Managing Editor

M. HAZEWINKEL, Amsterdam

Editorial Board

M. ARTIN, Cambridge
M. NAGATA, Okayama
C. PROCESI, Rome
R.G. SWAN, Chicago
P.M. COHN, London
A. DRESS, Bielefeld
J. TITS, Paris
N.J.A. SLOANE, Murray Hill
C. FAITH, New Brunswick
S.I. AD’YAN, Moscow
Y. IHARA, Tokyo
L. SMALL, San Diego
E. MANES, Amherst
I.G. MACDONALD, Oxford
M. MARCUS, Santa Barbara
L.A. BOKUT’, Novosibirsk
Basic philosophy

Algebra, as we know it today (2005), consists of a great many 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 other convenient designation. Even a nonspecialist is quite likely to encounter most of these, either somewhere in the published literature in the form of an idea, definition, theorem, algorithm, . . . somewhere, or to hear about them, often in somewhat vague terms, and to feel the need for more information. In such a case, if the concept relates to algebra, then one should be able to find something in this Handbook; at least enough to judge whether it is worth the trouble to try to find out more. In addition to the primary information the numerous references to important articles, books, or lecture notes should help the reader find out more.

As a further tool the index is perhaps more extensive than usual, and is definitely not limited to definitions, (famous) named theorems and the like.

For the purposes of this Handbook, “algebra” is more or less defined as the union of the following areas of the Mathematics Subject Classification Scheme:

- 20 (Group theory)
- 19 ($K$-theory; this 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 category 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 definitely not a topic that will be dealt with in this Handbook; there will, hopefully, one day be a separate Handbook on that topic)
- 12 (Field theory and polynomials)
- 11 The part of that also used to be classified under 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 (1992), we expected to cover the whole field in a systematic way. Volume 1 would be devoted to what is now called Section 1 (see below), Volume 2 to Section 2, and so on. A quite detailed and comprehensive plan was made in terms of topics that 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. Instead there is now a dynamic evolving plan. This also permits to take new developments into account.

Chapters are still by invitation only according to the then current version of the plan, but the various chapters are published as they arrive, allowing for faster publication. Thus in this Volume 4 of the Handbook of Algebra the reader will find contributions from 5 sections.

As the plan is dynamic suggestions from users, both as to topics that could or should be covered, and authors, are most welcome and will be given serious consideration by the board and editor.

The list of sections looks as follows:

Section 1: Linear algebra. Fields. Algebraic number theory
Section 2: Category theory. Homological and homotopical algebra. Methods from logic (algebraic model theory)
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 the detailed plan (2005 version), the reader is referred to the Outline of the Series following this 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. Indeed, the treatments of the various topics will be much too dense and professional for that. Basically, the level should be graduate and up, and such material as can be found in P.M. Cohn’s three volume textbook ‘Algebra’ (Wiley) should, as a rule, be assumed known. The most important function of the articles in this Handbook is to provide professional mathematicians working in a different area with a sufficiency of information on the topic in question if and when it is needed.

Each of the chapters 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:
Preface

The present

Volume 1 appeared in 1995 (copyright 1996), Volume 2 in 2000, Volume 3 in 2003. Volume 5 is planned for 2006. Thereafter, we aim at one volume every two years (or better).

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, CDROM (DVD)-based version certainly exist but the realization is still a nontrivial number of years in the future.

Kvoseliai, July 2005

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)

Elias Canetti
This page intentionally left blank
Outline of the Series
(as of July 2005)

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 the 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 and does change in response to new developments and availability of authors.

The new policy, starting with Volume 2, is 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 (i.e. chapters) as they become available from the invited authors.

The coding below is by style and is as follows.

– Author(s) in bold, followed by chapter title: articles (chapters) that have been received and are published or are being published in this volume.
– Chapter title in italic: chapters that are being written.
– Chapter title in plain text: topics that should be covered but for which no author has yet been definitely contracted.

Chapters that are included in Volumes 1–4 have a (x; yy pp.) after them, where ‘x’ is the volume number and ‘yy’ is the number of pages.

Compared to the plan that appeared in Volume 1 the section on “Representation and invariant theory” has been thoroughly revised. The changes of this current version compared to the one in Volume 2 (2000) and Volume 3 (2003) are relatively minor: mostly the addition of quite a few topics.

Editorial set-up
Managing editor: M. Hazewinkel.

Planned publishing schedule (as of July 2005)
1996: Volume 1 (published)
2001: Volume 2 (published)
2003: Volume 3 (published)
Outline of the series

2005: Volume 4 (last quarter)
Further volumes at the rate of one every year.

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.)
   - Correction to the chapter by L. Rodman, Matrix functions (3; 1 p.)
   - J.A. Hermida-Alonso, Linear algebra over commutative rings (3, 59 pp.)

Linear inequalities (also involving matrices)
Orderings (partial and total) on vectors and matrices
Positive matrices
Structured matrices such as Toeplitz and Hankel
Integral matrices, Matrices over other rings and fields
Quasideterminants, and determinants over noncommutative fields
Nonnegative matrices, positive definite matrices, and doubly nonnegative matrices
Linear algebra over skew 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, 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, 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)
Dessins d’enfants
Hopf Galois theory
Outline of the series

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, 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, I. Moerdijk, Topos theory (1; 28 pp.)
   R. Street, Categorical structures (1; 50 pp.)
   B.I. Plotkin, Algebra, categories and databases (2; 68 pp.)
   P.S. Scott, Some aspects of categories in computer science (2; 73 pp.)
   E. Manes, Monads of sets (3; 87 pp.)
   Operads

   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
   A. Kuku, Classical algebraic K-theory: the functors $K_0$, $K_1$, $K_2$ (3; 40 pp.)
   A. Kuku, Algebraic K-theory: the higher K-functors (4; 72 pp.)
   Grothendieck groups
   $K_2$ and symbols
   KK-theory and EXT
   Hilbert C*-modules
   Index theory for elliptic operators over C* algebras
   Simplicial algebraic K-theory
   Chern character in algebraic K-theory
   Noncommutative differential geometry
   K-theory of noncommutative rings
   Algebraic L-theory
Cyclic cohomology
Asymptotic morphisms and $E$-theory
Hirzebruch formulae

D. Model Theoretic Algebra

(See also P.C. Eklof, Whitehead modules, in Section 3B)
M. Prest, Model theory for algebra (3; 28 pp.)
M. Prest, Model theory and modules (3; 27 pp.)
Logical properties of fields and applications
Recursive algebras
Logical properties of Boolean algebras
F.O. Wagner, Stable groups (2; 40 pp.)
The Ax–Ershov–Kochen theorem and its relatives and applications

E. Rings up to Homotopy

Rings up to homotopy
Simplicial algebras

Section 3. Commutative and associative rings and algebras

A. Commutative Rings and Algebras

(See also C. Faith, Coherent rings and annihilator conditions in matrix and polynomial rings, in Section 3B)
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
R.H. Villareal, Monomial algebras and polyhedral geometry (3; 58 pp.)

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.)
Outline of the series

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
S.K. Sehgal, Group rings and algebras (3; 87 pp.)
Dimension theory
V. Bavula, Filter dimension (4; 29 pp.)
A. Facchini, The Krull–Schmidt theorem (3; 41 pp.)
Duality. Morita-duality
Commutants of differential operators
E.E. Enochs, Flat covers (3; 14 pp.)
C. Faith, Coherent rings and annihilator conditions in matrix and polynomial rings
(3; 30 pp.)
Rings of differential operators
Graded and filtered rings and modules (also commutative)
P.C. Eklof, Whitehead modules (3; 25 pp.)
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)
Finite rings and modules
T.Y. Lam, Hamilton’s quaternions (3; 26 pp.)
A.A. Tuganbaev, Semiregular, weakly regular, and π-regular rings (3; 22 pp.)
Hamiltonian algebras
A.A. Tuganbaev, Max rings and V-rings (3; 20 pp.)
Algebraic asymptotics
(See also “Freeness theorems in groups and rings and Lie algebras” in Section 5A)

C. Coalgebras

W. Michaelis, Coassociative coalgebras (3; 202 pp.)
Co-Lie-algebras
D. Deformation Theory of Rings and Algebras (Including Lie Algebras)

Deformation theory of rings and algebras (general)

Yu. Khakimdzanov, Varieties of Lie algebras (2; 31 pp.)
Deformation theoretic quantization

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
A. Pultr, Frames (3; 67 pp.)
Quantales

B. Boolean Algebras

C. Universal Algebra

Universal algebra

D. Varieties of Algebras, Groups, ...

(See also Yu. Khakimdzanov, Varieties of Lie algebras, in Section 3D)
V.A. Artamonov, Varieties of algebras (2; 29 pp.)
Varieties of groups
V.A. Artamonov, Quasivarieties (3; 23 pp.)
Varieties of semigroups

E. Lie Algebras

Yu.A. Bahturin, M.V. Zaitsev, A.A. Mikhailov, Infinite-dimensional Lie superalgebras (2; 34 pp.)
General structure theory
Ch. Reutenauer, Free Lie algebras (3; 17 pp.)
Classification theory of semisimple Lie algebras over \( \mathbb{R} \) and \( \mathbb{C} \)
The exceptional Lie algebras
M. Goze, Y. Khakimdjanov, Nilpotent and solvable Lie algebras (2; 47 pp.)
Universal enveloping algebras
Modular (ss) Lie algebras (including classification)
Infinite-dimensional Lie algebras (general)
Kac–Moody Lie algebras
Affine Lie algebras and Lie super algebras and their representations
Finitary Lie algebras
Standard bases
A.I. Molev, Gelfand–Tsetlin bases for classical Lie algebras (4; 62 pp.)
Kostka polynomials

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

M. Cohen, S. Gelaki, S. Westreich, Hopf algebras (4; 67 pp.)
Classification of pointed Hopf algebras
Recursive sequences from the Hopf algebra and coalgebra points of view
Quantum groups (general)
A.I. Molev, Yangians and their applications (3; 53 pp.)
Formal groups
p-divisible groups
F. Patras, Lambda-rings (3; 26 pp.)
Ordered and lattice-ordered groups, rings and algebras
Rings and algebras with involution. C*-algebras
A.B. Levin, Difference algebra (4; 94 pp.)
Differential algebra
Ordered fields
Hypergroups
Stratified algebras
Combinatorial Hopf algebras
Symmetric functions
Special functions and q-special functions, one and two variable case
Quantum groups and multiparameter q-special functions
Hopf algebras of trees and renormalization theory
Noncommutative geometry à la Connes
Noncommutative geometry from the algebraic point of view
Noncommutative geometry from the categorical point of view
Solomon descent algebras

I. Witt Vectors

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

Section 5. Groups and semigroups

A. Groups

A.V. Mikhalev, A.P. Mishina, Infinite Abelian groups: methods and results (2; 36 pp.)
Simple groups, sporadic groups
Representations of the finite simple groups
Diagram methods in group theory
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

M. Geck, G. Malle, Reflection groups (4; 47 pp.)
M.C. Tamburini, M. Vsemirnov, Hurwitz groups and Hurwitz generation (4; 42 pp.)
Groups with BN-pair, Tits buildings, ...
Groups and (finite combinatorial) geometry
“Additive” group theory
Probabilistic techniques and results in group theory
V.V. Vershinin, Braids, their properties and generalizations (4; 39 pp.)
L. Bartholdi, R.I. Grigorchuk, Z. Šunič, Branch groups (3; 124 pp.)
Frobenius groups
Just infinite groups
V.I. Senashov, Groups with finiteness conditions (4; 27 pp.)
Automorphism groups of groups
Automorphism groups of algebras and rings
Freeness theorems in groups and rings and Lie algebras
Groups with prescribed systems of subgroups
(see also “Groups and semigroups of automata transformations” in Section 5B)
Automatic groups
Groups with minimality and maximality conditions (school of Chernikov)
Lattice-ordered groups
Linearly and totally ordered groups
Finitary groups
Random groups
Hyperbolic groups

