_{H \in \mathcal{X}} M{\downarrow}^G_H{\uparrow}^G_H \to M.$$

We define M to be *$\mathcal{X}$-projective*, or *projective relative to $\mathcal{X}$* if and only if this morphism is a split epimorphism (in the category of Mackey functors). What this means is that for each subgroup J of G the sum of induction maps

$$\bigoplus_{H \in \mathcal{X}} \bigoplus_{x \in [J\backslash G/H]} M(J \cap {}^xH) \xrightarrow{(I^J_{J\cap {}^xH})} M(J)$$

is surjective, and furthermore each of these surjections can be split in a manner compatible with inductions restrictions and conjugations. It is possible to write out these compatibility conditions explicitly, but not entirely illuminating. It is usually better to work with the abstract formalism.

Dually, we may apply M^* instead of M_* as above to obtain a morphism

$$M \to \bigoplus_{H \in \mathcal{X}} M{\downarrow}^G_H{\uparrow}^G_H.$$

We say that M is *$\mathcal{X}$-injective* if and only if this morphism is a split monomorphism.

It turns out to be convenient to express induction in a notationally different form, using G-sets. If X is a G-set and M is a Mackey functor we define a new Mackey functor M_X by $M_X(\Omega) = M(\Omega \times X)$ on objects and on morphisms as follows: if $\alpha : \Omega_1 \to \Omega_2$ then

$M_{X*}(\alpha) = M_*(\alpha \times 1)$, $M_X^*(\alpha) = M^*(\alpha \times 1)$. The point about this is that in the special case when X is the G-set G/H we have $\Omega \times G/H \cong \Omega \downarrow_H^G \uparrow_H^G$ from which it follows that $M_{G/H} \cong M \downarrow_H^G \uparrow_H^G$. We define natural transformations

$$\theta_X : M_X \to M, \qquad \theta^X : M \to M_X$$

by putting

$$(\theta_X)_\Omega = M_*(\mathrm{pr}) : M_X(\Omega) = M(\Omega \times X) \to M(\Omega),$$
$$\left(\theta^X\right)_\Omega = M^*(\mathrm{pr}) : M(\Omega) \to M(\Omega \times X) = M_X(\Omega),$$

where $\mathrm{pr} : \Omega \times X \to \Omega$ is projection onto the first coordinate. There are some details to check to see that θ_X and θ^X are indeed natural transformations.

PROPOSITION 3.1. *Let $\mathcal{X}$ be a set of subgroups of G and let $X = \bigsqcup_{H \in \mathcal{X}} G/H$ be the disjoint union of the transitive G-sets G/H. The following are equivalent.*
(i) *M is $\mathcal{X}$-projective.*
(ii) *M is $\mathcal{X}$-injective.*
(iii) *θ_X is split surjective.*
(iv) *θ^X is split injective.*
(v) *M is a direct summand of M_X.*

In view of this we may take any of these equivalent conditions as the definition of $\mathcal{X}$-projectivity. There is also an equivalent definition analogous to Higman's criterion [51].

For each Mackey functor M it may be quite important to know whether M is projective relative to some proper set of subgroups, and we need techniques to determine whether or not this is so. The following straightforward result is useful in this connection.

LEMMA 3.2. *Let $\mathcal{X} \subseteq \mathcal{Y}$ be sets of subgroups of G.*
(i) *If M is $\mathcal{X}$-projective then M is $\mathcal{Y}$-projective.*
(ii) *If M is $\mathcal{X}$-projective then M is $\mathcal{X}_{\max}$-projective, where $\mathcal{X}_{\max}$ is a set of representatives up to conjugacy of the maximal elements of $\mathcal{X}$.*

In view of this, for each set of subgroups $\mathcal{X}$, M is $\mathcal{X}$-projective if and only if M is projective relative to the closure of $\mathcal{X}$ under taking subgroups and conjugates. Provided we do this, there is in fact a unique minimal set of subgroups relative to which M is projective. This may be deduced from the next result.

PROPOSITION 3.3. (i) *Let X and Y be G-sets and M a Mackey functor. If M is X-projective and also Y-projective then M is $X \times Y$-projective.*

(ii) *Let $\mathcal{X}$ and $\mathcal{Y}$ be sets of subgroups closed under taking subgroups and conjugation. If M is $\mathcal{X}$-projective and also $\mathcal{Y}$-projective then M is $\mathcal{X} \cap \mathcal{Y}$-projective.*

PROOF. We leave (i) as an exercise. Then (ii) follows from (i) because when we put $X = \bigsqcup_{H \in \mathcal{X}} G/H$ and $Y = \bigsqcup_{H \in \mathcal{Y}} G/H$ the stabilizers of G acting on $X \times Y$ are the groups in $\mathcal{X} \cap \mathcal{Y}$. □

We see from this that there is a unique minimal set of subgroups closed under conjugation and taking subgroups relative to which M is projective. This set is called a *defect set* (or *defect base*) for M. Thus a defect set of M is (informally) a set of subgroups $\mathcal{X}$ minimal such that the sum of the maps I from subgroups in $\mathcal{X}$ is surjective and split, in the sense previously discussed. This implies in particular that $M(G) = \sum_{H \in \mathcal{X}} I_H^G M(H)$ and it was this condition alone which Green used in his definition of a defect set. However, he was working in the context of Green functors, and in that case this condition is sufficient to imply everything else as we are about to see.

THEOREM 3.4 (Dress [25, Theorem 1]). *Let M be a Green functor and $\mathcal{X}$ a set of subgroups of G. For M to be $\mathcal{X}$-projective (as a Mackey functor) it suffices that the sum of the induction maps*

$$(I_H^G): \bigoplus_{H \in \mathcal{X}} M(H) \to M(G)$$

be surjective.

This makes it easy to deduce the relative projectivity of many familiar examples of Mackey functors, and in many cases to find their defect sets. So, for example, the character rings $\mathbb{Q} \otimes_{\mathbb{Z}} G_0(kG)$ where k is a field have as their defect sets all cyclic subgroups of G in case $\operatorname{char} k = 0$, and all cyclic p'-subgroups in case $\operatorname{char} k = p$. In the characteristic zero case, Artin's induction theorem coupled with Theorem 3.4 gives projectivity relative to cyclic subgroups. No smaller set of subgroups is possible since for no cyclic group is the sum of the induction maps from proper subgroups surjective. For a full discussion of these defect sets, as well as those of $G_0(kG)$ and the Green ring A(G) see Sections 9 and 10 of [56].

Provided $n \geqslant 1$ the cohomology Mackey functor $H^n(G, U)$ is the direct sum of functors giving the p-torsion subgroup $H^n(G, U)_p$. Each p-torsion functor is projective relative to p-subgroups of G. We may see this either using Theorem 3.4 applied to $H^*(G, \mathbb{Z})_p$ since the corestriction from a Sylow p-subgroup is surjective, and then quoting further theory to do with the fact that cohomology in general is a Green module over $H^*(G, \mathbb{Z})$; or for a different approach, see Section 7. We deduce that $H^n(G, U)$ is projective relative to the set of all p-subgroups for all the prime divisors of $|G|$.

In general $H^n(G, U)_p$ may have a defect set smaller than all p-subgroups, depending on the module U. For example, if $K \leqslant G$ then $H^0(-, V\uparrow_K^G) \cong (H^0(-, V))\uparrow_K^G$ as Mackey functors [62, 5.2], and so this functor is projective relative to K. On the other hand $H^0(-, \mathbb{Z}/p\mathbb{Z})$ has defect set all p-subgroups since if H is a p-subgroup which is not a Sylow p-subgroup then the corestriction map $I_H^G = 0$. From this we may see by dimension shifting that for each n there is a choice of module U so that $H^n(-, U)_p$ has defect set all p-subgroups of G.

The Burnside ring Mackey functor $B(G)$ has defect set all subgroups of G since the G-set consisting of a single point is never an orbit in a properly induced G-set.

Dress observed [25] that under the hypothesis of relative projectivity, not only is the value of a Mackey functor the sum of the images of induction maps, but that also the kernel of this map is determined. To show this he studied a resolution of the Mackey functor which he called an *Amitsur complex*, and which we now describe.

We suppose that X is a finite G-set and let $X^r = X \times \cdots \times X$ denote the r-fold product of X with itself. Let $\mathrm{pr}_{\bar{i}} : X^r \to X^{r-1}$ denote projection off component i. We consider the complex of Mackey functors

$$C\colon \ \cdots \xrightarrow{d_2} M_{X^2} \xrightarrow{d_1} M_X \xrightarrow{d_0} M \longrightarrow 0$$

which evaluated on Ω is

$$\begin{array}{ccccccccc} & & M(\Omega \times X \times X) & & M(\Omega \times X) & & & & \\ & & \| & & \| & & & & \\ C(\Omega)\colon \ \cdots & \xrightarrow{d_2} & M_{X \times X}(\Omega) & \xrightarrow{d_1} & M_X(\Omega) & \xrightarrow{d_0} & M(\Omega) & \longrightarrow & 0 \end{array}$$

where

$$d_r = \sum_{i=0}^{r} (-1)^i M_*(1 \times \mathrm{pr}_{\bar{i}}).$$

Thus $d_0 = \theta_X$, and by a standard calculation we verify that $d_r d_{r-1} = 0$.

There is a similar construction using M^* which gives a complex

$$D\colon \ 0 \longrightarrow M \xrightarrow{d_0} M_X \xrightarrow{d_1} M_{X^2} \xrightarrow{d_2} \cdots$$

with $d_0 = \theta^X$. The next result implies that the complexes C and D are acyclic in case M is X-projective or, equivalently, X-injective.