B. Semigroups
Semigroup theory. Ideals, radicals, structure theory
Semigroups and automata theory and linguistics
Groups and semigroups of automata transformations
Cohomology of semigroups
Outline of the series

C. Algebraic Formal Language Theory. Combinatorics of Words

D. Loops, Quasigroups, Heaps, …

Quasigroups in combinatorics

E. Combinatorial Group Theory and Topology

(See also “Diagram methods in group theory” in Section 5A)

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 topological groups
- Representation theory of locally compact topological groups
- Representation theory of \( SL_2(\mathbb{R}) \), …
- 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
  (See also A.I. Molev, Gelfand–Tsetlin bases for classical Lie algebras, in Section 4E)
- 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 $GL_\infty$
Metaplectic representation theory

D. Representation Theory of Algebras
Representations of rings and algebras by sections of sheaves
Representation theory of algebras (Quivers, Auslander–Reiten sequences, almost
split sequences, …)
Quivers and their representations
Tame algebras
Ringel–Hall algebras

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

H. Hecke Algebras
Hecke–Iwahori algebras

I. Invariant Theory


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
Combinatorial sums and counting algebraic structures
Groebner bases and their applications
Section 8. Applied algebra

Section 9. History of algebra

(See also K.T. Lam, Hamilton’s quaternions, in Section 3B)
History of coalgebras and Hopf algebras
Development of algebra in the 19-th century
This page intentionally left blank
Contents

Preface v

Outline of the Series ix

List of Contributors xxiii

Section 2C. Algebraic K-theory 1
   A. Kuku, Higher algebraic K-theory 3

Section 3B. Associative Rings and Algebras 75
   V. Bavula, Filter dimension 77

Section 4E. Lie Algebras 107
   A.I. Molev, Gelfand–Tsetlin bases for classical Lie algebras 109

Section 4H. Rings and Algebras with Additional Structure 171
   M. Cohen, S. Gelaki and S. Westreich, Hopf algebras 173
   A.B. Levin, Difference algebra 241

Section 5A. Groups and Semigroups 335
   M. Geck and G. Malle, Reflection groups 337
   M.C. Tamburini and M. Vsemirnov, Hurwitz groups and Hurwitz generation 385
   V.V. Vershikin, Braids, their properties and generalizations 427
   V.I. Senashov, Groups with finiteness conditions 467

Subject Index 495
This page intentionally left blank
List of Contributors

V. Bavula, University of Sheffield, Sheffield, e-mail: v.bavula@sheffield.ac.uk
M. Cohen, Ben Gurion University of the Negev, Beer Sheva, e-mail: mia@cs.bgu.ac.il
M. Geck, King’s College, Aberdeen University, Aberdeen, e-mail: geck@maths.abdn.ac.uk
S. Gelaki, Technion, Haifa, e-mail: gelaki@tx.technion.ac.il
A. Kuku, Institute for Advanced Study, Princeton, NJ, e-mail: kukuao@muohio.edu
A.B. Levin, The Catholic University of America, Washington, DC, e-mail: Levin@cua.edu
G. Malle, Universität Kaiserslautern, Kaiserslautern, e-mail: malle@mathematik.uni-kl.de
A.I. Molev, University of Sydney, Sydney, e-mail: alexm@maths.usyd.edu.au
V.I. Senashov, Institute of Computational Modelling of Siberian Division of Russian Academy of Sciences, Krasnoyarsk, e-mail: sen@icm.krasn.ru
M.C. Tamburini, Università Cattolica del Sacro Cuore, Brescia, e-mail: c.tamburini@dmf.unicatt.it
V.V. Vershinin, Université Montpellier II, Montpellier, e-mail: vershini@math.univ-montp2.fr
Sobolev Institute of Mathematics, Novosibirsk, e-mail: versh@math.nsc.ru
M. Vsemirnov, St. Petersburg Division of Steklov Institute of Mathematics, St. Petersburg, e-mail: vsemir@pdmi.ras.ru
S. Westreich, Bar-Ilan University, Ramat-Gan, e-mail: swestric@mail.biu.ac.il
This page intentionally left blank
Section 2C
Algebraic $K$-theory
This page intentionally left blank
Higher Algebraic $K$-Theory

Aderemi Kuku

Institute for Advanced Study, Princeton, NJ 08540, USA
E-mail: kukuao@muohio.edu

Contents
Introduction ........................................ 5
1. Simplicial objects, classifying spaces, and spectra ........................................ 7
   1.1. Simplicial objects and classifying spaces ........................................ 7
   1.2. Spectra – brief introduction ........................................ 10
2. Definitions of and relations between several higher algebraic $K$-theories (for rings) .... 11
   2.1. $K^Q_n$ – the $K$-theory of Quillen ........................................ 11
   2.2. $K^S_n$ – the $K$-theory of Swan and Gersten ................................ 12
   2.3. $K^n_{n}$ – the $K$-theory of Karoubi and Villamayor ............................... 13
   2.4. $K^V_n$ – Volodin $K$-theory ......................................... 14
   2.5. $K^M_n$ – Milnor $K$-theory ........................................ 14
3. Higher $K$-theory of exact, symmetric monoidal and Waldhausen categories ............ 15
   3.1. Higher $K$-theory of exact categories – definitions and examples .............. 15
   3.2. Higher $K$-theory of symmetric monoidal categories – definitions and examples .... 18
   3.3. Higher $K$-theory of Waldhausen categories – definitions and examples ....... 20
4. Some fundamental results and exact sequences in higher $K$-theory ..................... 24
   4.1. Resolution theorem ............................................ 24
   4.2. Additivity theorem (for exact and Waldhausen categories) ....................... 24
   4.3. Devissage ................................................. 25
   4.4. Localization .................................................. 26
   4.5. Fundamental theorem for higher $K$-theory .................................... 30
   4.6. Some exact sequences in the $K$-theory of Waldhausen categories .......... 31
   4.7. Excision; relative and Mayer–Vietoris sequences ................................ 32
5. Higher $K$-theory and connections to Galois, étale and motivic cohomology theories ... 34
   5.1. Higher $K$-theory of fields ....................................... 34
   5.2. Galois cohomology ............................................. 36
   5.3. Zariski and étale cohomologies ...................................... 38
   5.4. Motivic cohomology ............................................ 41
   5.5. Connections to Bloch’s higher Chow groups ................................... 43
6. Higher $K$-theory of rings of integers in local and global fields ........................ 44
   6.1. Some earlier general results on the higher $K$-theory of ring of integers in global fields .... 44
   6.2. Étale and motivic Chern characters ....................................... 46
   6.3. Higher $K$-theory and zeta functions ...................................... 51
   6.4. Higher $K$-theory of $\mathbb{Z}$ ...................................... 53
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.</td>
<td>Higher $K$-theory of orders, group-rings and modules over 'EI'-categories</td>
<td>55</td>
</tr>
<tr>
<td>7.1.</td>
<td>Higher $K$-theory of orders and group-rings</td>
<td>55</td>
</tr>
<tr>
<td>7.2.</td>
<td>Higher class groups of orders and group-rings</td>
<td>60</td>
</tr>
<tr>
<td>7.3.</td>
<td>Profinite higher $K$-theory of orders and group-rings</td>
<td>62</td>
</tr>
<tr>
<td>7.4.</td>
<td>Higher $K$-theory of modules over EI categories</td>
<td>63</td>
</tr>
<tr>
<td>8.</td>
<td>Equivariant higher algebraic $K$-theory together with relative generalizations</td>
<td>65</td>
</tr>
<tr>
<td>8.1.</td>
<td>Equivariant higher algebraic $K$-theory</td>
<td>65</td>
</tr>
<tr>
<td>8.2.</td>
<td>Relative equivariant higher algebraic $K$-theory</td>
<td>66</td>
</tr>
<tr>
<td>8.3.</td>
<td>Interpretation in terms of group-rings</td>
<td>67</td>
</tr>
<tr>
<td>8.4.</td>
<td>Some applications</td>
<td>68</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td></td>
<td>70</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>70</td>
</tr>
</tbody>
</table>
Introduction

This chapter is a sequel to “Classical Algebraic K-Theory: The Functors $K_0$, $K_1$, $K_2$” published in Volume 3 of Handbook of Algebra [79]. The unexplained notions in this chapter are those of [79]. Here, we shall provide higher-dimensional analogues to quite a number of results in [79].

As already observed in [79], the functor $K_0$ was defined by A. Grothendieck, $K_1$ by H. Bass and $K_2$ by J. Milnor. The definition of $K_2$ by Milnor in 1967 inspired many mathematicians to search for higher $K$-groups and the next five years (1967–1972) witnessed a lot of vigorous activity in this respect. During this period, several higher $K$-theories were proposed; notably by D. Quillen, [114,111], S. Gersten, [36], R.G. Swan, [138], I. Volodin, [151], J. Milnor, [105], and F. Keune, [64]. These theories are briefly reviewed in Section 2 with connections between them highlighted. By far the most successful among the theories are those of D. Quillen. Hence, a substantial part of this chapter is devoted to developments of the subject arising from Quillen’s work.

We now review the contents of this chapter. Section 1 is a brief discussion of some of the central notions in most constructions of higher $K$-theories – simplicial objects, classifying spaces and spectra (see 1.1, 1.2).

In Section 2, we define Quillen $K$-theory, $K^Q_n$ (2.1); $K$-theory of Gersten and Swan, $K^S_n$ (2.2); $K$-theory of Karoubi and Villamayor $K^{k-v}_n$ (2.3); Volodin $K$-theory, $K^V_n$ (2.4); and Milnor $K$-theory, $K^M_n$ (2.5) – also highlighting some connections between them, e.g., that $K^Q_n(A)$ coincides with $K^V_n(A)$ and $K^S_n(A)$ while $K^Q_n(A)$ coincides with $K^{k-v}_n(A)$ when $A$ is regular.

In Section 3, we define higher $K$-theory of exact, symmetric monoidal and Waldhausen categories, providing copious examples in each situation (see 3.1, 3.2, 3.3). Thus we discuss for exact categories, higher $K$-theory of rings and schemes; mod-$m$ and profinite higher $K$-theory; equivariant higher $K$-theory, etc. For Waldhausen categories, for instance, we discuss $K$-theory of perfect complexes and stable derived categories (see 3.3.10).

In Section 4, we highlight, with copious examples, some fundamental results in higher $K$-theory, most of which have classical analogues at the zero-dimensional level. The topics covered include the resolution theorem for exact categories (4.1); the additivity theorem for exact and Waldhausen categories (4.2), the devissage theorem (4.3); localization sequences (4.4) leading to the Gersten conjecture and fundamental theorems for higher $K$- and $G$-theories (4.4.3 and 4.5). We also discuss Waldhausen’s fibration sequence, localization sequence for Waldhausen’s $K$-theory and a long exact sequence which realizes the cofibre of the Cartan maps as $K$-theory of a Waldhausen category. Finally, we discuss excision, Mayer–Vietoris sequences and long exact sequence associated to an ideal.

In Section 5, we define Galois, étale and motivic cohomologies and discuss their interconnections as well as connections with $K$-theory. We discuss in 5.2 the Bloch–Kato conjecture including parts of it earlier proved – the Milnor conjecture and the Merkurjev–Suslin theorem. In 5.3 we discuss Zariski and étale cohomology as well as connections between them (see 5.3.6). Next we define motivic cohomology which we identify with Lichtenbaum (étale) cohomology groups for smooth $k$-varieties as well as the connection
of this to a special case of Bloch–Kato conjecture 5.4.8. Next we discuss Bloch’s higher Chow groups and their connections with $K$-theory and motivic cohomology.

In Section 6, we discuss higher $K$-theory of rings of integers in local and global fields. In 6.2 we define étale Chern characters of Soulé (6.2.3) with the observation that there are alternative approaches through “anti-Chern” characters defined by B. Kahn, [56], and maps from étale $K$-theory to étale cohomology due to Dwyer and Friedlander, [26]. We discuss the Quillen–Lichtenbaum conjecture and record for all $n \geq 2$ computations of $K_n(O_S)(2)$, $K_n(O_S)(l)$ in terms of étale cohomology groups where $O_S$ is a ring of integers in number field $F$ as well as $K_n(O_S)$ when $F$ is totally imaginary. We also briefly discuss the motivic Chern characters of Pushin used to identify $K$-groups of number fields and their integers with motivic cohomology groups (see 6.2.18). In 6.3, we treat higher $K$-theory and zeta functions including the Lichtenbaum conjecture, Wiles theorem and their consequences. Finally we review in 6.4 some more explicit computations of $K_n(\mathbb{Z})$.

Section 7 deals with the higher $K$-theory of orders, group rings and modules over EI categories. In 7.1, we review some finiteness results due to Kuku, e.g.: If $F$ is a number field with integers $O_F$ and $A$ any $O_F$-order in a semi-simple $F$-algebra, then for all $n \geq 1$, $K_n(A)$, $G_n(A)$ are finitely generated, $SK_n(A)$, $SG_n(A)$ are finite and rank $K_n(A) = \text{rank} G_n(A)$ (see 7.1.4, 7.1.6 and 7.1.11). We also discuss the result due to R. Laubenbacher and D. Webb that $SG_n(O_FG) = 0$ for all $n \geq 1$ (see 7.1.8) as well as the result of Kuku and Tang that for all $n \geq 1$, $G_n(O_FV)$ is finitely generated where $V$ is virtually infinite cyclic group (see 7.1.10). We also exhibit D. Webb’s decomposition of $G_n(RG)$, $R$ a Noetherian ring and $G$ finite Abelian group as well as extensions of the decomposition to some non-Abelian groups, e.g., quaternion and dihedral groups.

Next we review in 7.2 results on higher class groups $Cl_n(A)$, $n \geq 0$, for orders. First we observe Kuku’s result that $Cl_n(A)$ is finite for all $n \geq 1$, as well as a result due to Laubenbacher and Kolster that the only $p$-torsion possible in odd-dimensional class groups $Cl_{2n-1}(A)$ are for primes $p$ lying below prime ideals $q$ for which $\hat{A}_q$ are not maximal. An analogous result due to Guo and Kuku for even-dimensional class groups $Cl_{2n}(A)$ is given in 7.2.9 for Eichler orders in quaternion $F$-algebras and hereditary orders in semi-simple $F$-algebras. In 7.3, we discuss Kuku’s results on profinite $K$-theory of orders and group-rings providing several $l$-completeness results for orders in algebras over number fields and $p$-adic fields as well as showing that for $p$-adic orders $A$, $G_n(A)_l$, $K_n(\Sigma)_l$ are finite groups if $l \neq p$ and that $K_n(O_FG)_l$ is also finite for any finite group $G$. In 7.4, we exhibit several finiteness results on higher $K$-theory of modules over ‘EI’ categories.

The last Section 8 deals with equivariant higher algebraic $K$-theory together with relative generalizations – for finite group action – due to Dress and Kuku with the observation that there are analogous theories for profinite groups and compact Lie group actions due to Kuku, [70,77]. Time and space prevented us from including the latter two cases. We also remark that K. Shimakawa, [127], provided a $G$-spectrum formulation of the absolute part of the theory discussed in Section 8, but again time and space has prevented us from going into this.
1. Simplicial objects, classifying spaces, and spectra

In this opening section, we briefly review some of the central notions in the construction of higher \(K\)-theories.

1.1. Simplicial objects and classifying spaces

1.1.1. Definition. Let \(\Delta\) be the category defined as follows: \(\text{ob}(\Delta) = \{0 < 1 < \cdots < n\}\). The set \(\text{Hom}_\Delta(m, n)\) of morphisms from \(m\) to \(n\) consists of maps \(f : m \to n\) such that \(f(i) \leq f(j)\) for \(i < j\).

Let \(\mathcal{A}\) be any category. A simplicial object in \(\mathcal{A}\) is a contravariant functor \(X : \Delta \to \mathcal{A}\) where we write \(X_n\) for \(X(n)\). Thus, a simplicial set (resp. group; resp. ring; resp. space, etc.) is a simplicial object in the category of sets (resp. groups; resp. rings; resp. spaces, etc.). A co-simplicial object is a covariant functor \(X : \Delta \to \mathcal{A}\).

Equivalently one could define a simplicial object in a category \(\mathcal{A}\) as a set of objects \(X_n(\geq 0)\) in \(\mathcal{A}\) and a set of morphisms \(\delta_i : X_n \to X_{n-1}\) (\(0 \leq i \leq n\)) called face maps as well as a set of morphisms \(s_j : X_n \to X_n+1\) (\(0 \leq j \leq n\)) called degeneracy maps satisfying certain “simplicial identities” – [165, p. 256]. We shall denote the category of simplicial sets by \(S\) sets.

1.1.2. Definition. The geometric \(n\)-simplex is the topological space \(\hat{\Delta}_n = \{(x_0, x_1, \ldots, x_n) \in \mathbb{R}^{n+1} | 0 \leq x_i \leq 1 \forall i\text{ and } \Sigma x_i = 1\}\).

The functor \(\hat{\Delta} : \Delta \to \text{spaces given by } n \to \hat{\Delta}^n\) is a co-simplicial space.

1.1.3. Definition. Let \(X_s\) be a simplicial set. The geometric realization of \(X_s\) written \(|X_s|\) is defined by \(|X_s| := X \times_\Delta \hat{\Delta} = \bigcup_{n \geq 0} (X_n \times \hat{\Delta}^n)/\approx\) where the equivalence relation ‘\(\approx\)’ is generated by \((x, \varphi_n(y)) \approx (\varphi^*(x), y)\) for any \(x \in X_n, y \in Y_m,\) and \(\varphi : m \to n\) in \(\Delta\) and where \(X_n \times \hat{\Delta}^n\) is given the product topology and \(X_n\) is considered as a discrete space.

1.1.4. Examples/Remarks.

(i) Let \(T\) be a topological space, \(\text{Sing}_s T = \{\text{Sing}_n T\}\) where \(\text{Sing}_n T = \{\text{continuous maps } \hat{\Delta}^n \to T\}\). A map \(f : n \to m\) determines a linear map \(\hat{\Delta}^n \to \hat{\Delta}^m\) and hence induces a map \(\hat{f} : \text{Sing}_m T \to \text{Sing}_n T\). So, \(\text{Sing}_s T : \Delta \to \text{sets is a simplicial set.}\)

(ii) For any simplicial set \(X_s, |X_s|\) is a CW-complex with \(X_n\) in one-one correspondence with \(n\)-cells in \(|X_s|\).

(iii) For any simplicial sets \(X_s, Y_s, |X_s| \times |Y_s| \cong |X_s \times Y_s|\) where the product is such that \((X_s \times Y_s)_n = X_n \times Y_n\).

1.1.5. Definition. Let \(\mathcal{A}\) be a small category. The nerve of \(\mathcal{A}\) written \(N\mathcal{A}\) is the simplicial set whose \(n\)-simplices are diagrams

\[\mathcal{A}_n = \{A_0 \xrightarrow{f_1} A_2 \xrightarrow{\cdots} A_n\},\]
where the $A_i$ are $A$-objects and the $f_i$ are $A$-morphisms. The classifying space of $A$ is defined as $|N\mathcal{A}|$ and is denoted by $B\mathcal{A}$.

### 1.1.6. Properties of $B\mathcal{A}$.

(i) $B\mathcal{A}$ is a CW-complex whose $n$-cells are in one-one correspondence with the diagrams $A_n$ above (see 1.1.4(ii)).

(ii) From 1.1.4(iii), we have, for small categories $\mathcal{C}, \mathcal{D}$ (I) $B(\mathcal{C} \times \mathcal{D}) \cong B\mathcal{C} \times B\mathcal{D}$ where $B\mathcal{C} \times B\mathcal{D}$ is given the compactly generated topology (see [128]). In particular we have the homeomorphism (I) if either $B\mathcal{C}$ or $B\mathcal{D}$ is locally compact (see [128]).

(iii) Let $F, G$ be functors, $\mathcal{C} \to \mathcal{D}$ (where $\mathcal{C}, \mathcal{D}$ are small categories). A natural transformation of functors $\eta : F \to C$ induces a homotopy $B\mathcal{C} \times I \to B\mathcal{D}$ from $B\mathcal{C}$ to $B\mathcal{D}$.

(iv) If $F : \mathcal{C} \to \mathcal{D}$ has a left or right adjoint, then $F$ induces a homotopy equivalence.

(v) If $\mathcal{C}$ is a category with initial or final object, then $B\mathcal{C}$ is contractable.

### 1.1.7. Examples.

(i) A discrete group $G$ can be regarded as a category with one object $G$ whose morphisms can be identified with the elements of $G$.

The nerve of $G$, written $N_n(G)$ is defined as follows: $N_n(G) = G^n$, with face maps $\delta_i$ given by

\[
\delta_i(g_1, \ldots, g_n) = \begin{cases} 
(g_2, \ldots, g_n), & i = 0, \\
(g_1, \ldots, g_i g_{i+1}, \ldots, g_n), & 1 \leq i < n - 1, \\
(g_1, \ldots, g_{n-1}), & i = n - 1,
\end{cases}
\]

and degeneracies $s_i$ given by

\[
s_i(g_1, g_2, \ldots, g_n) = (g_1, g_i, 1, g_{i+1}, g_n).
\]

The classifying space $BG$ of $G$ is defined as $|N_n(G)|$ and it is a connected CW-complex characterized up to homotopy type by the property that $\pi_1(BG, *) = G$ and $\pi_n(BG, *) = 0$ for all $n > 0$ where $*$ is some basepoint of $BG$. Note that $BG$ has a universal covering space usually denoted by $EG$ (see [165]).

Note that the term classifying space of $G$ comes from the theory of fibre bundles. So, if $X$ is a finite cell complex, the set $[X, BG]$ of homotopy classes of maps $X \to BG$ gives a complete classification of the fibre bundles over $X$ with structure group $G$.

(ii) Let $G$ be a topological group (possibly discrete) and $X$ a topological $G$-space. The translation category $\underline{X}$ of $X$ is defined as follows: $\text{ob}(\underline{X}) =$ elements of $X$; $\text{Hom}_A(x, x') = \{g \in G \mid gx = x'\}$. Then the nerve of $\underline{X}$ is the simplicial space equal to $G^n \times X$ in dimension $n$. $B\underline{X} = \text{nerve of } X$ is the Borel space $EG \times X$ (see [94]).