THEOREM 3.5. *If M is X-projective then both C and D are chain homotopic to the zero complex.*

PROOF. It is useful to say that a chain complex is *contractible* if it is chain homotopic to the zero complex. Summands of contractible complexes are contractible. Since M is a summand of M_X it suffices to prove the result for the functor M_X. The complex we obtain replacing M by M_X evaluated at a G-set Ω is

$$\begin{array}{ccccccc} & & M(\Omega \times X^{r+1} \times X) & & M(\Omega \times X^r \times X) & & \\ & & \| & & \| & & \\ \cdots & \xrightarrow{d_{r+1}} & M_X(\Omega \times X^{r+1}) & \xrightarrow{d_r} & M_X(\Omega \times X^r) & \xrightarrow{d_{r-1}} & \cdots. \end{array}$$

It is a routine check that the degree 1 mapping $s_r : M(\Omega \times X^r \times X) \to M(\Omega \times X^{r+1} \times X)$ given by $s_r = (-1)^{r-1} M_*(1 \times 1^r \times \Delta)$, where $\Delta : X \to X \times X$ is the diagonal, satisfies

$s_{r-1}d_r + d_{r+1}s_r = 1$, showing that the identity mapping on this complex is chain homotopic to 0. □

To say that the complexes C and D are contractible is equivalent to saying that they are isomorphic to a direct sum of complexes of the form $\cdots \to 0 \to A \xrightarrow{\alpha} A \to 0 \to \cdots$ where α is an isomorphism. This means that the complexes are acyclic, and are everywhere split. The acyclicity implies that the values $M(K)$ of the Mackey functor are given as the cokernel (if we use M_*) or the kernel (if we use M^*) of the explicitly given map d_1, and this description is compatible with inclusions of subgroups and conjugations in G.

The cokernel of the map $M_{X^2}(G) \to M_X(G)$ in the complex C may be described as a colimit, and the kernel of $M_X(G) \to M_{X^2}(G)$ in complex D may be described as a limit, as we may see in an elementary fashion. We will see in the next section that the other homology groups of C and D can be interpreted as derived functors of colimit and limit functors. For these results we only really need to work with half of the Mackey functor, either the covariant half M_* – which is known as a *coefficient system* – or the contravariant half M^*.

PROPOSITION 3.6. (i) *In the complex C,* $\mathrm{Coker}(M_{X^2}(G) \xrightarrow{d_1} M_X(G))$ *is the colimit of the diagram made up of all possible morphisms of R-modules*

$$M(H^g \cap K) \xrightarrow{I^K_{H^g\cap K}} M(K) \quad \textit{and} \quad M(H^g \cap K) \xrightarrow{I^H_{H\cap {}^gK} c_g} M(H),$$

where $H, K \in \mathcal{X}$ and $g \in G$.

(ii) *In the complex D,* $\mathrm{Ker}(M_X(G) \xrightarrow{d_1} M_{X^2}(G))$ *is the limit of the diagram made up of all possible morphisms of R-modules*

$$M(K) \xrightarrow{R^K_{H^g\cap K}} M(H^g \cap K) \quad \textit{and} \quad M(H) \xrightarrow{c_{g^{-1}} R^H_{H\cap {}^gK}} M(H^g \cap K),$$

where $H, K \in \mathcal{X}$ and $g \in G$.

PROOF. In a similar way to Lemma 2.1 (note that $X \times X$ is the pullback of $X \to \mathrm{pt} \leftarrow X$) we see in the covariant case that $M_{X^2} \to M_X$ is the direct sum of terms with component maps $M_{G/H^g\cap K} \to M_{G/H} \oplus M_{G/K}$ specified by $(-I^H_{H\cap {}^gK} c_g, I^K_{H^g\cap K})$. The cokernel of this is the stated colimit, and the contravariant case is similar. □

This observation provides the connection with one of the main examples which motivates this general development, namely the 'stable elements' formula of Cartan and Eilenberg [16]. If we take M to be the p-part of group cohomology and $\mathcal{X}$ to be all p-subgroups of G then the assertion that $H^n(G, U)_p$ is isomorphic to the stable elements in $H^n(P, U)$, where P is a Sylow p-subgroup is exactly the assertion that it is isomorphic to the limit of the above-mentioned diagram. We thus see how to generalize this formula to arbitrary Mackey functors.

COROLLARY 3.7. *If M is $\mathcal{X}$-projective then $M(G) \cong \operatorname{colim}_{\mathcal{X}} M_* \cong \lim_{\mathcal{X}} M^*$ where these terms denote the colimit and limit described in the last result. In particular, if M is a Green functor and the sum of the induction maps from subgroups in $\mathcal{X}$ to $M(G)$ is surjective, then these isomorphisms hold.*

4. Complexes obtained from G-spaces

A deficiency of the Amitsur complex considered by Dress is that it has infinite length, and it is often more useful to have a resolution of finite length. The Amitsur complex is in fact a particular case of a theory in which we obtain exact sequences of Mackey functors from the action of G on a suitable space and we now describe this. We first have to say with what kind of G-spaces we will work, and the most elegant formulation is to define a *G-space* to be a simplicial G-set, that is, a simplicial object in the category of G-sets. For the reader unfamiliar with these we may equally consider admissible G-CW complexes (or admissible G-simplicial complexes), namely CW complexes (simplicial complexes) equipped with a cellular (simplicial) action of G and satisfying the condition that for each cell (simplex) σ the stabilizer G_σ fixes σ pointwise. If Z is a G-space we denote the set of (non-degenerate) simplices (or cells) in dimension i by Z_i. For each i this is a G-set.

Given a Mackey functor M we may construct a covariant functor $F_M : G\text{-set} \to \operatorname{Mack}_R(G)$ defined by $F_M(X) = M_X$, using the covariant part of the functor M_* to give the functorial dependence on X. We may also define a contravariant functor $F^M : G\text{-set}^{\mathrm{op}} \to \operatorname{Mack}_R(G)$ defined again by $F^M(X) = M_X$, but using the contravariant part of the functor M^* to give the functorial dependence on X. Now given a G-space Z we may construct complexes of Mackey functors

$$\cdots \to M_{Z_1} \to M_{Z_0} \to M \to 0$$

and

$$\cdots \leftarrow M_{Z_1} \leftarrow M_{Z_0} \leftarrow M \leftarrow 0$$

as follows. Regarding Z as a functor $Z : \Delta^{\mathrm{op}} \to$G-set (where Δ is the category of sets of the form $\{0, \dots, n\}$ with monotone maps as morphisms) we obtain by composition a simplicial Mackey functor $F_M \circ Z$. The first sequence is now the normalized chain complex of this simplicial object, augmented by the map $\theta_{Z_0} : M_{Z_0} \to M$. The second sequence is obtained similarly from $F^M \circ Z^{\mathrm{op}}$, augmenting by the map $\theta^{Z_0} : M \to M_{Z_0}$. For future reference, let us denote by M_Z the sequence $\cdots \to M_{Z_1} \to M_{Z_0} \to 0$ without augmentation, and by M^Z the sequence $\cdots \leftarrow M_{Z_1} \leftarrow M_{Z_0} \leftarrow 0$ again without augmentation.

As an example we indicate how the Amitsur complex which Dress considered may be constructed in this way. Given a G-set X, we construct a space Z as the nerve of the category in which the objects are the elements of X, and in which there is precisely one morphism (an isomorphism) between each ordered pair of objects. Thus the r-simplices (including the degenerate ones) are in bijection with X^{r+1}. Dress's complex is the (unnormalized) chain complex of $F_M \circ Z$, augmented by θ_{Z_0}. The augmented normalized complex is a quotient of this by a contractible subcomplex, and has the same homology.

The following is the theorem which ties all this together.

THEOREM 4.1 [65,9,27]. *Let G be a finite group, M a Mackey functor for G, $\mathcal{X}$ and $\mathcal{Y}$ sets of subgroups of G which are closed under taking subgroups and conjugation, and Z a G-space. Suppose that*

(i) *M is projective relative to $\mathcal{X}$.*

(ii) *For every $Y \in \mathcal{Y}$, $M(Y) = 0$.*

(iii) *For every subgroup $H \in \mathcal{X} - \mathcal{Y}$, the fixed points Z^H are contractible.*

Then the complexes of Mackey functors

$$\cdots \to M_{Z_1} \to M_{Z_0} \to M \to 0 \quad \text{and} \quad \cdots \leftarrow M_{Z_1} \leftarrow M_{Z_0} \leftarrow M \leftarrow 0$$

are contractible; that is, they are acyclic and everywhere split.

In the case of the Amitsur complex we take $X = \bigsqcup_{H \in \mathcal{X}} G/H$ and $\mathcal{Y} = \emptyset$. Then for every $H \in \mathcal{X}$ the space Z previously constructed satisfies the condition that Z^H is contractible, since it is the nerve of the category whose objects are the elements of X^H and where there is a single morphism between each pair of objects. This category is equivalent to a category with only one object and morphism, and its nerve is contractible.

When we evaluate the sequences of the theorem at G we get sequences which express $M(G)$ in terms of the values of M on the stabilizer groups of the simplices in Z. Thus the covariant sequence may be written

$$\cdots \to \bigoplus_{\sigma \in [G \backslash Z_1]} M(G_\sigma) \to \bigoplus_{\sigma \in [G \backslash Z_0]} M(G_\sigma) \to M(G) \to 0.$$

This is particularly useful when the G-space Z has finite dimension, in which case the sequences have finite length and the acyclicity and splitting mean that the isomorphism type of $M(G)$ is determined by the isomorphism types of the remaining terms.

This approach has been used quite extensively to assist in the computation of group cohomology, and a description of these applications is given in [1] (in a more rudimentary version phrased only in terms of group cohomology, and without the force of the exact sequence). For this we fix a prime p and let $M(G) = H^n(G, U)_p$ be the Sylow p-subgroup of the group cohomology in degree n of G with coefficients in the $\mathbb{Z}G$-module U, for some $n > 0$ and U. For Z we may take various spaces, for example the order complex (i.e. the nerve) of the poset

$$\mathcal{S}_p(G) = \{H \leqslant G \mid 1 \neq H \text{ is a } p\text{-subgroup}\}$$

with G acting by conjugating the subgroups, or equally one of a number of other G-spaces (see [27,64]). We take $\mathcal{X}$ to be all p-subgroups of G and $\mathcal{Y}$ to contain just the identity subgroup. Then the conditions of the theorem are satisfied, and the isomorphism type of $H^n(G, U)_p$ is conveniently expressed by the equation

$$H^n(G, U)_p = \sum_{\sigma \in [G \backslash |\mathcal{S}_p(G)|]} (-1)^{\dim \sigma} H^n(G_\sigma, U)_p$$

the sum being over representatives of the G-orbits of (non-degenerate) simplices in the order complex $|\mathcal{S}_p(G)|$. This equation holds in the Grothendieck group of finite Abelian groups with relations given by direct sum decompositions.

We should mention also that in this context the truncated sequences $M_Z(G)$ and $M^Z(G)$ make up the E^1 and E_1 pages of what we may call (cf. [27]) the 'isotropy spectral sequences' for the equivariant homology and cohomology of G acting on Z.

There is another interpretation of the sequences M_Z and M^Z, which is that they compute the derived functors of certain limit and colimit functors. These have importance because they appear in the spectral sequence of Bousfield and Kan (see [27]). It is also interesting to have an interpretation like this of the homology of complexes such as Dress's Amitsur complex. The general framework is that we have a small category $\mathcal{C}$ and another category $\mathcal{D}$. Let $\mathcal{D}^{\mathcal{C}}$ denote the category of functors from $\mathcal{C}$ to $\mathcal{D}$, and let $\lim, \operatorname{colim}: \mathcal{D}^{\mathcal{C}} \to \mathcal{D}$ denote the limit and colimit functors (assuming they exist). If $\mathcal{D}$ is an Abelian category then so is $\mathcal{D}^{\mathcal{C}}$, and we may consider the right derived functors $\lim^i$ of lim and the left derived functors colim_i of colim.

We need to consider the situation where our G-space Z is constructed as the homotopy colimit $Z = \operatorname{hocolim} \alpha$ of a diagram $\alpha: \mathcal{C} \to G$-set. For this construction see [27], where it is shown that $|\mathcal{S}_p(G)|$ and many other spaces may be constructed in this way up to equivariant homeomorphism.

As an example we show how to construct the space which gives the Amitsur complex in this way. Starting with a G-set X we let $\mathcal{C}$ be the category whose objects are the orbits of G on the various sets X^{r+1}, $r \geqslant 0$ and where the morphisms $\Omega \to \Psi$ are the restrictions of all possible projection mappings $\mathrm{pr}_I : X^{r+1} \to X^{|I|}$ where I is a subset of the set $\{0, 1, \ldots, r\}$ indexing the product X^{r+1}. Let $\alpha: \mathcal{C} \to G$-set be the inclusion functor. Now $\operatorname{hocolim} \alpha$ is G-homeomorphic to the space described earlier which gives Dress's Amitsur complex. This can be proved using the methods of [27], by showing that the 'Grothendieck construction' of α is a category whose nerve is the desired space.

Given a Mackey functor M and a diagram of G-sets $\alpha : \mathcal{C} \to G$-set, we obtain diagrams of R-modules by composition $M_* \circ \alpha : \mathcal{C} \to R$-mod and $M^* \circ \alpha^{\mathrm{op}} : \mathcal{C}^{\mathrm{op}} \to R$-mod.

PROPOSITION 4.2. *Let* $Z = \operatorname{hocolim} \alpha$ *where* $\alpha: \mathcal{C} \to G$*-set. Then*

$$\operatorname{colim}_i(M_* \circ \alpha) \cong H_i(M_Z),$$

and

$$\lim{}^i(M^* \circ \alpha) \cong H_i\big(M^Z\big).$$

PROOF. The idea of the proof is that the construction of the homotopy colimit of the diagram α of sets may also be done with the diagram $M_* \circ \alpha$ of R-modules, in which case the result is a simplicial R-module whose degree i homology is $\operatorname{colim}_i(M_* \circ \alpha)$ [28, App. II, 3.3]. We get the same answer if we form $\operatorname{hocolim} \alpha$, apply M_* and take homology. The argument for the second isomorphism involving $\lim^i$ is dual. □

As a consequence of this identification and the previous theorem we obtain the following corollary.

COROLLARY 4.3. *Let M be a Mackey functor for G and $\mathcal{X}$ and $\mathcal{Y}$ be sets of subgroups of G which are closed under taking subgroups and conjugation. Suppose that*

(i) *M is projective relative to $\mathcal{X}$.*

(ii) *For every $Y \in \mathcal{Y}$, $M(Y) = 0$.*

Let $\mathcal{C}$ be the full subcategory of G-set whose objects are the coset spaces G/H where $H \in \mathcal{X} - \mathcal{Y}$. Then $M_* : \mathcal{C} \to R$-mod and $M^* : \mathcal{C}^{\mathrm{op}} \to R$-mod may be regarded as diagrams of R-modules and we have $\lim^i M^* = 0$ and $\mathrm{colim}_i\, M_* = 0$ for all $i > 0$.

PROOF. Let $\alpha : \mathcal{C} \to G$-set be the inclusion functor. The space $\mathrm{hocolim}\,\alpha$ is considered in [27] where it is denoted $X^{\beta}_{\mathcal{C}}$, and it is shown that if $H \in \mathcal{X} - \mathcal{Y}$ then $(\mathrm{hocolim}\,\alpha)^H$ is contractible. Thus the conditions of the previous theorem are satisfied and we deduce that $H_i(M_{\mathrm{hocolim}\,\alpha}) = 0 = H_i(M^{\mathrm{hocolim}\,\alpha})$ when $i > 0$. □

5. Mackey functors as representations of the Mackey algebra

The structure of Mackey functors may be analyzed in a similar way to the representation theory of finite groups with many similarities in the results. We describe this approach in this section.

To start with, Mackey functors really are the same thing as modules for a certain finite-dimensional algebra defined as follows. As always we work over a commutative ground ring R, which has a 1. We consider the free algebra on non-commuting variables $I_H^K, R_H^K, c_{x,H}$ where H and K range over subgroups of G with $H \leqslant K$, and x ranges over elements of G. The *Mackey algebra* $\mu_R(G)$ is the quotient of this algebra by the ideal given by the following relations.

(0) $I_H^H = R_H^H = c_{h,H}$ for all subgroups H and $h \in H$,

$$\left.\begin{array}{l}(1)\ R_J^K R_K^H = R_J^H \\ (2)\ I_K^H I_J^K = I_J^H\end{array}\right\} \text{for all subgroups } J \leqslant K \leqslant H,$$

(3) $c_{g,{}^hK} c_{h,K} = c_{gh,H}$ for all $g, h \in G$ and subgroups K,

$$\left.\begin{array}{l}(4)\ R^{{}^gH}_{{}^gK} c_{g,H} = c_{g,K} R_K^H \\ (5)\ I^{{}^gH}_{{}^gK} c_{g,K} = c_{g,H} I_K^H\end{array}\right\} \text{for all subgroups } K \leqslant H \text{ and } g \in G,$$

(6) $R_J^H I_K^H = \sum_{x \in [J\backslash H/K]} I^J_{J\cap {}^xK} c_x R^K_{J^x \cap K}$ for all subgroups $J, K \leqslant H$,

(7) $\sum_{H \leqslant G} I_H^H = 1$,

(8) All other products of I_H^L, R_J^K, $c_{h,Q}$ are zero.

A Mackey functor M may be regarded as the $\mu_R(G)$-module $\bigoplus_{H \leqslant G} M(H)$ where the generators I_H^K, R_H^K and $c_{x,H}$ act on each summand in this direct sum as the mappings $I_H^K : M(H) \to M(K)$, $R_H^K : M(K) \to M(H)$, $c_{x,H} : M(H) \to M({}^xH)$ where this is possible, and as zero on other summands.

It is immediate to see that the ideal of relations will act as zero, since these relations are part of the definition of the Mackey functor. Conversely we note that $1 = \sum_{H \leqslant G} I_H^H$ is a sum of orthogonal idempotents, and so if we have a $\mu_R(G)$-module V we may write it as $V = \bigoplus_{H \leqslant G} I_H^H \cdot V$. The specification $M(H) = I_H^H \cdot V$ defines a Mackey functor, with the action of I_H^K, R_H^K and $c_{x,H}$ coming from the module structure.

The standard approaches to the representation theory of finite-dimensional algebras may now be applied to Mackey functors. This has been done in [37,60,62,67,9,51] to name just a few sources, and there is a summary in [57]. We will describe some of the principal ideas. We will refer to Mackey functors, but we could equally refer to $\mu_R(G)$-modules, and similarly we will refer to subfunctors instead of submodules.

There are simple Mackey functors (having no proper subfunctors), and if R is a field (or a complete local ring) they have projective covers, which form a complete list of indecomposable projective Mackey functors. These simple Mackey functors are parametrized and explicitly described in [60], where it is also proved that if R is a field of characteristic 0 or of characteristic not dividing $|G|$ then $\mu_R(G)$ is semisimple (see also [62, (14.4)]). There is a decomposition map analogous to that for group representations. It is surjective, the Cartan matrix satisfies the equation $C = D^t D$ (where D is the decomposition matrix) and hence it is symmetric and non-singular [62]. This also provides a very effective way to compute the Cartan matrix. Further information about the projectives of a rather deep and fundamental nature is given in [9].

The notion of relative projectivity was developed in [51] into a theory of vertices and sources, as well as Green correspondence. We have seen that for each Mackey functor M there is a unique set of subgroups $\mathcal{X}$ closed under conjugation and taking subgroups, minimal with respect to the property that M is $\mathcal{X}$-projective. If R is a field or a complete discrete valuation ring and M is indecomposable this set consists of a single conjugacy class of subgroups together with their subgroups. A representative of this single conjugacy class is called a *vertex* of M. The notions of source and Green correspondence are now formulated in the usual way. Unlike the situation with group representations in characteristic p, the vertex of an indecomposable Mackey functor need not be a p-subgroup of G; in fact any subgroup of G may be the vertex of an indecomposable Mackey functor, even when the Mackey functor is projective. This points to another difference with group representations, which is that whereas an indecomposable Mackey functor whose vertex is the identity and whose values are projective R-modules is necessarily projective, the converse is not true (assuming $G \neq 1$).

Various techniques are available to analyze in detail the subfunctor structure of a specific Mackey functor. A method is described in [62] to find the composition factors of the Mackey functor, and there is developed a way to compute Ext groups between the simple functors. We generally expect the subfunctor structure of a Mackey functor to be more complicated than the submodule structure of representations of the same groups, but still in small cases it can be done. In [62] it is proved that when R is a field of characteristic p and p divides $|G|$ to the first power, but not the second, $\mu_R(G)$ is a direct sum of semisimple algebras and Brauer tree algebras in an explicitly given way, so that all Mackey functors can be completely described in this situation. Such algebras are self-injective and of finite representation type. It is proved that if $p^2 \mid |G|$ then $\mu_R(G)$ is neither self-injective, nor of finite representation type.

The Mackey algebra is a direct sum of indecomposable ideal summands, and these are the *blocks* of Mackey functors. In [62] these are explicitly parametrized in terms of the blocks of G and its sections, and properties are described which enable us to determine the block to which a given indecomposable Mackey functor belongs.