(iii) Let $\mathcal{C}$ be a small category, $F : \mathcal{C} \to \text{Sets}$ a functor, then $\mathcal{C}_F$ is the category defined as follows

\[
\text{ob}\mathcal{C}_F = \{(C, x) \mid C \in \text{ob}\mathcal{C}, x \in F(C)\}.
\]
A morphism from \((C, x)\) to \((C', x')\) is a morphism \(f : C \to C'\) in \(\mathcal{C}\) such that \(f_a(x) = x'\).

The homotopy colimit of \(F\) is defined as \(\text{hocolim} F := BC_F\). This construction is also called the Bonsfield–Kan construction. If the functor \(F\) is trivial, we have \(BC_F = BC\).

1.1.8. Let \(C = C^{\text{top}}\) be a topological category (i.e. the objects in \(\mathcal{C}\) as well as \(\text{Hom}_{\mathcal{C}}(X, Y)\) \((X, Y \in \mathcal{C})\) are topological spaces). Then \(\text{NC}^{\text{top}}\) is a simplicial topological space and \(BC^{\text{top}} = |\text{NC}^{\text{top}}|\) the geometric realization of \(\text{NC}^{\text{top}}\). We could regard the identity map as a continuous function \(C^\delta \to C^{\text{top}}\) between topological categories and get an induced continuous maps \(BC^\delta \to BC^{\text{top}}\). (Here \(C^\delta\) is a discrete category, i.e. \(C\) with discrete topology on objects.)

1.1.9. **Examples.**

(i) Any topological group \(G = G^{\text{top}}\) is a topological category: \(\pi_1(BG^\delta) = G^\delta\), \(\pi_j(BG^\delta) = 0\) if \(j \neq 1\), \(\Omega BG^{\text{top}} = G^{\text{top}}\). Hence \(\pi_i(BG^{\text{top}}) = \pi_{i-1}G^{\text{top}}\) for \(i > 0\).

(ii) If \(A\) is a \(C^*-\)algebra with identity, then put \(G\) in (i) as \(G = \text{GL}(A) = \bigcup_n \text{GL}_n(A)\), and \(\pi_i(BGL(A)) = \pi_{i-1}(\text{GL}(A))\) which is by definition \(K_i^{\text{top}}(A)\) (higher ‘topological’ \(K\)-theory of \(A\)). \(K_0^{\text{top}}(A) = \pi_0(\text{GL}(A)) = K_0(\mathcal{P}(A)^\delta)\), the usual Grothendieck group of \(A\) and \(K_1(A) = \text{GL}_\infty(A)/\text{GL}(A)\) where \(\text{GL}(A)\) is the connected component of the identity in \(\text{GL}(A)\). In fact, Bott periodicity is satisfied, i.e. \(K_n(A) \cong K_{n+2}(A)\) for all \(n \geq 0\) (see [18]) and so, this theory is \(\mathbb{Z}_2\)-graded, having only \(K_0^{\text{top}}(A) = K_0(A)\) and \(K_1^{\text{top}}(A)\).

(iii) If \(A = \mathbb{C}\) in (ii) and we denote by \(U_n\) the unitary groups, then \(BU_n\) is homotopy equivalent to \(B\text{GL}_n(\mathbb{C})^{\text{top}}\) (because \(U_n\) is a deformation retract of \(\text{GL}_n(\mathbb{C})^{\text{top}}\)). Since \(\text{GL}_n(\mathbb{C})\) is connected, we have \(K_1^{\text{top}}(\mathbb{C}) = 0\), and \(K_0(\mathbb{C}) = K_0^{\text{top}}(\mathbb{C}) \cong \mathbb{Z}\).

1.1.10. **Remarks.**

(i) Given a simplicial object \(A = \{A_n\}\) in an Abelian category, there exists a chain complex \((C(A), d)\), i.e.

\[
C(A) : \cdots \to C_n \to C_{n-1} \to C_{n-2} \to \cdots,
\]

where \(C_n = A_n\) and \(d_n : C_n \to C_{n-1}\) is given by \(d_n = \delta_0 - \delta_1 + \cdots + (-1)^n \delta_n\).

(ii) If \(R\) is a ring, then there exists a functor

\[
\text{Sets} \to R\text{-Mod} : X \to R[X] = \text{free } R\text{-module on } X.
\]

If \(X = \{X_n\}\) is a simplicial set, then \(R[X] = \{R[X_n]\}\) is a simplicial \(R\)-module and \(H_\ast(X, R) := \text{homology of the chain complex associated to } R[X]\) (see (i) above).

Also \(H_\ast(X, R) = H_\ast(|X|, R)\), the singular homology of \(X\).
1.1.11. Let $G = \{G_n\}$ be a simplicial group with face maps $\delta_i : G_n \to G_{n-1}$ and degeneracies $s_i : G_n \to G_{n+1}$ ($0 \leq i \leq n$). Define $\pi_n G = H_n/d_{n+1}K_n$ where $H_n \subset K_n \subset G_n$ are defined by

$$K_n := \ker(\delta_0) \cap \cdots \cap \ker(\delta_{n-1})$$

and

$$H_n = K_n \cap (\ker(\delta_n)).$$

Say that $G$ is acyclic if $\pi_n(G) = 0 \ \forall \ n$.

We can regard a simplicial ring as a simplicial group using its additive structure and we say that a simplicial ring is acyclic if $\pi_n R = 0$ for all $n$.

1.1.12. A simplicial ring $R = \{R_i\}$ is said to be free if there exists a basis $B_n$ of $R_n$ as a free ring for all $n$ and $s_i(B_n) \subset B_{n+1}$ for all $i$ and all $n$.

A simplicial ring $R = \{R_i\}$ is said to have a unit if each $R_i$ has a unit and all $\delta_i$ and $s_i$ are unit preserving.

1.2. Spectra – brief introduction

1.2.1. REMARKS. The importance of spectra for this chapter has to do with the fact that higher $K$-groups are often expressed as homotopy groups of spectra $E = \{E_i\}$ whose spaces $E_i \approx \Omega^k E_{i+k}$ (for $k$ large) are infinite loop spaces. (It is usual to take $i = 0$ and consider $E_0$ as an infinite loop space.) Also to each spectrum can be associated generalized cohomology theory and vice-versa. Hence algebraic $K$-theory can always be endowed with the structure of a generalized cohomology theory. We shall come across these notions copiously in later sections.

1.2.2. DEFINITION. A spectrum $E = \{E_i\}, i \in \mathbb{Z}$, is a sequence of based spaces $E_n$ and based homeomorphisms $E_i \approx \Omega E_{i+1}(I)$. If we regard $E_i = 0$ for negative $i$, call $E$ a connective spectrum.

A map $f : E = \{E_i\} \to \{F_i\} = F$ of spectra is a sequence of based continuous maps strictly compatible with the given homeomorphism (I). The spectra form a category which we shall denote by $\text{Spectra}$.

1.2.3. From the adjunction isomorphism $[\Sigma X, Y] = [X, \Omega Y]$ for spaces $X, Y$, we have $\pi_n(\Omega E_i) \cong \pi_{n+1}(E_1)$, and so, we can define the homotopy groups of a connective spectrum $E$ as $\pi_n(E) = \pi_n(E) = \pi_{n+1}(E_1) = \cdots = \pi_{n+i}(E_i)$.

1.2.4. Each spectrum $E = \{E_n\}$ gives rise to an extraordinary cohomology theory $E^n$ in such a way that if $X_+$ is the space obtained from $X$ by adjoining a base point, $E^n(X) = [X_+, E_n]$ and conversely.
One can also associate to $E$ a homology theory defined by

$$E_n(X) = \lim_{k \to \infty} \pi_{n+k}(E_k \wedge X).$$

1.2.5. Examples.

(i) Eilenberg–MacLane spectrum.

Let $E_s = K(A, s)$ where each $K(A, s)$ is an Eilenberg–MacLane space where $A$ is an Abelian group and $\pi_n(K(A, s)) = \delta_{is}(A)$. By adjunction isomorphism, we have $K(A, n) \approx \Omega K(A, n + 1)$, and get the Eilenberg–MacLane spectrum whose associated cohomology theory is ordinary cohomology with coefficients in $A$, otherwise defined by means of singular chain complexes.

(ii) The suspension spectrum.

Let $X$ be a based space. The $n$-th space of the suspension spectrum $\Sigma^\infty X$ is $\Omega^\infty \Sigma^{n}(X)$ and the homotopy groups are $\pi_n(\Sigma^\infty X) = \lim_{k \to \infty} \pi_{n+k}(\Sigma^k X)$. When $X = S^0$, we obtain the sphere spectrum $\Sigma^\infty (S^0)$ and $\pi_n(\Sigma^\infty (S^0)) = \lim_{k \to \infty} \pi_{n+k}(S^k)$ is called the stable $n$-stem and denote by $\pi_{S}^n$.

Note that there is an adjoint pair $(\Sigma^\infty, \Omega^\infty)$ of functors between spaces and spectra and we can write $\Sigma^\infty X = \{X, \Sigma X, \Sigma^2 X, \ldots\}$. Also if $E$ is an $\Omega$-spectrum, $\Omega^\infty E$ is an infinite loop space. (Indeed, every infinite loop space is the initial space of an $\Omega$-spectrum and $\pi_n(E) = [\Sigma^\infty S^n, E] = \pi_n(\Omega^\infty E)$.)

2. Definitions of and relations between several higher algebraic $K$-theories (for rings)

In this section, we define the higher $K$-functors $K^Q_n$ (Quillen $K$-theory), $K^S_n$ ($K$-theory of Swan), $K^{k-v}_n$ (Karoubi–Villamayor $K$-theory), $K^M_n$ (Milnor $K$-theory) and $K^V_n$ (Volodin $K$-theory) for arbitrary rings with identity and discuss connections between the theories. Because $K^Q_n$ has been the most successful and has been most often used, we shall eventually write $K_n$ for $K^Q_n$.

2.1. $K^Q_n$ – the $K$-theory of Quillen

The definition of $K^Q_n(A)$, $A$ any ring with identity, will make use of the following result.

2.1.1. Theorem [94,111]. Let $X$ be a connected CW-complex, $N$ a perfect normal subgroup of $\pi_1(X)$. Then there exists a CW-complex $X^+$ (depending on $N$) and a map $i : X \to X^+$ such that

(i) $i_* : \pi_1(X) \to \pi_1(X^+)$ is the quotient map $\pi_1(X) \to \pi_1(X^+)/N$.

(ii) For any $\pi_1(X^+)/N$-module $L$, there is an isomorphism $i_* : H_*(X, i^*L) \to H_*(X^+, L)$ where $i^*L$ is $L$ considered as a $\pi_1(X)$-module.

(iii) The space $X^+$ is universal in the sense that if $Y$ is any CW-complex and $f : X \to Y$ is a map such that $f_* : \pi_1(X) \to \pi_1(Y)$ satisfies $f_*(N) = 0$, then there exists a unique map $f^+ : X^+ \to Y$ such that $f^+ i = f$. 

2.1.2. Definition. Let $A$ be a ring and take $X = \text{BGL}(A)$ in Theorem 2.1.1. Then $\pi_1 \text{BGL}(A) = \text{GL}(A)$ contains $E(A)$ as a perfect normal subgroup. Hence by Theorem 2.1.1, there exists a space $\text{BGL}(A)^+$. Define $K_n(A) = \pi_n(\text{BGL}(A)^+)$. 