6. Induction theorems and the action of the Burnside ring

The Burnside ring plays a particularly important role with regard to Mackey functors. On the one hand it provides an example of a Mackey functor $M(G) = B(G)$, which is in fact a Green functor. As a Mackey functor, B is generated by the (isomorphism class of the) G-set which consists of a single point. Furthermore it satisfies a universal property, that given any Mackey functor N, every assignment $\eta(\text{point}) \in N(G)$ extends uniquely to a morphism of Mackey functors $\eta : B \to N$. It follows from this that B is a projective object in $\mathrm{Mack}_R(G)$. More generally, if we denote by B^H the Burnside ring functor as a Mackey functor on H and its subgroups, then $B^H \uparrow_H^G$ is a projective Mackey functor (since induction carries projectives to projectives). If we assume R is a field or a complete discrete valuation ring then every indecomposable projective Mackey functor is a summand of some $B^H \uparrow_H^G$ (see [62, 8.6]).

Turning to another structure, there is an action of the ring $B(G)$ as a ring of endomorphisms of every Mackey functor for G. This action may defined in several equivalent ways. In terms of G-set notation, if X is a finite G-set and M a Mackey functor we have previously defined (in the context of relative projectivity) natural transformations

$$M \xrightarrow{\theta^X} M_X \xrightarrow{\theta_X} M.$$

We define X to act on M as the composite $\theta_X \theta^X$. It is hard to see at first what this composite is doing. In the particular case when $X = G/K$ for some subgroup K, the effect on $M(G/H)$ is a composite of maps $M(G/H) \to M(G/H \times G/K) \to M(G/H)$ where we have an identification

$$M(G/H \times G/K) \cong \bigoplus_{g \in [H \backslash G/K]} M(G/H \cap {}^gK).$$

From this we may see that if $x \in M(G/H)$ then

$$G/K \cdot x = \sum_{g \in [H \backslash G/K]} I^H_{H \cap {}^gK} R^H_{H \cap {}^gK}(x).$$

Yet another way to specify the action of the Burnside ring is to observe that there is an R-algebra homomorphism $B(G) \to \mu_R(G)$ specified on basis elements by

$$G/K \mapsto \sum_{H \leqslant G} \sum_{g \in [H \backslash G/K]} I^H_{H \cap {}^gK} R^H_{H \cap {}^gK}.$$

This homomorphism is injective [62], and because the resulting action commutes with the Mackey functor operations, it embeds $B(G)$ as a subalgebra of the center of $\mu_R(G)$.

It follows from this that any expression $1 = e_1 + \cdots + e_r$ in $B(G)$ as a sum of orthogonal idempotents gives a decomposition of every Mackey functor as $M = e_1 M \oplus \cdots \oplus e_r M$, and that the indecomposable summands of each $e_i M$ lie in distinct blocks from the summands of the other e_jM with $j \neq i$. This is because blocks may be identified as the primitive central idempotents in $\mu_R(G)$, and each e_i is a sum of blocks.

It becomes important to have a description of the primitive idempotents in $B(G)$. When $|G|$ is invertible in R, $B(G)$ is semisimple and an explicit description of the idempotents appears in [68] and [29]. In practical applications with Mackey functors whose values have torsion it is helpful to know the result of Dress [23] which shows that when p is a prime and all prime divisors of $|G|$ other than p are invertible in R, the primitive idempotents in $B(G)$ are in bijection with conjugacy classes of p-perfect subgroups of G. (A subgroup J is p-perfect if it has no non-identity p-group as a homomorphic image.) Writing f_J for the corresponding primitive idempotent of $B(G)$, several descriptions are given in [62] which characterize the summand f_JM of M. In particular when R is additionally a field or complete discrete valuation ring, the summands of f_JM are precisely the summands of M which have a vertex containing J as a normal subgroup of p-power index. This means we can tell which of the summands f_JM are non-zero by knowing a defect set of M. For example, if $M(G) = H^n(G, U)_p$ is the Mackey functor given by taking the p-torsion subgroup of group cohomology in degree n with some $\mathbb{Z}G$-module U, a defect set will consist entirely of p-subgroups of G. Here we may regard M as a Mackey functor over $R = \mathbb{Z}_p$, the p-adic integers. Every indecomposable summand of M has a p-subgroup as a vertex. Since the only p-perfect subgroup of a p-group is the identity subgroup, we have $f_1M = M$, all other summands being zero, and so the decomposition of M given by Burnside ring idempotents is of no help in examining the structure of M. On the other hand, when the defect set of M is larger there may be more summands and useful information may be obtained. This is exemplified very nicely in [49] with the algebraic K-theory of $\mathbb{Z}G$.

We now describe an induction theorem of Dress, and for this we introduce certain subfunctors of a Mackey functor M. Let $\mathcal{X}$ be a set of subgroups of G which is closed under conjugation and taking subgroups and put $X = \coprod_{H\in\mathcal{X}} G/H$. We will again use the natural transformations $\theta_X : M_X \to M$ and $\theta^X : M \to M_X$ from Section 3, and write $I_\mathcal{X}M = \theta_X(M)$ and $R_\mathcal{X}M = \operatorname{Ker}\theta^X$. These are subfunctors of M and the letters in the notation are suggested by the fact that they are specified as the image of induction maps from subgroups in $\mathcal{X}$, and the kernel of restriction maps to subgroups in $\mathcal{X}$, respectively.

LEMMA 6.1. *Let $\mathcal{X}$ be a set of subgroups of G closed under conjugation and taking subgroups. Then*

$$I_\mathcal{X}M(K) = \sum_{J\leqslant K,\, J\in\mathcal{X}} I_J^K M(J)$$

and

$$R_\mathcal{X}M(K) = \bigcap_{J\leqslant K,\, J\in\mathcal{X}} \operatorname{Ker} R_J^K.$$

If π is a set of primes (possible empty) we write π' for the complementary set of primes and let $|G| = |G|_\pi \cdot |G|_{\pi'}$ be the product of numbers whose prime divisors lie respectively in π and π'. Given a set of subgroups $\mathcal{X}$ we will write

$$\mathcal{H}_\pi \mathcal{X} = \{K \leqslant G \mid \text{there exist } p \in \pi \text{ and } H \lhd K \text{ with } H \in \mathcal{X} \text{ and } K/H \text{ a } p\text{-group}\}.$$

THEOREM 6.2 (Dress [25, Theorems 2 and 4], [24, Theorem 7.1]). *Let $\mathcal{X}$ be a set of subgroups of G closed under conjugation and taking subgroups and let π be a set of primes. We have*

$$|G|_{\pi'} M \subseteq I_{\mathcal{H}_\pi \mathcal{X}} M + R_{\mathcal{X}} M$$

and

$$|G|_{\pi'} \cdot (I_{\mathcal{X}} M \cap R_{\mathcal{H}_\pi \mathcal{X}} M) = 0.$$

In his original formulation Dress stated this result only for the evaluation of the Mackey functors at G. In view of the identifications given in the preceding lemma, the form of the result we have given is immediate.

This theorem is perhaps most useful when we take π either to be empty, or to be all primes. Evidently $\mathcal{H}_\emptyset \mathcal{X} = \mathcal{X}$. When π consists of all primes let us simply write $\mathcal{H}\mathcal{X}$ instead of $\mathcal{H}_\pi \mathcal{X}$. In these cases we obtain:

COROLLARY 6.3. *For any set of subgroups $\mathcal{X}$ closed under conjugation and taking subgroups we have*

$$|G| \cdot M \subseteq I_{\mathcal{X}} M + R_{\mathcal{X}} M, \quad |G| \cdot (I_{\mathcal{X}} M \cap R_{\mathcal{X}} M) = 0$$

and

$$M = I_{\mathcal{H}\mathcal{X}} M + R_{\mathcal{X}} M, \quad I_{\mathcal{X}} M \cap R_{\mathcal{H}\mathcal{X}} M = 0.$$

As a consequence:

COROLLARY 6.4. *Suppose that $|G|$ is invertible in R. Then*

$$M = I_{\mathcal{X}} M \oplus R_{\mathcal{X}} M.$$

An example of the application of this is Conlon's theorem [19, (81.36)], giving a decomposition of the Green ring $A(G)$ of finitely generated kG-modules where k is either a field of characteristic p or a complete discrete valuation ring with residue field of characteristic p. If $P \leqslant G$ is a p-subgroup we consider the subspace U of $A(G)$ spanned by the modules which are relatively P-projective and the subspace V spanned by all expressions $[B] - [C] - [A]$ arising from short exact sequences of kG-modules $0 \to A \to B \to C \to 0$ which split on restriction to P. Then taking

$$\mathcal{X} = \{H \mid H/O_p(H) \text{ is cyclic}, O_p(H) \text{ is conjugate to a subgroup of } P\}$$

it follows from the discussion in [19] that $U = I_{\mathcal{X}}A(G)$ and $V = R_{\mathcal{X}}A(G)$. We have $A(G) = U \oplus V$.

At this point we should mention that the semisimplicity of the category of Mackey functors over a field R of characteristic 0 mentioned in Section 5 allows us to say that every subfunctor of a Mackey functor M over R, and in particular $I_{\mathcal{X}}M$, is a direct summand of M. The extra information in Dress's result is that it identifies $R_{\mathcal{X}}M$ as a direct complement.

We also point out that if we have several sets of subgroups of G in a chain

$$\mathcal{X}_0 \subset \mathcal{X}_1 \subset \cdots \subset \mathcal{X}_n$$

then evidently

$$I_{\mathcal{X}_0}M \subseteq I_{\mathcal{X}_1}M \subseteq \cdots \subseteq I_{\mathcal{X}_n}M$$

and

$$R_{\mathcal{X}_0}M \supseteq R_{\mathcal{X}_1}M \supseteq \cdots \supseteq R_{\mathcal{X}_n}M.$$

When $|G|$ is invertible in R we have

$$M = I_{\mathcal{X}_i}M \oplus R_{\mathcal{X}_i}M$$

for each i and by the modular law we have

$$I_{\mathcal{X}_i}M = I_{\mathcal{X}_{i-1}}M \oplus (I_{\mathcal{X}_i}M \cap R_{\mathcal{X}_{i-1}}M).$$

Hence

$$M = \bigoplus_{i=0}^{n} I_{\mathcal{X}_i}M \cap R_{\mathcal{X}_{i-1}}M,$$

where $R_{\mathcal{X}_{-1}}M = M$. This decomposition is exemplified by a different part of Conlon's theorem [19, (81.36)].