2.1.3. Hurewitz map. For any ring $A$ with identity, there exist Hurewitz maps:

(i) $h_n : K_n(A) = \pi_n(\text{BGL}(A)^+) \to H_n(\text{BGL}(A)^+ , \mathbb{Z}) \approx H_n(\text{GL}(A) , \mathbb{Z}) \forall n \geq 1$,

(ii) $h_n : K_n(A) = \pi_n(\text{BE}(A)^+) \to H_n(\text{BE}(A)^+ , \mathbb{Z}) \approx H_n(\text{E}(A) , \mathbb{Z}) \forall n \geq 2$,

(iii) $h_n : K_n(A) = \pi_n(\text{BST}(A)^+) \to H_n(\text{BST}(A)^+ , \mathbb{Z}) \approx H_n(\text{St}(A) , \mathbb{Z}) \forall n \geq 3$.

Note that $\text{BGL}(A)^+$ is connected, $\text{BE}(A)^+$ is simply connected (i.e. one-connected) and $\text{BST}(A)^+$ is 2-connected.

For a comprehensive discussion of Hurewitz maps, see [6].

2.1.4. Examples/Remarks. For $n = 0, 1, 2$, $K_n(A)$ as defined in Section 2.1.2 can be identified respectively with the classical $K_n(A)$.

(i) $\pi_1(\text{BGL}(A)^+) = \text{GL}(A)/E(A) = K_1(A)$.

(ii) Note that $\text{BE}(A)^+$ is the universal covering space of $\text{BGL}(A)^+$ and so, we have

$$\pi_2(\text{BGL}(A)^+) \approx \pi_2(\text{BE}(A)^+) \approx H_2(\text{BE}(A)^+) \cong H_2(\text{BE}(A)) \cong H_2(\text{E}(A)) \approx K_2(A).$$

(iii) $K_3(A) = H_3(\text{St}(A))$. For a proof, see [38].

(iv) If $A$ is a finite ring, then $K_n(A)$ is finite (see [73] for a proof).

(v) For a finite field $\mathbb{F}_q$, $K_{2n}(\mathbb{F}_q) = 0$, $K_{2n-1}(\mathbb{F}_q) = \mathbb{Z}/(q^n - 1)$ (see [112]). In later Sections 3–8, we shall come across many computations of $K_n(A)$, for various rings, fields, etc.

2.2. $K^S_n$ – the K-theory of Swan and Gersten

2.2.1. In [138], R.G. Swan defined higher $K$-functors by resolving the functor $\text{GL}$ in the category of functors and S.M. Gersten in [36] defined higher $K$-functors by introducing a cotriple construction in the category $\text{Ring}$ of rings. Swan showed in [142] that Gersten’s resolution applied to $\text{GL}$ gives Swan’s groups. As has been the tradition, we denote this theory by $K^S_n(A)$.

2.2.2. Cotriples. A cotriple $(T, \varepsilon, \delta)$ in a category $\mathcal{A}$ is an endofunctor $T : \mathcal{A} \to \mathcal{A}$ together with natural transformations $\varepsilon : T \to \text{id}_{\mathcal{A}}$ and $\delta : T \to T^2$ such that the following diagrams commute for every object $A$.

\[
\begin{array}{ccc}
TA & \xrightarrow{T \delta} & T(TA) \\
\downarrow & & \downarrow \delta_{TA} \\
T(T(A)) & \xrightarrow{\delta} & TT(TA)
\end{array} \quad \begin{array}{ccc}
TA & \xleftarrow{T \varepsilon A} & T(TA) \\
\downarrow & & \downarrow \varepsilon_{TA} \\
TA & \xleftarrow{\varepsilon} & TA
\end{array}
\]
2.2.3. **Remarks.**

(i) If $\mathcal{A} \xleftarrow{L} \mathcal{B}$ is an adjoint situation where $L$ is left adjoint to $V$, then $T = LV : \mathcal{B} \to \mathcal{B}$ is part of a cotriple $(T, \varepsilon, \delta)$, where $\varepsilon : LV \to \text{id}_B$ is the counit of the adjunction.

(ii) Given a cotriple $T$ on $\mathcal{A}$, and $A \in \text{ob} \mathcal{A}$, we have a simplicial object $T^*A = \{T^nA\}$ of $\mathcal{A}$ with face maps $\delta_i = T^n A \to T^{n+1} A$ and degeneracy maps $s_i = T^n A \to T^{n-1} A$.

2.2.4. Let $\text{Ring}$ be the category of rings and for any ring $A$, let $FA$ be the free ring on the underlying set of $A$. Then $FA$ is a functor $\mathcal{S}et \to \text{Ring}$ adjoint to the forgetful functor and the adjointness yields a morphism $\varepsilon : FA \to A$ and a morphism $\delta : FA \to F^2A$ such that $(F, \varepsilon, \delta)$ is a cotriple in $\text{Ring}$.

Let $|r|$ be the free generator of $FA$ corresponding to $r \in A$. Then $\varepsilon(|r|) = r$ and $\delta(|r|) = \|r\|$. So, we obtain the augmented simplicial ring:

$$F^nA : R \leftarrow FA \rightarrow F^2A \rightarrow F^3A \cdots.$$

2.2.5. Define $K_{n}^S(A) = \hat{\pi}_n(\text{GL}(F^nA))$ where $\hat{\pi}_n = \pi_n$, $n \geq 1$, $\hat{\pi}_0(\text{GL}(F^nA)) = \ker(\pi_0(\text{GL}(F^nA)) \to \text{GL}(A))$ and $\hat{\pi}_{-1}(\text{GL}(F^nA)) = \text{Coker}(\text{GL}(F^nA) \to \text{GL}(A))$.

2.2.6. **Theorem** [138]. $K_n^S(FA) = 0$.

2.2.7. **Theorem** [38]. $K_n^Q(FA) = 0$.

2.2.8. **Theorem** [3]. For any ring $A$, there exists an exact sequence

$$\rightarrow K_{n+1}(A) \to K_{n+1}^S(A) \to K_n(FA) \to K_n(A) \to K_n^S(A) \to .$$

2.2.9. **Corollary** (Connection with Quillen $K$-theory). $K_n^S(A) = K_n^Q(A)$ for any ring $A$.

**Proof.** This follows from 2.2.6, 2.2.7 and 2.2.8.

2.3. $K_{n}^{k-v} - the K$-theory of Karoubi and Villamayor

2.3.1. Let $R(\Delta^a) = R[t_0, t_1, t_n]/(\Sigma t_i - 1) \simeq R[t_1, \ldots, t_n]$. Applying the functor GL to $R(\Delta^a)$ yields a simplicial group $\text{GL}(R(\Delta^a))$.

2.3.2. **Definition.** Let $R$ be a ring with identity. Define the Karoubi–Villamayor $K$-groups by $K_{n}^{k-v}(R) = \pi_{n-1}(\text{GL}(R(\Delta^a))) = \pi_n(\text{BGL}(R(\Delta^a)))$ for all $n \geq 1$. Note that $\pi_0(\text{GL}(R(\Delta^a)))$ is the quotient $\text{GL}(R)/\text{uni}(R)$ of $K_1(R)$ where $\text{uni}(R)$ is the subgroup of $\text{GL}(R)$ generated by unipotent matrices, i.e. matrices of the form $1 + N$ for some nilpotent matrix $N$. 


2.3.3. **Theorem** [151].
(i) For \( p \geq 1, \ q \geq 0 \), there is a spectral sequence \( E^1_{pq} = K_p(R[\Delta^q]) \Rightarrow K^{k-v}_{p+q}(R) \).
(ii) If \( R \) is regular, then the spectral sequence in (i) above degenerates and \( K_n(R) = K^{k-v}_n(R) \) for all \( n \geq 1 \).

2.3.4. **Definition.** A functor \( F : \text{Rings} \to \text{Z-mod} \) (Chain complexes etc.) is said to be homotopy invariant if for any ring \( R \), the natural map \( R \to R[t] \) induces an isomorphism \( F(R) \cong F(R[t]) \). Note that if \( F \) is homotopy invariant, then the simplicial object \( F(R[\Delta^n]) \) is constant.

2.3.5. **Theorem** [38]. The functors \( K^{k-v}_n : \text{Rings} \to \text{Z-mod} \) are homotopy invariant, i.e. \( K^{k-v}_n(R) \cong K^{k-v}_n(R[t]) \) for all \( n \geq 1 \).

2.4. \( K^V_n \) – Volodin \( K \)-theory

2.4.1. Let \( A \) be a ring with identity, \( \gamma \) a partial ordering of \( \{1, 2, \ldots, n\} \) and \( T^\gamma(A) := \{1 + (a_{ij}) \in \text{GL}_n(A) \mid a_{ij} = 0 \forall i < j \} \). Note that if \( \gamma \) is the standard ordering \( \{1 < \cdots < n\} \), then \( T^\gamma(A) \) is the subgroup of upper triangular matrices. The inclusion \( T^\gamma(A) \hookrightarrow \text{GL}(A) \) induces a cofibration on the classifying space \( BT^\gamma(A) \hookrightarrow B\text{GL}(A) \).

2.4.2. Define the Volodin space \( X(A) \) by \( X(A) := \bigcup_\gamma BT^\gamma(A) \).

2.4.3. **Theorem** [151]. For any ring \( A \) with identity, the connected space \( X(A) \) is acyclic (\( H_n(X(A)) = 0 \)) and is simple in dimension \( \geq 2 \).

2.4.4. **Definition.** Define \( K^V_n(A) := \pi_{n-1}(X(A)) \).

2.4.5. Connections with Quillen \( K \)-theory.

**Theorem** [94]. There exists a natural homotopy fibration \( X(A) \to B\text{GL}(A) \to B\text{GL}(A)^+ \) and hence \( \pi_1(X(A)) = \text{St}(A), \pi_n(X(A)) = K_{n+1}(A) \) for all \( n \geq 2 \), i.e.

\[
K^V_n(A) = \pi_n(X(A)) = K_{n+1}(A) \quad \forall n \geq 2.
\]

2.5. \( K^M_n \) – Milnor \( K \)-theory

2.5.1. Let \( A \) be a commutative ring with identity and \( T(A^*) \) the tensor algebra over \( \mathbb{Z} \) where \( A^* \) is the Abelian group of invertible elements of \( A \). For any \( x \in A^* - \{1\} \), the elements \( x \otimes (1 - x) \) and \( x \otimes (-x) \) generate a 2-sided ideal \( I \) of \( T(A^*) \). The quotient \( T(A^*)/I \) is a graded Abelian group whose component in degree 0, 1, 2 are respectively \( \mathbb{Z} \), \( A^* \) and \( K^M_2(A) \) where \( K^M_2(A) \) is the classical \( K_2 \)-group, see [105,79].
2.5.2. Connections with Quillen $K$-theory.

(i) As remarked above $K^M_n(A) = K^Q_n(A)$ for $n \leq 2$.

(ii) First observe that there is a well defined product $K^Q_m(A) \times K^Q_n(A) \to K^Q_{m+n}(A)$, due to J.L. Loday (see [95]). Now, there exists a map $\varphi: K^M_n(A) \to K^Q_n(A)$ constructed as follows: We use the isomorphism $K_1(A) \simeq A^\times$ to embed $A^\times$ in $K_1(A)$ and use the product in Quillen $K$-theory to define inductively a map $(A^\times)^n \to K_1(A)^n \to K_n(A)$, which factors through the exterior power $A^n A^\times$ over $\mathbb{Z}$, and hence through the Milnor $K$-groups $K^M_n(A)$ yielding the map $\varphi: K^M_n(A) \to K_n(A)$.