The penultimate equality in Corollary 6.3 is very useful in obtaining induction theorems in situations where we know $R_{\mathcal{X}}M$ to be zero. So for example there is an induction theorem due to Conlon [19, (80.50)], for the Green ring $A(G)$ over $\mathbb{Q}$ of finitely generated kG-modules where k is either a field of characteristic p or a complete discrete valuation ring with residue field of characteristic p. It says that if $\mathcal{X} = \{H \leqslant G \mid H/O_p(H) \text{ is cyclic}\}$ then $A(G)$ is the sum of the images of the induction maps from the subgroups in $\mathcal{X}$, or in our language $A = I_{\mathcal{X}}A$. From this and Corollary 6.4 we obtain $R_{\mathcal{X}}A = 0$. If we now let $a(G)$ denote the Green ring over $\mathbb{Z}$ of finitely generated kG-modules, allowing only linear combinations over $\mathbb{Z}$, we deduce that $R_{\mathcal{X}}a = 0$, since it embeds in $R_{\mathcal{X}}A = 0$. We deduce from Corollary 6.3 that $a = I_{\mathcal{H}\mathcal{X}}a$, which is the integral form of Conlon's induction theorem, due to Dress in [26].

7. Cohomological Mackey functors

A Mackey functor is said to be *cohomological* if for every pair of subgroups $H \leqslant K$ of G the map $I_H^K R_H^K : M(K) \to M(K)$ is multiplication by $|K : H|$. These functors take their name because group cohomology $M(K) = H^n(K, U)$ satisfies this condition. The functor which assigns to each subgroup of the Galois group of an extension of number fields the class group of the ring of integers of the fixed field is another example of a cohomological Mackey functor, since the Mackey functor operations derive from taking fixed points.

Perhaps the most striking result about cohomological Mackey functors is the theorem of Yoshida which identifies them as modules for the Hecke algebra

$$\mathcal{E} = \mathrm{End}_{RG}\Bigl(\bigoplus_{H \leqslant G} R[G/H]\Bigr),$$

the endomorphism ring of the direct sum of all permutation modules $R[G/H]$ where H ranges over the subgroups of G. They are also related to the category $\mathcal{H}_G$ defined to be the full subcategory of RG-modules whose objects are the finitely generated permutation RG-modules.

THEOREM 7.1 (Yoshida [69]). *The following categories are equivalent:*

(i) *the full subcategory of* $\mathrm{Mack}_R(G)$ *whose objects are the cohomological Mackey functors,*
(ii) *the category of R-linear functors $\mathcal{H}_G \to R$*-mod*, and*
(iii) *the category of $\mathcal{E}$-modules.*

This result identifies $\mathcal{E}$ with what we might call the 'cohomological Mackey algebra', obtained by imposing the relations $I_H^K R_H^K = |K : H| \cdot I_K^K$ on the Mackey algebra. The equivalence of (ii) and (iii) is a routine piece of category theory, immediate from the definitions.

What is behind the equivalence of (i) and (ii) is that all morphisms in $\mathcal{H}_G$ can be expressed as linear combinations of composites of three kinds of morphism: the morphisms

$$i_H^K : R[G/H] \to R[G/K],$$
$$r_H^K : R[G/K] \to R[G/H],$$
$$c_g : R[G/H] \to R[G/{}^gH]$$

specified by $i_H^K(xH) = xK$ and $r_H^K(xK) = \sum_{k \in [K/H]} xkH$ whenever H is a subgroup of K, and $c_g(xH) = xg^{-1}{}^gH$ whenever $g \in G$ and $H \leqslant G$. These morphisms satisfy all the relations satisfied by the corresponding Mackey functor operations, and also the relation that $i_H^K r_H^K$ is multiplication by $|K : H|$ whenever H is a subgroup of K. Further, all relations between the morphisms are deducible from the relations just mentioned. It follows that any R-linear functor $\mathcal{H}_G \to R$-mod can be regarded as a cohomological Mackey functor, and conversely, any cohomological Mackey functor can be regarded as being defined on $\mathcal{H}_G$.

The use of this approach is that any isomorphism between direct sums of permutation modules yields a relationship between the values of a cohomological Mackey functor. Such relationships are exploited in [50,63] and [8].

The following is an exercise in working from the definitions.

PROPOSITION 7.2. *Suppose that M is a cohomological Mackey functor and that $H \leqslant G$ is a subgroup such that $|G:H|$ acts invertibly on all of the R-modules $M(K)$, where $K \leqslant G$. Then M is projective relative to $\{H\}$.*

In view of this, when $R = \mathbb{Q}$ all cohomological Mackey functors are 1-projective, and when $R = \mathbb{Z}_p$ cohomological Mackey functors are projective relative to the p-subgroups of G. In this case, according to the section on Burnside ring action, they are acted on as the identity by the Burnside ring idempotent f_1.

It is shown in [62] that every cohomological Mackey functor is in fact a homomorphic image of a Mackey functor of the form $M(G) = H^0(G, U)$, where U is a permutation module. In fact, the indecomposable summands of these functors are precisely the indecomposable projective cohomological Mackey functors. Further information about the projective and simple cohomological Mackey functors is given in [62].

8. Globally-defined Mackey functors

We describe now another kind of Mackey functor which has appeared more recently and which appears to be important. These are the globally-defined Mackey functors. In some ways they are more general than the original Mackey functors, and in some ways more restrictive. One main difference is that instead of being defined just on the subgroups of a particular group, they are defined on all finite groups. This is in keeping with many of the natural examples of Mackey functor, such as group cohomology with trivial coefficients, or the character ring, which are in fact defined on all groups. A second main difference is that whereas the original Mackey functors only possess operations corresponding to inclusions of subgroups and conjugations, the globally-defined Mackey functors may possess operations for all group homomorphisms. This possibility necessitates slightly more restrictive axioms to make it work.

Let R be a commutative ring with a 1. By saying that a group K is a *section* of a group G we mean that there is a subgroup H of G and a surjective homomorphism $H \to K$. Let $\mathcal{X}$ and $\mathcal{Y}$ be classes of finite groups satisfying the following two conditions: (1) if G lies in $\mathcal{X}$ and K is a section of G then K lies in $\mathcal{X}$; and (2) if $1 \to A \to B \to C \to 1$ is a short exact sequence of groups with $A \in \mathcal{X}$ and $C \in \mathcal{X}$ then $B \in \mathcal{X}$. We say that a *globally-defined Mackey functor* over R, with respect to $\mathcal{X}$ and $\mathcal{Y}$, is a structure M which specifies an R-module $M(G)$ for each finite group G, together with for each homomorphism $\alpha : G \to K$ with $\operatorname{Ker}\alpha \in \mathcal{Y}$ an R-module homomorphism $\alpha_* : M(G) \to M(K)$ and for each homomorphism $\beta : G \to K$ with $\operatorname{Ker}\beta \in \mathcal{X}$ an R-module homomorphism $\beta^* : M(K) \to M(G)$. These morphisms should satisfy the following relations:

(1) $(\alpha\gamma)_* = \alpha_*\gamma_*$ and $(\beta\delta)^* = \delta^*\beta^*$ always, whenever these are defined;
(2) whenever $\alpha : G \to G$ is an inner automorphism then $\alpha_* = 1 = \alpha^*$;

(3) for every commutative diagram of groups

$$\begin{array}{ccc} G & \xrightarrow{\beta} & H \\ {\scriptstyle\gamma}\uparrow & & \uparrow{\scriptstyle\alpha} \\ \beta^{-1}(K) & \xrightarrow[\delta]{} & K \end{array}$$

in which α and γ are inclusions and β and δ are surjections we have $\alpha^*\beta_* = \delta_*\gamma^*$ whenever $\operatorname{Ker}\beta \in \mathcal{Y}$, and $\beta^*\alpha_* = \gamma_*\delta^*$ whenever $\operatorname{Ker}\beta \in \mathcal{X}$;

(4) for every commutative diagram

$$\begin{array}{ccc} G & \xrightarrow{\gamma} & H/\operatorname{Ker}\alpha\operatorname{Ker}\beta \\ {\scriptstyle\beta}\uparrow & & \uparrow{\scriptstyle\delta} \\ H & \xrightarrow[\alpha]{} & K \end{array}$$

in which α, β, γ and δ are all surjections, with $\operatorname{Ker}\beta \in \mathcal{Y}$ and $\operatorname{Ker}\alpha \in \mathcal{X}$, we have $\beta_*\alpha^* = \gamma^*\delta_*$;

(5) (Mackey axiom) for every pair of subgroups $J, K \leqslant H$ of every group H we have

$$(\iota_K^H)^*(\iota_J^H)_* = \sum_{h\in[K\backslash H/J]} (\iota_{K\cap {}^hJ}^K)_* c_{h*} (\iota_{K^h\cap J}^J)^*,$$

where $\iota_K^H : K \hookrightarrow H$ and $\iota_J^H : J \hookrightarrow H$ etc. are the inclusion maps and $c_h : K^h \cap J \to K \cap {}^hJ$ is the homomorphism $c_h(x) = hxh^{-1}$.

These globally defined Mackey functors form an Abelian category denoted $\mathrm{Mack}_R^{\mathcal{X},\mathcal{Y}}$.

At first sight some conditions appear to have been omitted which are necessary to make things work. Thus in both (3) and (4), if $\operatorname{Ker}\beta \in \mathcal{Y}$ it follows that $\operatorname{Ker}\delta \in \mathcal{Y}$ also; in (1), if $\operatorname{Ker}\alpha \in \mathcal{Y}$ and $\operatorname{Ker}\gamma \in \mathcal{Y}$ then $\operatorname{Ker}\alpha\gamma \in \mathcal{Y}$ also; and so on. Axiom (4) implies that if $\alpha : H \to K$ is an isomorphism, then $(\alpha^{-1})_* = \alpha^*$ and $(\alpha^{-1})^* = \alpha_*$. The automorphisms of each group G act on $M(G)$, and because the inner automorphisms act trivially each $M(G)$ has the structure of an $R[\operatorname{Out} G]$-module.

The main reason for having the classes $\mathcal{X}$ and $\mathcal{Y}$ as part of the definition is that a globally-defined Mackey functor need not possess all operations α_* and α^* when α is a surjective group homomorphism, and with each example we discuss the possibilities for $\mathcal{X}$ and $\mathcal{Y}$. It is always possible to take $\mathcal{X}$ and $\mathcal{Y}$ to consist only of the identity group, which is the same as saying that α_* and α^* are only defined when α is injective. Sometimes it is possible to take $\mathcal{X}$ and $\mathcal{Y}$ to be larger classes of groups.