If $F$ is a field, we have the following more precise result due to A. Suslin.

2.5.3. THEOREM [132]. The kernel of $\varphi: K^M_n(F) \to K_n(F)$ is annihilated by $(n-1)!$.

We shall discuss more connections between Milnor and Quillen $K$-theories (especially for fields) in Section 5.

3. Higher $K$-theory of exact, symmetric monoidal and Waldhausen categories

3.1. Higher $K$-theory of exact categories – definitions and examples

In [79, Section 3], we discussed $K_0$ of exact categories $\mathcal{C}$, providing copious examples. In this section, we define $K_n(\mathcal{C})$ for all $n \geq 0$ with the observation that this definition generalizes to higher dimensions the earlier ones at the zero-dimensional level.

3.1.1. DEFINITION. Recall [108], [79, 3.1], that an exact category is a small additive category $\mathcal{C}$ (which is embeddable as a full subcategory of an Abelian category $\mathcal{A}$) together with a family $\mathcal{E}$ of short exact sequences $0 \to C' \xrightarrow{i} C \xrightarrow{j} C'' \to 0$ (I) such that $\mathcal{E}$ is the class of sequences (I) in $\mathcal{C}$ that are exact in $\mathcal{A}$ and $\mathcal{C}$ is closed under extensions (i.e. for any exact sequence $0 \to C' \xrightarrow{i} C \xrightarrow{j} C'' \to 0$ in $\mathcal{A}$ with $C', C''$ in $\mathcal{C}$, we have $C \in \mathcal{C}$).

In the exact sequence (I) above, we shall refer to $i$ as an inflation or admissible monomorphism, $j$ as a deflation or admissible epimorphism; and to the pair $(i, j)$ as a conflation.

Let $\mathcal{C}$ be an exact category. We form a new category $QC$ whose objects are the same as objects of $\mathcal{C}$ such that for any two objects $M, P \in \text{ob}(QC)$, a morphism from $M$ to $P$ is an isomorphism class of diagrams $M \xleftarrow{i} N \xrightarrow{j} P$ where $i$ is admissible monomorphism and $j$ is an admissible epimorphism in $\mathcal{C}$, i.e. $i$ and $j$ are part of some exact sequences $0 \to N' \xrightarrow{i} P' \xrightarrow{j} 0$ and $0 \to N' \xrightarrow{i} N \xrightarrow{j} M \to 0$ respectively.

Composition of arrows $M \xleftarrow{i} N \xrightarrow{j} P$ and $P \xleftarrow{i} R \xrightarrow{j} T$ is defined by the following diagram which yields an arrow

\[
\begin{array}{c}
M \xleftarrow{i} N \times_P R \xrightarrow{j} T \\
in QC
\end{array}
\]
3.1.2. **Definition.** For all $n \geq 0$, define

$$K_n(C) := \pi_{n+1}(BQC, o)$$

(see [114]).

3.1.3. We could also obtain $K_n(C)$ via spectra. For example, we could take the $\Omega$-spectrum (see 1.2) $BQC = \{\Omega BQC, BQC, BQ^2C, \ldots\}$ where $Q^iC$ is the multicategory defined in [154] and $\pi_n(BQC) = K_n(C)$.

3.1.4. **Examples.**

(i) For any ring $A$ with identity, the category $\mathcal{P}(A)$ of finitely generated projective modules over $A$ is exact and we shall write $K_n(A)$ for $K_n(\mathcal{P}(A))$.

Note that for all $n \geq 1$, $K_n(A)$ coincides with the groups $\pi_n(BGL(A)^\times)$ defined in 2.1.2.

(ii) Let $A$ be a left Noetherian ring. Then $\mathcal{M}(A)$, the category of finitely generated (left) $A$-modules is an exact category and we denote $K_n(\mathcal{M}(A))$ by $G_n(A)$. The inclusion functor $\mathcal{P}(A) \to \mathcal{M}(A)$ induces a homomorphism $K_n(A) \to G_n(A)$.

If $A$ is regular, then $K_n(A) \approx G_n(A)$ (see 4.1.2).

(iii) Let $X$ be a scheme, see [128], $\mathcal{P}(X)$ the category of locally free sheaves of $OX$-modules of finite rank (or equivalently category of finite-dimensional (algebraic) vector bundles on $X$). Then $\mathcal{P}(X)$ is an exact category and we write $K_n(X)$ for $K_n(\mathcal{P}(X))$, see [114].

If $X = \text{Spec}(A)$ for some commutative ring $A$, then we have an equivalence of categories:

$$\mathcal{P}(X) \to \mathcal{P}(A) : E \to \Gamma(X, E) = \{A\text{-modules of global sections}\}$$

with an inverse equivalence $\mathcal{P}(A) \to \mathcal{P}(X)$ given by

$$P \to \tilde{P} : U \to O_X(U) \otimes_A P.$$ 

So,

$$K_n(A) \approx K_n(X).$$
(iv) If $X$ is a Noetherian scheme, then the category $\mathcal{M}(X)$ of coherent sheaves of $O_X$-modules is exact. We write $G_n(X)$ for $K_n(\mathcal{M}(X))$. If $X = \text{Spec}(A)$, then we have an equivalence of categories $\mathcal{M}(X) \approx \mathcal{M}(A)$ and $G_n(X) \approx G_n(A)$.

(v) Let $R$ be a commutative ring with identity, $A$ an $R$-algebra that is finitely generated as an $R$-module, $\mathcal{P}_R(A)$ the category of left $A$-lattices. Then $\mathcal{P}_R(A)$ is an exact category and we write $G_n(R, A)$ for $K_n(\mathcal{P}_R(A))$. If $A = RG$, $G$ finite group, write $G_n(R, G)$ for $G_n(R, RG)$. If $R$ is regular, then $G_n(R, A) \approx G_n(A)$, see [67].

(vi) Let $G$ be a finite group, $S$ a $G$-set, $\mathcal{S}$ the translation category of $S$ (or category associated to $S$) see [72] or [25], or 1.1.7(ii). Then, the category $[\mathcal{S}, \mathcal{C}]$ of functors from $\mathcal{S}$ to an exact category $\mathcal{C}$ is also an exact category. We denote by $K^G_n(S, \mathcal{C})$ the Abelian group $K_n([\mathcal{S}, \mathcal{C}])$. As we shall see later $K^G_n(-, \mathcal{C}): G\text{Set} \to \text{Ab}$ is a ‘Mackey’ functor see [24] or [25] or [80].

If $S = G/G$, and $\mathcal{C}_G$ denotes the category of representations of $G$ in $\mathcal{C}$, then $[G/G, \mathcal{C}] \approx \mathcal{C}_G$. In particular, $[G/G, \mathcal{P}(R)] \approx \mathcal{P}(R)_G \approx \mathcal{P}_R(RG)$ and so $K^G_n(G/G, \mathcal{P}(R)) \approx K_n(\mathcal{P}(R)_G) \approx G_n(R, G) \approx G_n(RG)$ if $R$ is regular. As explained in [79,72], when $R = \mathbb{C}$, $K_0(\mathcal{P}(\mathbb{C})_G) \approx G_0(\mathbb{C}, G) \approx G_0(\mathbb{C}G) = \text{Abelian group of characters } X: G \to \mathbb{C}$.

We shall discuss relative generalizations of this in Section 8.

(vii) Let $X$ be a compact topological space, $F = \mathbb{R}$ or $\mathbb{C}$. Then the category $\text{VB}_F(X)$ of vector bundles on $X$ is an exact category and we can write $K_n(\text{VB}_F(X))$ as $K_n^F(X)$.

(viii) Let $X$ be an $H$-space, $m, n$ positive integers, $M^n_m$ an $n$-dimensional mod-$m$ Moore space, i.e. the space obtained from $S^{n-1}$ by attaching an $n$-cell via a map of degree $m$ (see [16] or [107]). Write $\pi_n(X, \mathbb{Z}/m)$ for $[M^n_m, X]$ the set of homotopy classes of maps form $M^n_m$ to $X$. If $X = \mathbb{B}Q\mathbb{C}$ where $\mathbb{C}$ is an exact category, write $K_n(\mathbb{C}, \mathbb{Z}/m)$ for $\pi_{n+1}(\mathbb{B}Q\mathbb{C}, \mathbb{Z}/m)$, $n \geq 1$, and call this group the mod-$m$ higher $K$-theory of $\mathbb{C}$. This theory is well defined for $\mathbb{C} = \mathcal{P}(A)$ where $A$ is any ring with identity and we write $K_n(A, \mathbb{Z}(m))$ for $K_n(\mathcal{P}(A), \mathbb{Z}/m)$. If $X$ is a scheme write $K_n(X, \mathbb{Z}/m)$ for $K_n(\mathcal{P}(X), \mathbb{Z}/m)$. For a Noetherian ring $A$, write $G_n(A, \mathbb{Z}/m)$ for $K_n(A, \mathbb{Z}/m)$ while for a Noetherian scheme $X$, we shall write $G_n(X, \mathbb{Z}/m)$ for $K_n(\mathcal{M}(X), \mathbb{Z}/m)$. For the applications, it is usual to consider $m = \ell^s$ where $\ell$ is a prime and $s$ a positive integer (see [16] or [78]).

(ix) Let $G$ be a discrete Abelian group, $M^n(G)$ the space with only one non-zero reduced integral cohomology group $\tilde{H}^n(M^n(G))$. Suppose that $\tilde{H}^n(M^n(G)) = G$. If we write $\pi_n(X, G)$ for $[M^n(G), X]$, and we put $G = \mathbb{Z}/m$, we recover (viii) above since $M^n_m = M^n(\mathbb{Z}/m)$. If $G = \mathbb{Z}$, $M^n(\mathbb{Z}) = S^n$ and so, $\pi_n(X, \mathbb{Z}) = [S^n, X] = \pi_n(X)$.

(x) With notations as in (ix), let $M^{n+1}_\ell = \lim_{\rightarrow} M^{n+1}_\ell$. For all $n \geq 0$, we shall denote $[M^{n+1}_\ell, BC]$ (C an exact category) by $K^{n}_\ell(C, \mathbb{Z}_\ell)$ and call this group the profinite (higher) $K$-theory of $C$. By way of notation, we shall write $K^{n}_\ell(A, \mathbb{Z}_\ell)$ if $P = \mathcal{M}(A)$. A any ring with identity: $G^n_\ell(A, \mathbb{Z}_\ell)$ if $C = \mathcal{M}(A)$, A Noetherian; $K^{pr}_\ell(X, \mathbb{Z}_\ell)$ if $C = \mathcal{P}(X)$, $X$ any scheme and $G^{pr}_\ell(X, \mathbb{Z}_\ell)$ if $C = \mathcal{M}(X)$, $X$ a Noetherian scheme. For a comprehensive study of these constructions and applications especially to orders and grouprings, see [78] or 7.3.
3.2. Higher $K$-theory of symmetric monoidal categories – definitions and examples

3.2.1. A symmetric monoidal category is a category $S$ equipped with a functor $\perp : S \times S \to S$ and a distinguished object ‘0’ such that ‘$\perp$’ is ‘coherently’ associative and commutative in the sense of MacLane (i.e. satisfying properties and diagrams in [79, 1.4.1]). Note that $BS$ is an $H$-space (see [39]).