Some of the examples of ordinary Mackey functors we have previously discussed also give examples of globally-defined Mackey functors; and some do not. The following are examples of globally-defined Mackey functors.

- $G_0(kG)$ and $B(G)$: in both these examples we may take $\mathcal{X}$ and $\mathcal{Y}$ to be all finite groups. Whenever $\alpha : G \to H$ is a group homomorphism we may restrict both representations of H and H-sets along α. Also we may form $RH \otimes_{RG} U$ and $H \times_G \Omega$ whenever U is an RG-module and Ω a G-set, and this allows us to form α_*.

- $H^n(G, R)$, $H_n(G, R)$, $\hat{H}^n(G, R)$: the cohomology, homology and Tate cohomology of G in some dimension n with trivial coefficients (arbitrary coefficient modules are not possible since they must be modules for every finite group). For cohomology we may take $\mathcal{X}$ to be all finite groups and $\mathcal{Y}$ to be the identity group. If $\alpha : G \to H$ is a surjective group homomorphism then α^* is inflation. However, provided we allow such inflations it is not possible to define α_* (except on isomorphisms) so as to satisfy the axioms. To see this, consider the fixed point functor $M(G) = H^0(G, R)$ for each finite group G, and also the homomorphisms $1 \xrightarrow{\iota} G \xrightarrow{\beta} 1$. We know that $\iota_* = |G| \cdot \mathrm{id}$ and $\iota^* = \mathrm{id}$. Since $\beta\iota = \mathrm{id}$ we have $\beta_*\iota_* = \mathrm{id}$ and $\iota^*\beta^* = \mathrm{id}$. From this we deduce that $\beta^* = \mathrm{id}$ and $\beta_* = |G|^{-1} \cdot \mathrm{id}$. At this point if $|G|^{-1}$ does not exist in R we see that β_* cannot be defined. Even when $|G|^{-1}$ does exist in R, consider axiom (4) applied to the square

$$\begin{array}{ccc} 1 & \longrightarrow & 1 \\ {\scriptstyle\beta}\uparrow & & \uparrow \\ G & \underset{\beta}{\longrightarrow} & 1 \end{array}$$

This allows us to deduce that $\beta_*\beta^* = \mathrm{id}$, that is $|G|^{-1} = 1$, which cannot hold for all finite groups G.
- $K_n(\mathbb{Z}G)$, the algebraic K-theory of $\mathbb{Z}G$. Here we may take $\mathcal{Y}$ = all finite groups but put $\mathcal{X} = 1$ (see [49]), the point being that if $\alpha : G \to H$ is a group homomorphism and P is a projective $\mathbb{Z}G$-module then $\mathbb{Z}H \otimes_{\mathbb{Z}G} P$ is a projective $\mathbb{Z}H$-module, but the restriction of a projective module along a homomorphism α is only projective when α is injective.
- $\mathbb{Q} \otimes D(G)$ when G is a p-group and $D(G)$ is the Dade group of endopermutation kG-modules, where k is a field of characteristic p (see [14]). Here we have to consider a modified version of the theory where we consider functors defined only on p-groups. We may take $\mathcal{X}$ and $\mathcal{Y}$ to be all p-groups.

As is the case with ordinary Mackey functors, there is another definition of globally-defined Mackey functors [10] which is less immediately transparent, but which is usually easier to work with. Given a pair of groups G and H we consider the finite (G, H)*-bisets*. These are finite sets Ω with a left action of G and a right action of H so that the two actions commute: $g(\omega h) = (g\omega)h$ for all $g \in G$, $h \in H$ and $\omega \in \Omega$. By analogy with the Burnside ring, let $A^{\mathcal{X},\mathcal{Y}}(G, H)$ be the Grothendieck group with respect to disjoint unions of all finite (G, H)-bisets Ω with the property that $\mathrm{Stab}_G(\omega) \in \mathcal{X}$ and $\mathrm{Stab}_H(\omega) \in \mathcal{Y}$ for all $\omega \in \Omega$. This is the free Abelian group with the isomorphism classes of transitive such bisets as a basis – we say Ω is *transitive* if given $\omega \in \Omega$, every element of Ω may be written $g\omega h$ for some $g \in G$ and $h \in H$. We now define $A_R^{\mathcal{X},\mathcal{Y}}(G, H) = R \otimes_{\mathbb{Z}} A^{\mathcal{X},\mathcal{Y}}(G, H)$.

Given a third group K there is a product

$$A_R^{\mathcal{X},\mathcal{Y}}(G, H) \times A_R^{\mathcal{X},\mathcal{Y}}(H, K) \to A_R^{\mathcal{X},\mathcal{Y}}(G, K)$$

defined on basis elements as $(\Omega, \Psi) \mapsto \Omega \times_H \Psi$ where the latter amalgamated product is the set of equivalence classes under the relation $(\omega h, \psi) \sim (\omega, h\psi)$ whenever $\omega \in \Omega$, $\psi \in \Psi$ and $h \in H$. This product is associative, and provides in particular a ring structure on $A_R^{\mathcal{X},\mathcal{Y}}(G,G)$. When $\mathcal{X}$ consists of all finite groups and $\mathcal{Y}$ consists of the identity group, this ring is known as the *double Burnside ring* of G, see [3] or [43]. (We have chosen the opposite convention to many authors, who take $\mathcal{X}=1$ and $\mathcal{Y}=$ all finite groups.)

With all this we associate a category $\mathcal{C}_R^{\mathcal{X},\mathcal{Y}}$ whose objects are all finite groups and where $\operatorname{Hom}_{\mathcal{C}_R^{\mathcal{X},\mathcal{Y}}}(H,G)=A_R^{\mathcal{X},\mathcal{Y}}(G,H)$. The composition of morphisms is the product we have defined, and because we have (apparently perversely!) reversed the order of G and H this composition is correct for applying mappings from the left. Finally, a *globally-defined Mackey functor* (with respect to $\mathcal{X}$ and $\mathcal{Y}$) is an R-linear functor $M:\mathcal{C}_R^{\mathcal{X},\mathcal{Y}}\to R$-mod.

The key to understanding why this definition is equivalent to the first one is to consider for each group homomorphism $\alpha:G\to K$ the bisets ${}_KK_G$ and ${}_GK_K$ where in the first case K acts on K from the left by left multiplication and G acts on K from the right via α and right multiplication, and in the second case the reverse happens. Given a functor M as just defined we define $\alpha_*=M({}_KK_G)$ and $\alpha^*=M({}_GK_K)$. It is the case that these bisets satisfy relations which imply the axioms we have given. Conversely, every transitive biset is a composite of bisets of this special form, and the axioms are sufficient to imply that a Mackey functor defined in the first way gives rise to an R-linear functor $M:\mathcal{C}_R^{\mathcal{X},\mathcal{Y}}\to R$-mod.

9. The computation of globally-defined Mackey functors using simple functors

We describe the first of two applications of globally-defined Mackey functors together with further properties which seem to suggest their importance. The applications depend on a description of the simple globally-defined Mackey functors and we start with this. This material is developed in [66] and [10].

As with ordinary Mackey functors, we can speak of subfunctors of globally-defined Mackey functors, kernels and so forth. We thus have the notion of a simple functor, namely one which has no proper non-zero subfunctors.

THEOREM 9.1 [10,66]. *The simple globally-defined Mackey functors are in bijection with pairs (H,U) where H is a finite group and U is a simple $R[\operatorname{Out} H]$-module (both taken up to isomorphism). The corresponding simple functor $S_{H,U}$ has the property that $S_{H,U}(H)\cong U$ as $R[\operatorname{Out} H]$-modules, and that if G is a group for which $S_{H,U}(G)\neq 0$ then H is a section of G. Provided R is a field or a complete discrete valuation ring each simple functor $S_{H,U}$ has a projective cover $P_{H,U}$, and these form a complete list of the indecomposable projective functors.*

It is a feature of this classification that it is independent of the choice of $\mathcal{X}$ and $\mathcal{Y}$, although the particular structure of the simple functors changes as we vary $\mathcal{X}$ and $\mathcal{Y}$. In the special case when $\mathcal{X}=\mathcal{Y}=1$ an explicit description (stated below) of the values $S_{H,U}(G)$ was given in [66], as well as a less transparent description in the case when $\mathcal{X}$ is all finite groups and $\mathcal{Y}=1$. In this latter case it is shown that the dimension of the $S_{H,U}(G)$ is related to the stable decomposition of the classifying space BG (as will be explained later) and existing computations of these decompositions are really equivalent to computing this dimension. When $\mathcal{X}$ and $\mathcal{Y}$ consist of all finite groups it appears to be rather difficult to describe the simple functors explicitly, in general, but we will return to this question in the last section of this article.

THEOREM 9.2 [66]. *When $\mathcal{X} = \mathcal{Y} = 1$ the simple globally-defined Mackey functors are given explicitly by*

$$S_{H,U}(G) = \bigoplus_{\substack{\alpha: H \cong L \leqslant G \\ \text{up to } G\text{-conjugacy}}} \operatorname{tr}_L^{N_G(L)}({}^{\alpha}U)$$

where H ranges over finite groups and U ranges over simple $R[\mathrm{Out}(H)]$-modules.

Here the direct sum is taken over G-orbits of isomorphisms α from H to subgroups L of G, and ${}^{\alpha}U$ means U with the action transported to $N_G(L)/L$ via α. The symbol tr means the relative trace, i.e. multiplication by the sum of coset representatives of L in $N_G(L)$.

This straightforward description of the simple functors when $\mathcal{X} = \mathcal{Y} = 1$ gives rise to a method of computing the values of globally-defined Mackey functors which has been applied in the case of group cohomology in [66] and [18]. We work with the p-torsion subgroup $M(G) = H^n(G, R)_p$ for a fixed prime p. Although M can be defined with inflation operations α^* when α is a surjective group homomorphism we choose to forget these and regard M as a globally-defined Mackey functor with $\mathcal{X} = \mathcal{Y} = 1$. In the category of such functors we consider a 'composition series' of M, namely a filtration

$$\cdots \subset M_{i-1} \subset M_i \subset M_{i+1} \subset \cdots \subset M$$

such that $\bigcap M_i = 0$, $\bigcup M_i = M$ and M_{i+1}/M_i is always a simple functor. It is shown that such a series exists, and the multiplicity of each simple functor as a factor is determined independently of the choice of the series. Furthermore, the fact that – as an ordinary Mackey functor – cohomology is projective relative to p-subgroups implies that the only simple functors which arise as composition factors in the global situation are $S_{H,U}$ where H is a p-group; and furthermore the multiplicities as composition factors are determined by knowledge of the cohomology of p-groups. Putting all this together, we get a formula for the size of $M(G)$ knowing the composition factor multiplicities and the values $S_{H,U}(G)$, and it is expressed in terms of the cohomology of the p-subgroups of G and conjugacy of p-elements. Given explicit information about the cohomology of the p-subgroups, the formula for the cohomology of G gives completely explicit numerical results. A remarkable feature of this approach is that we obtain a uniform formula which applies at once to all finite groups G with a given Sylow p-subgroup.