3.2.2. Examples.
(i) Let $(\text{Iso} S)$ denotes the subcategory of isomorphisms in $S$, i.e. $\text{ob}(\text{Iso} S) = \text{ob} S$; morphisms are isomorphisms in $S$. $\pi_0(\text{Iso} S) =$ set of isomorphism classes of objects of $S$. Then $S^{\text{iso}} := \pi_0(\text{Iso} S)$ is a monoid.

$\text{Iso}(S)$ is equivalent to the disjoint union $\bigsqcup S \text{Aut}_S(S)$ and $B(\text{Iso} S)$ is homotopy equivalent to $\bigsqcup B(\text{Aut}_S(S))$, $S \in S^{\text{iso}}$.

(ii) If $S = \text{FSet}$ in (1), $\text{Aut}_{\text{FSet}}(S) \simeq \Sigma_n$ (symmetric group on $n$ letters) $\text{Iso}(\text{FSet})$ is equivalent to the disjoint union $\bigsqcup \Sigma_n$. $B(\text{Iso}(\text{FSet}))$ is homotopy equivalent to $\bigsqcup B\Sigma_n$.

(iii) $B(\text{Iso P}(R))$ is equivalent to the disjoint union $\bigsqcup B\text{Aut}(P)$, $P \in \text{P}(R)$.

(iv) Let $\mathcal{F}(R) =$ category of free $R$-modules $(\text{Iso} \mathcal{F}(R)) = [\Sigma GL_n(R) \text{ and } B(\text{Iso} \mathcal{F}(R))$ is equivalent to the disjoint union $\bigsqcup BGL_n(R)$. If $R$ satisfies the invariant basis property, then $\text{Iso}(\mathcal{F}(R))$ is a full subcategory of $\text{Iso}(\mathcal{P}(R))$ and $\text{Iso}(\mathcal{F}(R))$ is cofinal in $\text{Iso} \mathcal{P}(R)$.

3.2.3. Suppose that every map in $S$ is an isomorphism and every translation $S \perp : \text{Aut}_S(T) \to \text{Aut}_S(S \perp T)$ is an injection. We now define a category $S^{-1}S$ such that $K(S) = B(S^{-1}S)$ is a ‘group completion’ of $BS$.

Recall that a group completion of a homotopy commutative and homotopy associative $H$-space $X$ is an $H$-space $Y$ together with an $H$-space map $X \to Y$ such that $\pi_0(Y)$ is the group completion of (i.e. the Grothendieck group associated to) the monoid $\pi_0(X)$ (see [79, 1.1]) and the homology ring $H_n(Y, R)$ is isomorphic to the localization $\pi_0(X)^{-1}H_n(X, R)$ of $H_n(X, R)$.

3.2.4. Definition. Define $S^{-1}S$ as follows:

\[ \text{ob}(S^{-1}S) = \{ (S, T) \mid S, T \in \text{ob} S \}, \]

\[ \text{mor}_{S^{-1}S}(\langle S_1, T_1 \rangle, \langle S_1^1, T_1^1 \rangle) = \left\{ \frac{\text{equivalence class of composites}}{(S_1, T_1) \xrightarrow{S \perp} (S \perp S_1, S \perp T_1) \xrightarrow{(f,g)} (S_1^1, T_1^1)} \right\} \]

Notes.
(i) The composite $\langle S_1, T_1 \rangle \xrightarrow{S \perp} (S \perp S_1, S \perp T_1) \xrightarrow{(f,g)} (S_1^1, T_1^1)$ is said to be equivalent to $\langle S_1, T_1 \rangle \xrightarrow{T \perp} (T \perp S_1, T \perp T_1) \xrightarrow{(f',g')} (S_1', T_1')$ if there exists an isomorphism $\alpha : S \approx T$ in $S$ such that composition with $\alpha \perp S_1, \alpha \perp T_1$ sends $f'$ and $g'$ to $f$.

(ii) Since we have assumed that every translation is an injection in 3.2.3, it means that $S^{-1}S$ determines its objects up to unique isomorphism.
(iii) $S^{-1}S$ is a symmetric monoidal category with $(S, T) \perp (S', T') = (S \perp S', T \perp T')$ and the functor $S \to S^{-1}S : S \to (o, S)$ is monoidal. Hence $B(S^{-1}S)$ is an $H$-space (see [39]).

(iv) $BS \to B(S^{-1}S)$ is an $H$-space map and $\pi_0(S) \to \pi_0(S^{-1}S)$ is a map of Abelian monoids.

(v) $\pi_0(S^{-1}S)$ is an Abelian group.

3.2.5. **Examples.**

(i) If $S = \bigsqcup \text{GL}_n(R) = \text{Iso} \mathcal{F}(R)$, then $B(S^{-1}S)$ is a group completion of $BS$ and $B(S^{-1}S)$ is homotopy equivalent to $\mathbb{Z} \times B\text{GL}(R)^+$, see [39] or [167], for a proof. See theorem 3.2.8 below for a more general formulation of this.

(ii) For $S = \text{Iso}(\mathcal{F}\text{Set})$, $B(S^{-1}S)$ is homotopy equivalent to $\mathbb{Z} \times B\Sigma^+$ where $\Sigma$ is the infinite symmetric group (see [167]).

3.2.6. **Definition.** Let $S$ be a symmetric monoidal category in which every morphism is an isomorphism.

Define

$$K_n^+(S) := \pi_n(B(S^{-1}S)).$$

**Note.** $K_n^+(S)$ as defined above coincides with $K_n^+(S)$ as defined in [79, 1.4]. This is because $K_0^+(S) = \pi_0(B(S^{-1}S))$ is the group completion of the Abelian monoid $\pi_0(S) = S^{\text{iso}}$. For a proof, see [167].

3.2.7. **Remarks.** Suppose that $S$ is a symmetric monoidal category which has a countable sequence of objects $S_1, S_2, \ldots$ such that $S_{n+1} = S_n \perp T_n$ for some $T_n \in S$ and satisfying the cofinality condition, i.e. for every $S \in S$, there exist an $S'$ and an $n$ such that $S \perp S' \approx S_n$. If this situation obtains, then we can form $\text{Aut}(S) = \text{colim}_{n \to \infty} \text{Aut}_S(S_n)$.

3.2.8. **Theorem** [167]. Suppose that $S = \text{Iso}(S)$ is a symmetric monoidal category whose translations are injections, and that the conditions of 3.2.7 are satisfied so that the group $\text{Aut}(S)$ exists. Then the commutator subgroup $E$ of $\text{Aut}(S)$ is a perfect normal subgroup; $K_1(S) = \text{Aut}(S)/E$ and $B\text{Aut}(S)^+$ is the connected component of the identity in the group completion of $B(S^{-1}S)$.

Hence $B(S^{-1}S) \cong K_0(S) \times B\text{Aut}(S)^+$.

3.2.9. **Example.** Let $R$ be a commutative ring with identity. We saw in [79, 1.43] that $(S = \text{Pic}(R), \otimes)$ is a symmetric monoidal category. Since $\pi_0(S)$ is a group, $S$ and $S^{-1}S$ are homotopy equivalent (see [167]). Hence we get $K_0\text{Pic}(R) = \text{Pic}(R), \ K_1(\text{Pic}(R)) = U(R) \ (\text{units of } R)$, and $K_n(\text{Pic}(R)) = 0$ for all $n \geq 2$. 
3.3. Higher $K$-theory of Waldhausen categories – definitions and examples

3.3.1. Definition. A category with cofibrations is a category $\mathcal{C}$ with zero object together with a subcategory $\text{co} (\mathcal{C})$ whose morphisms are called cofibrations written $A \rightarrow B$ and satisfying the axioms

(C1) Every isomorphism in $\mathcal{C}$ is a cofibration.

(C2) If $A \rightarrow B$ is a cofibration and $A \rightarrow C$ any $\mathcal{C}$-map, then the pushout $B \cup_A C$ exists in $\mathcal{C}$

$A \xrightarrow{\sim} B$

$\downarrow$

$C \xrightarrow{\sim} B \cup_A C$

- Hence coproducts exists in $\mathcal{C}$ and each cofibration $A \rightarrow B$ has a cokernel $C = B/A$.
- Call $A \rightarrow B \rightarrow B/A$ a cofibration sequence.

(C3) The unique map $0 \rightarrow B$ is a cofibration $\forall \mathcal{C}$-objects $B$.

3.3.2. Definition. A Waldhausen category (or $W$-category for short) $\mathcal{C}$ is a category with cofibrations together with a subcategory $w(\mathcal{C})$ of weak equivalences (w.e. for short) containing all isomorphisms and satisfying.

Gluing axiom for weak equivalences (W1). For any commutative diagram

$C \leftarrow A \rightarrow B$

$\sim \downarrow \sim \downarrow \sim$

$C' \leftarrow A' \rightarrow B'$

in which the vertical maps are weak equivalences and the two right horizontal maps are cofibrations, the induced map $B \cup_A C \rightarrow B' \cup_{A'} C'$ is also a weak equivalence.

We shall sometimes denote $\mathcal{C}$ by $(\mathcal{C}, w)$.

3.3.3. Definition. A Waldhausen subcategory $\mathcal{A}$ of a $W$-category $\mathcal{C}$ is a subcategory which is also $W$-category such that (a) the inclusion $\mathcal{A} \subseteq \mathcal{C}$ is an exact functor, (b) the cofibrations in $\mathcal{A}$ are the maps in $\mathcal{A}$ which are cofibrations in $\mathcal{C}$ and whose cokernel lies in $\mathcal{A}$ and (c) the weak equivalences in $\mathcal{A}$ are the weak equivalences of $\mathcal{C}$ which lie in $\mathcal{A}$.

3.3.4. Definition. A $W$-category $\mathcal{C}$ is said to be saturated if whenever $(f, g)$ are composable maps and $fg$ is a w.e., then $f$ is a w.e., iff $g$ is.

- The cofibrations sequences in a $W$-category $\mathcal{C}$ form a category $\mathcal{E}$. Note that $\text{ob}(\mathcal{E})$ consists of cofibrations sequences $E : A \rightarrow B \rightarrow C$ in $\mathcal{C}$. A morphism $E \rightarrow E' : A' \rightarrow B' \rightarrow C'$ in $\mathcal{E}$ is a commutative diagram
To make $\mathcal{E}$ a $W$-category, we define a morphism $E \to E'$ in $\mathcal{E}$ to be a cofibration if $A \to A', C \to C'$ and $A' \cup_A B \to B'$ are cofibrations in $\mathcal{C}$ while $E \to E'$ is a w.e. if its component maps $A \to A', B \to B', C \to C'$ are w.e. in $\mathcal{C}$.

3.3.5. EXTENSION AXIOM. A $W$-category $\mathcal{C}$ is said to satisfy the extension axiom if for any morphism $f : E \to E'$ as in 3.3.4, maps $A \to A', C \to C'$ being w.e. implies that $B \to B'$ is also a w.e.

3.3.6. EXAMPLES.
(i) Any exact category $\mathcal{C}$ is a $W$-category where the cofibrations are the admissible monomorphisms and the w.e. are isomorphisms.
(ii) If $\mathcal{C}$ is any exact category, then the category $\text{Ch}_b(\mathcal{C})$ of bounded chain complexes in $\mathcal{C}$ is a $W$-category where the w.e. are quasi-isomorphisms (i.e. isomorphisms on homology) and a chain map $A_i \to B_i$ is a cofibration if each $A_i \to B_i$ is a cofibration (admissible monomorphism) in $\mathcal{C}$.
(iii) Let $\mathcal{C}$ = category of finite based CW-complexes. Then $\mathcal{C}$ is a $W$-category where the cofibrations are cellular inclusion and the w.e. are homotopy equivalences.
(iv) If $\mathcal{C}$ is a $W$-category, define $K_0(\mathcal{C})$ as the Abelian group generated by objects of $\mathcal{C}$ with relations
   (i) $A \to B \Rightarrow [A] = [B]$.
   (ii) $A \to B \to C \Rightarrow [B] = [A] + [C]$.
Note that this definition agrees with the earlier $K_0(\mathcal{C})$ given in [79, 3.1] for an exact category.