10. Stable decompositions of BG

For surveys of the background material to this section see [3] and [43]. We denote by $(BG_+)^{\wedge}_p$ the p-completion of the suspension spectrum obtained from the classifying space

BG after first adjoining a disjoint base point to give a space BG_+. The problem of decomposing $(BG_+)^\wedge_p$ stably as a wedge of indecomposable spectra is a fundamental question which – thanks to Carlsson's proof of the Segal conjecture – comes down to an analysis of the double Burnside ring $A^{\mathrm{all},1}_{\mathbb{Z}_p}(G, G)$ defined in an earlier section. By studying the representations of this ring it was proved by Benson and Feshbach [4] and also by Martino and Priddy [44] that the indecomposable p-complete spectra which can appear as a summand of some $(BG_+)^\wedge_p$ (allowing G to vary over all finite groups) are parametrized by pairs (H, U) where H is a p-group and U is a simple $\mathbb{Z}_p[\mathrm{Out}(H)]$-module. They also gave a method for determining the multiplicity with which each spectrum in the parametrization occurs as a summand of a given BG_+.

We describe in this section how a proof of their theorem may be given entirely within the context of globally-defined Mackey functors (assuming Carlsson's theorem). The complete description can be found in [66]. We work with globally-defined Mackey functors where we take $\mathcal{X}$ to be all finite groups and $\mathcal{Y}$ to be the isomorphism class of the identity group, and we will see that there is an equivalence of categories between the full subcategory of the category of spectra whose objects are the summands of the $(BG_+)^\wedge_p$, and the full subcategory of $\mathrm{Mack}^{\mathcal{X},\mathcal{Y}}_{\mathbb{Z}_p}$ whose objects are the projective covers $P_{H,U}$ of the simple functors $S_{H,U}$ where H is a p-group. The advantage of this approach is that we work with the projective objects in a category – and such a situation is often felt to be well-understood – rather than a more mysterious subcategory of the category of spectra.

For each group K there is a representable functor $F^K = \mathrm{Hom}_{C^{\mathcal{X},\mathcal{Y}}_{\mathbb{Z}_p}}(K, -)$ which is a projective object in $\mathrm{Mack}^{\mathcal{X},\mathcal{Y}}_{\mathbb{Z}_p}$ by Yoneda's lemma. This decomposes as a sum of indecomposable projectives, and the particular form of the decomposition will be of use to us in what follows.

LEMMA 10.1. *The representable functor F^K decomposes as $F^K \cong \bigoplus P^{n_{H,U}}_{H,U}$ where*

$$n_{H,U} = \dim S_{H,U}(K)/\dim \mathrm{End}_{\mathbb{Z}_p[\mathrm{Out}\,H]}\, U.$$

Thus $P_{H,U}$ is only a summand of F^K if H is a section of K, and $P_{K,U}$ does occur as a summand of F^K with multiplicity $\dim U/\dim \mathrm{End}_{\mathbb{Z}_p[\mathrm{Out}\,H]}\, U$.

The dimensions are taken over $\mathbb{Z}/p\mathbb{Z}$ here. This is possible since the values of a simple functor over $\mathbb{Z}_p$ are actually $\mathbb{Z}/p\mathbb{Z}$-vector spaces.

PROOF. From the properties of a projective cover we have

$$\mathrm{Hom}(P_{H,U}, S_{J,V}) = \begin{cases} \mathrm{End}(S_{H,U}) & \text{if } (H, U) \cong (J, V), \\ 0 & \text{otherwise.} \end{cases}$$

If we write $F^K \cong \bigoplus P^{n_{H,U}}_{H,U}$ for some integers $n_{H,U}$ to be determined, we have

$$\dim \mathrm{Hom}\big(F^K, S_{H,U}\big) = n_{H,U} \cdot \dim \mathrm{End}(S_{H,U}).$$

Now

$$\mathrm{Hom}\big(F^K, S_{H,U}\big) \cong S_{H,U}(K)$$

by Yoneda's lemma, and also

$$\mathrm{End}(S_{H,U}) \cong \mathrm{End}_{\mathbb{Z}_p[\mathrm{Out}\,H]}(U)$$

from [66] or [10]. Rearranging these equations gives the claimed expression for $n_{H,U}$. We obtain the remaining statements from Theorem 9.1. □

Again by Yoneda's lemma, $\mathrm{Hom}(F^G, F^H) \cong \mathrm{Hom}(H, G) = A_{\mathbb{Z}_p}^{\mathcal{X},\mathcal{Y}}(G, H)$ and composition of morphisms on the left corresponds to the product on the right. It is a consequence of Carlsson's theorem that when P is a p-group we have $[(BP_+)^\wedge_p, (BG_+)^\wedge_p] \cong A_{\mathbb{Z}_p}^{\mathcal{X},\mathcal{Y}}(G, P)$, where the left hand side denotes the homotopy classes of maps of spectra. We have reversed the expected order of G and P on the right hand side so that composition of maps written on the left corresponds to the product of bisets.

We immediately have the first part of the next result.

THEOREM 10.2. *Let p be a prime.*

(i) *The assignment $(BP_+)^\wedge_p \to F^P$ gives an equivalence of categories between the full subcategory of the category of spectra whose objects are the $(BP_+)^\wedge_p$ where P is a p-group, and the full subcategory of $\mathrm{Mack}_{\mathbb{Z}_p}^{\mathcal{X},\mathcal{Y}}$ whose objects are the representable functors F^P.*

(ii) *The equivalence in* (i) *extends to an equivalence between the full subcategory of the category of spectra whose objects are stable summands of the classifying spaces $(BG_+)^\wedge_p$ as G ranges over finite groups, and the full subcategory of $\mathrm{Mack}_{\mathbb{Z}_p}^{\mathcal{X},\mathcal{Y}}$ whose objects are the indecomposable projectives $P_{H,U}$ with H a p-group.*

PROOF. To prove the second part we first observe that the equivalence in part (i) can be extended to summands of the objects, since these correspond to idempotents in the endomorphisms rings of objects, and corresponding objects have isomorphic endomorphism rings. We see from Lemma 10.1 that the indecomposable summands of the F^P with P a p-group are precisely the $P_{H,U}$ with H a p-group. Also it is well known that $(BG_+)^\wedge_p$ is stably a summand of $(BP_+)^\wedge_p$ where P is a Sylow p-subgroup of G, and so the summands of all the $(BG_+)^\wedge_p$ are the same as the summands of all the $(BP_+)^\wedge_p$. This completes the proof. □

We now deduce that the stable summands of the $(BG_+)^\wedge_p$ are parametrized the same way as the $P_{H,U}$ with H a p-group, and the multiplicities of these summands are given by 10.1. The properties given in 10.1 are exactly the properties of the summands of the $(BG_+)^\wedge_p$ given in [4] and [44]. By analyzing the structure of $S_{H,U}(G)$ we are also able to obtain their general formula for these multiplicities.

11. Some naturally-occurring globally-defined Mackey functors

We conclude by pointing out that some very important naturally-occurring Mackey functors are in fact simple in some cases and projective in another. It is remarkable that this highly technical theory encapsulates natural examples in this way. We state results only over $\mathbb{Q}$ but in fact they hold over any field of characteristic 0.

THEOREM 11.1 (Bouc [10], Bouc and Thévenaz [14]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be all finite groups.*
(i) *The Burnside ring Mackey functor $\mathbb{Q} \otimes_{\mathbb{Z}} B$ is the indecomposable projective $P_{1,\mathbb{Q}}$.*
(ii) *The functor $M(G) = \mathbb{Q} \otimes_{\mathbb{Z}} G_0(\mathbb{Q}G)$ which assigns the representation ring of $\mathbb{Q}G$-modules, tensored with $\mathbb{Q}$, is the simple functor $S_{1,\mathbb{Q}}$.*
(iii) *Let p be a prime. The kernel of the projective cover map $P_{1,\mathbb{Q}} \to S_{1,\mathbb{Q}}$, regarded as a functor only on p-groups, is the functor $\mathbb{Q} \otimes D$, where $D(P)$ is the Dade group of endopermutation modules of the p-group P. This functor is simple: $\mathbb{Q} \otimes D \cong S_{C_p \times C_p,\mathbb{Q}}$.*

The first statement in this theorem is straightforward. For each group G we have a representable functor F^G and in case $G = 1$ we may see from the definitions that $F^1 = \mathbb{Q} \otimes_{\mathbb{Z}} B$. We also know the decomposition of this functor as a direct sum of indecomposable projectives, as given in Lemma 10.1, and from this we see it is $P_{1,\mathbb{Q}}$. The second statement is less obvious and appears in [10]. We know from Artin's induction theorem that the natural map $\mathbb{Q} \otimes_{\mathbb{Z}} B \to \mathbb{Q} \otimes_{\mathbb{Z}} G_0(\mathbb{Q}G)$ is surjective. It requires some further argument to show that the target is simple. Statement (iii) is not obvious at all. Dade's group is described in [59] and [14], and we will not discuss it here. By regarding a functor as defined only on p groups, we mean that we are considering the restriction of the functor to the full subcategory of $\mathcal{C}_{\mathbb{Q}}^{\mathcal{X},\mathcal{Y}}$ whose objects are the p-groups, and this restriction is asserted to be simple in the category of functors on this subcategory.

The kernel of the projective cover map $P_{1,\mathbb{Q}} \to S_{1,\mathbb{Q}}$ is not simple as a functor defined on all finite groups, although it is not a straightforward matter to determine its composition factors. We have the following:

THEOREM 11.2 (Bouc [10]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be all finite groups. The composition factors of the Burnside functor $\mathbb{Q} \otimes_{\mathbb{Z}} B$ all have the form $S_{H,\mathbb{Q}}$ for various groups H. Each such simple functor appears with multiplicity at most* 1.

Bouc also characterizes in [10] those groups H for which the simple functor $S_{H,\mathbb{Q}}$ does appear as a composition factor of $\mathbb{Q} \otimes_{\mathbb{Z}} B$ by means of a certain combinatorial condition (he calls such groups 'b-groups'). There is also information in [14] about the composition factors of $k \otimes_{\mathbb{Z}} B$ when k is a field of prime characteristic.

References

[1] A. Adem and R.J. Milgram, *Cohomology of Finite Groups*, Grundlehren Math. Wiss., B. 309, Springer, Berlin (1994).

[2] A.-M. Aubert, *Systèmes de Mackey*, Preprint (1993).
[3] D.J. Benson, *Stably splitting BG*, Bull. Amer. Math. Soc. **33** (1996), 189–198.
[4] D.J. Benson and M. Feshbach, *Stable splitting of classifying spaces of finite groups*, Topology **31** (1992), 157–176.
[5] R. Boltje, *A canonical Brauer induction formula*, Astérisque **181–182** (1990), 31–59.
[6] R. Boltje, *Identities in representaton theory via chain complexes*, J. Algebra **167** (1994), 417–447.
[7] R. Boltje, *A general theory of canonical induction formulae*, J. Algebra **206** (1998), 293–343.
[8] R. Boltje, *Class group relations from Burnside ring idempotents*, J. Number Theory **66** (1997), 291–305.
[9] S. Bouc, *Résolutions de foncteurs de Mackey*, Proc. Symp. Pure Math. **63** (1998), 31–84.
[10] S. Bouc, *Foncteurs d'ensembles munis d'une double action*, J. Algebra **183** (1996), 664–736.
[11] S. Bouc, *Construction de foncturs entre catégories de G-ensembles*, J. Algebra **183** (1996), 737–825.
[12] S. Bouc, *Green Functors and G-sets*, Lecture Notes in Math., Vol. 1671, Springer, Berlin (1997).
[13] S. Bouc, *Non-additive exact functors and tensor induction for Mackey functors*, Preprint (1997).
[14] S. Bouc and J. Thévenaz, *The group of endopermutation modules*, Preprint (June 1998).
[15] G.E. Bredon, *Equivariant Cohomology Theories*, Lecture Notes in Math., Vol. 34, Springer, Berlin (1967).
[16] H. Cartan and S. Eilenberg, *Homological Algebra*, Princeton University Press (1956).
[17] A. Chin, *The cohomology rings of finite groups with semi-dihedral Sylow* 2*-subgroups*, Bull. Austral. Math. Soc. **51** (1995), 421–432.
[18] A. Chin, *Some rings of invariants at the prime two*, Internat. J. Algebra Comput.
[19] C.W. Curtis and I. Reiner, *Methods of Representation Theory* II, Wiley (1987).
[20] T. tom Dieck, *Transformation Groups and Representation Theory*, Lecture Notes in Math., Vol. 766, Springer, Berlin (1979).
[21] T. tom Dieck, *Transformation Groups*, Walter de Gruyter, Berlin (1987).
[22] R. Dipper and J. Du, *Harish-Chandra vertices*, J. Reine Angew. Math. **437** (1993), 101–130.
[23] A.W.M. Dress, *A characterization of solvable groups*, Math. Z. **110** (1969), 213–217.
[24] A.W.M. Dress, *Notes on the theory of representations of finite groups*, Duplicated notes, Bielefeld (1971).
[25] A.W.M. Dress, *Contributions to the theory of induced representations*, Algebraic K-Theory II, H. Bass, ed., Lecture Notes in Math., Vol. 342, Springer, Berlin (1973), 183–240.
[26] A.W.M. Dress, *Modules with trivial source, modular monomial representations and a modular version of Brauer's induction theorem*, Abh. Math. Sem. Univ. Hamburg **44** (1975), 101–109.
[27] W.G. Dwyer, *Sharp homology decompositions for classifying spaces of finite groups*, Proc. Symp. Pure Math. **63** (1998), 197–220.
[28] P. Gabriel and M. Zisman, *Calculus of Fractions and Homotopy Theory*, Ergeb. Math. Grenzgeb., B. 35, Springer, Berlin (1967).
[29] D. Gluck, *Idempotent formula for the Burnside algebra with applications to the p-subgroup simplicial complex*, Illinois J. Math. **25** (1981), 63–67.
[30] J.A. Green, *Axiomatic representation theory for finite groups*, J. Pure Appl. Algebra **1** (1971), 41–77.
[31] J.P.C. Greenlees, *Some remarks on projective Mackey functors*, J. Pure Appl. Algebra **81** (1992), 17–38.
[32] J.P.C. Greenlees and J.P. May, *Some remarks on the structure of Mackey functors*, Proc. Amer. Math. Soc. **115** (1992), 237–243.
[33] I. Hambleton, L.R. Taylor and B. Williams, *Induction theory*, MSRI Report (1990).
[34] I. Hambleton, L.R. Taylor and B. Williams, *Detection theorems for K-theory and L-theory*, J. Pure Appl. Algebra **63** (1990), 247–299.
[35] K. Kato, *A generalization of local class field theory using K-groups*, J. Fac. Sci. Univ. Tokyo **27**, 603–683.
[36] T.-Y. Lam, *Induction theorems for Grothendieck groups and Whitehead groups of finite groups*, Ann. École Norm. Sup. **1** (1968), 91–148.
[37] L.G. Lewis, Jr., *The theory of Green functors*, Unpublished notes (1980).
[38] L.G. Lewis, Jr., *The category of Mackey functors for a compact Lie group*, Proc. Symp. Pure Math. **63** (1998), 301–354.
[39] L.G. Lewis, Jr., J.P. May and J.E. McClure, *Classifying G-spaces and the Segal conjecture*, Current Trends in Algebraic Topology, R.M. Kane et al., eds, CMS Conference Proc., Vol. 2 (1982), 165–179.
[40] H. Lindner, *A remark on Mackey functors*, Manuscripta Math. **18** (1976), 273–278.
[41] F. Luca, *The algebra of Green and Mackey functors*, PhD thesis, University of Alaska, Fairbanks (1996).

[42] F. Luca, *The defect of the completion of a Green functor*, Proc. Symp. Pure Math. **63** (1998), 369–382.
[43] J. Martino, *Classifying spaces and their maps*, Homotopy Theory and its Applications, Adem, Milgram and Ravenal, eds, Contemp. Math., Vol. 188 (1995), 153–190.
[44] J. Martino and S. Priddy, *The complete stable splitting for the classifying space of a finite group*, Topology **31** (1992), 143–156.
[45] J. Martino and S. Priddy, *Classification of BG for groups with dihedral or quaternion Sylow 2-subgroups*, J. Pure Appl. Algebra **73** (1991), 13–21.
[46] J.P. May, *Stable maps between classifying spaces*, Conference on Algebraic Topology in honour of P. Hilton, R. Piccinini and D. Sjerve, eds., Contemp. Math., Vol. 37, Amer. Math. Soc. (1985), 121–129.
[47] N. Minami, *Hecke algebras and cohomotopical Mackey functors*, Trans. Amer. Math. Soc. **351** (1999), 4481–4513.
[48] F. Oda, *On defect groups of the Mackey algebras for finite groups*, Proc. Symp. Pure Math. **63** (1998), 453–460.
[49] R. Oliver, *Whitehead Groups of Finite Groups*, LMS Lecture Notes, Vol. 132, Cambridge University Press (1988).
[50] K.W. Roggenkamp and L.L. Scott, *Hecke actions on Picard groups*, J. Pure Appl. Algebra **26** (1982), 85–100.
[51] H. Sasaki, *Green correspondence and transfer theorems of Wielandt type for G-functors*, J. Algebra **79** (1982), 98–120.
[52] P. Symonds, *A splitting principle for group representations*, Comment. Math. Helv. **66** (1991), 169–184.
[53] P. Symonds, *Functors for a block of a group algebra*, J. Algebra **164** (1994), 576–585.
[54] D. Tambara, *Homological properties of the endomorphism ring of certain permutation modules*, Osaka J. Math. **26** (1989), 807–828.
[55] D. Tambara, *On multiplicative transfer*, Comm. Algebra **21** (1993), 1393–1420.
[56] J. Thévenaz, *Some remarks on G-functors and the Brauer morphism*, J. Reine Angew. Math. **384** (1988), 24–56.
[57] J. Thévenaz, *A visit to the kingdom of the Mackey functors*, Bayreuth. Math. Schr. **33** (1990), 215–241.
[58] J. Thévenaz, *Defect theory for maximal ideals and simple functors*, J. Algebra **140** (1991), 426–483.
[59] J. Thévenaz, *G-Algebras and Modular Representation Theory*, Oxford University Press (1995).
[60] J. Thévenaz and P.J. Webb, *Simple Mackey Functors*, Proc. of 2nd International Group Theory Conference, Bressanone (1989), Rend. Circ. Mat. Palermo Suppl. **23** (1990), 299–319.
[61] J. Thévenaz and P.J. Webb, *A Mackey functor version of a conjecture of Alperin*, Astérisque **181–182** (1990), 263–272.
[62] J. Thévenaz and P.J. Webb, *The structure of Mackey functors*, Trans. Amer. Math. Soc. **347** (1995), 1865–1961.
[63] P.J. Webb, *A local method in group cohomology*, Comment. Math. Helv. **62** (1987), 135–167.
[64] P.J. Webb, *Subgroup complexes*, The Arcata Conference on Representations of Finite Groups, P. Fong, ed., AMS Proc. of Symposia in Pure Mathematics, Vol. 47 (1987).
[65] P.J. Webb, *A split exact sequence of Mackey functors*, Comment. Math. Helv. **66** (1991), 34–69.
[66] P.J. Webb, *Two classifications of simple Mackey functors with applications to group cohomology and the decomposition of classifying spaces*, J. Pure Appl. Algebra **88** (1993), 265–304.
[67] A. Wiedemann, *Elementary construction of the quiver of the Mackey algebra for groups with cyclic normal p-Sylow subgroup*, J. Algebra **150** (1992), 296–307.
[68] T. Yoshida, *Idempotents of Burnside rings and Dress induction theorem*, J. Algebra **80** (1983), 90–105.
[69] T. Yoshida, *On G-functors (II): Hecke operators and G-functors*, J. Math. Soc. Japan **35** (1983), 179–190.
[70] T. Yoshida, *Idempotents and transfer theorems of Burnside rings, character rings and span rings*, Algebraic and Topological Theories (to the memory of T. Miyata) (1985), 589–615.

Subject Index