3.3.7. In order to define the $K$-theory space $K(\mathcal{C})$ such that

$$\pi_n(K(\mathcal{C})) = K_n(\mathcal{C})$$

for a $W$-category $\mathcal{C}$, we construct a simplicial $W$-category $S_n \mathcal{C}$, where $S_n \mathcal{C}$ is the category whose objects $A_*$ are sequences of $n$ cofibrations in $\mathcal{C}$, i.e.

$$A_* : 0 = A_0 \to A_1 \to A_2 \to \cdots \to A_n$$
together with a choice of every subquotient \( A_{ij} = A_j/A_i \) in such a way that we have a commutative diagram

\[
\begin{array}{c}
\xymatrix{
A_{n-1,n} & A_{n-2,n} & \cdots & A_{2,n} \\
& & & \\
& & & \\
& & & \\
A_1 & A_2 & A_3 & \cdots & A_n
} \end{array}
\]

By convention put \( A_{jj} = 0 \) and \( A_{0j} = A_j \).

A morphism \( A \rightarrow B \) is a natural transformation of sequences.

A weak equivalence in \( S_n(\mathcal{C}) \) is a map \( A \rightarrow B \) such that each \( A_i \rightarrow B_i \) (and hence each \( A_{ij} \rightarrow B_{ij} \)) is a w.e. in \( \mathcal{C} \). A map \( A \rightarrow B \) is a cofibration if for every \( 0 \leq i < j < k \leq n \) the map of cofibration sequences is a cofibration in \( E(\mathcal{C}) \).

For \( 0 < i \leq n \), define exact functors \( \delta_i : S_n(\mathcal{C}) \rightarrow S_{n+1}(\mathcal{C}) \) by omitting \( A_i \) from the notation and re-indexing the \( A_{jk} \) as needed. Define \( \delta_0 : S_n(\mathcal{C}) \rightarrow S_{n+1}(\mathcal{C}) \) where \( \delta_0 \) omits the bottom arrow. We also define \( s_i : S_n(\mathcal{C}) \rightarrow S_{n+1}(\mathcal{C}) \) by duplicating \( A_i \) and re-indexing (see [154]).

We now have a simplicial category \( n \rightarrow wS_n\mathcal{C} \) with degree-wise realization \( n \rightarrow B(wS_n\mathcal{C}) \), and denote the total space by \( |wS\mathcal{C}| \) (see [154]).

3.3.8. Definition. The \( K \)-theory space of a \( W \)-category \( \mathcal{C} \) is \( K(\mathcal{C}) = \Omega |wS\mathcal{C}| \). For each \( n \geq 0 \), the \( K \)-groups are defined as \( K_n(\mathcal{C}) = \pi_n(K\mathcal{C}) \).

3.3.9. By iterating the \( S_\bullet \) construction, one can show (see [154]) that the sequence

\[ \{ \Omega |wS_\bullet \mathcal{C}|, \Omega |wS_\bullet S_\bullet \mathcal{C}|, \ldots, \Omega |wS_\bullet \mathcal{C}| \ldots \} \]
forms a connective spectrum $\mathbb{K}(\mathcal{C})$ called the $K$-theory spectrum of $\mathcal{C}$. Hence $K(\mathcal{C})$ is an infinite loop space, see 1.2.2.

3.3.10. Examples.

(i) Let $\mathcal{C}$ be an exact category, $\text{Ch}_b(\mathcal{C})$ the category of bounded chain complexes over $\mathcal{C}$. It is a theorem of Gillet and Waldhausen that $K(\mathcal{C}) \cong K(\text{Ch}_b(\mathcal{C}))$ and so, $K_n(\mathcal{C}) \simeq K_n(\text{Ch}_b(\mathcal{C}))$ for every $n \geq 0$ (see [32]).

(ii) Perfect complexes. Let $R$ be any ring with identity and $\mathcal{M}'(R)$ the exact category of finitely presented $R$-modules. (Note that $\mathcal{M}'(R) = \mathcal{M}(R)$ if $R$ is Noetherian.) An object $M_\bullet \in \text{Ch}_b(\mathcal{M}'(R))$ is called a perfect complex if $M_\bullet$ is quasi isomorphic to an object in $\text{Ch}_b(\mathcal{P}(R))$. The perfect complexes form a Waldhausen subcategory $\text{Perf}(R)$ of $\text{Ch}_b(\mathcal{M}'(R))$. So, we have

$$K(R) \simeq K(\text{Ch}_b(\mathcal{P}(R))) \cong K(\text{Perf}(R)).$$

(iii) Derived categories. Let $\mathcal{C}$ be an exact category and $H^b(\mathcal{C})$ the (bounded) homotopy category of $\mathcal{C}$, i.e. the stable category of $\text{Ch}_b(\mathcal{C})$ (see [63]). So, $\text{ob}(H^b(\mathcal{C})) = \text{Ch}_b(\mathcal{C})$ and morphisms are homotopy classes of bounded complexes. Let $A(\mathcal{C})$ be the full subcategory of $H^b(\mathcal{C})$ consisting of acyclic complexes (see [63]). The derived category $D^b(\mathcal{C})$ of $\mathcal{C}$ is defined by $D^b(\mathcal{C}) = H^b(\mathcal{C})/A(\mathcal{C})$. A morphism of complexes in $\text{Ch}_b(\mathcal{C})$ is called a quasi-isomorphism if its image in $D^b(\mathcal{C})$ is an isomorphism. We could also define the unbounded derived category $D(\mathcal{C})$ from unbounded complexes $\text{Ch}(\mathcal{C})$.

Note that there exists a faithful embedding of $\mathcal{C}$ in an Abelian category $\mathcal{A}$ such that $\mathcal{C} \subseteq \mathcal{A}$ is closed under extensions and the exact functor $\mathcal{C} \to \mathcal{A}$ reflects exact sequences. So, a complex in $\text{Ch}(\mathcal{C})$ is acyclic iff its image in $\text{Ch}(\mathcal{A})$ is acyclic. In particular, a morphism in $\text{Ch}(\mathcal{C})$ is a quasi-isomorphism iff its image in $\text{Ch}(\mathcal{A})$ is a quasi-isomorphism. Hence, the derived category $D(\mathcal{C})$ is the category obtained from $\text{Ch}(\mathcal{C})$ formally inverting quasi-isomorphisms.

(iv) Stable derived categories and Waldhausen categories. Now let $\mathcal{C} = \mathcal{M}'(R)$. A complex $M_\bullet$ in $\mathcal{M}'(R)$ is said to be compact if the functor $\text{Hom}(M_\bullet, -)$ commutes with arbitrary set-valued coproducts. Let $\text{Comp}(R)$ denote the full subcategory of $D(\mathcal{M}'(R))$ consisting of compact objects. Then we have $\text{Comp}(R) \subseteq D^b(\mathcal{M}'(R)) \subseteq D(\mathcal{M}'(R))$.

Define the stable derived category of bounded complexes $D^b(\mathcal{M}'(R))$ as the quotient category of $D^b(\mathcal{M}'(R))$ with respect to $\text{Comp}(R)$. A morphism of complexes in $\text{Ch}_b(\mathcal{M}'(R))$ is called a stable quasi-isomorphism if its image in $D^b(\mathcal{M}'(R))$ is an isomorphism. The family of stable quasi-isomorphisms in $\mathcal{A} = \text{Ch}_b(\mathcal{M}'(R))$ is denoted $\omega \mathcal{A}$.

(v) Theorem.

1. $w(\text{Ch}_b(\mathcal{M}'(R)))$ forms a set of weak equivalences and satisfies the saturation and extension axioms.

2. $\text{Ch}_b(\mathcal{M}'(R))$ together with the family of stable quasi-isomorphisms is a Waldhausen category.
4. Some fundamental results and exact sequences in higher $K$-theory

4.1. Resolution theorem

4.1.1. Resolution theorem for exact categories [114]. Let $\mathcal{P} \subset \mathcal{H}$ be full exact subcategories of an Abelian category $\mathcal{A}$, both closed under extensions and inheriting their exact structure from $\mathcal{A}$. Suppose that (1) every object $M$ of $\mathcal{H}$ has a finite $\mathcal{P}$-resolution and (2) $\mathcal{P}$ is closed under kernels in $\mathcal{H}$, i.e. if $L \rightarrow M \rightarrow N$ is an exact sequence in $\mathcal{H}$ with $M, N \in \mathcal{P}$, then $L$ is also in $\mathcal{P}$. Then $K_n\mathcal{P} \cong K_n\mathcal{H}$ for all $n \geq 0$.

4.1.2. Remarks and examples.

(i) Let $R$ be a regular Noetherian ring. Then by taking $\mathcal{H} = \mathcal{M}(R)$, $\mathcal{P} = \mathcal{P}(R)$ in 4.1.1, we have $K_n(R) \cong G_n(R)$ for all $n \geq 0$.

(ii) Let $R$ be any ring with identity and $\mathcal{H}(R)$ the category of all $R$-modules having finite homological dimension (i.e. having a finite resolution by finitely generated projective $R$-modules). $H_n(R)$ the subcategory of modules in $\mathcal{H}(R)$ having resolutions of length $\leq s$. Then by 4.1.1 applied to $\mathcal{P}(R) \subseteq H_s(R) \subseteq H(R)$ we have $K_n(R) \cong K_n(H(R)) \cong K_n(H_s(R))$ for all $s \geq 1$.

(iii) Let $T = \{T_i\}$ be an exact connected sequence of functors from an exact category $\mathcal{C}$ to an Abelian category, i.e. given an exact sequence $0 \rightarrow M' \rightarrow M \rightarrow M'' \rightarrow 0$ in $\mathcal{C}$ there exists a long exact sequence $\cdots \rightarrow T_2M'' \rightarrow T_1M' \rightarrow T_1M \rightarrow$. Let $\mathcal{P}$ be the full subcategory of $T$-acyclic objects (i.e. objects $M$ such that $T_n(M) = 0$ for all $n \geq 1$), and assume that for each $M \in \mathcal{C}$, there is a map $P \rightarrow M$ such that $P \in \mathcal{P}$ and that $T_nM = 0$ for $n$ sufficiently large. Then $K_nP \cong K_n\mathcal{C}$ $\forall n \geq 0$ (see [114]).

(iv) As an example of (iii) let $A, B$ be a Noetherian rings, $f : A \rightarrow B$ a homomorphism, $B$ a flat $A$-module, then we have a homomorphism of $K$-groups: $G_n(A) \rightarrow G_n(B)$ (since $B \otimes_A ?$ is exact). Let $B$ be of finite tor-dimension as a right $A$-module. Then by applying (iii) above, to $\mathcal{C} = \mathcal{M}(A)$, $T_i(M) = \text{Tor}^A_i(B, M)$ and taking $\mathcal{P}$ as the full subcategory of $\mathcal{M}(A)$ consisting of $M$ such the $T_iM = 0$ for $i > 0$, we have $K_n\mathcal{P} \cong G_n(A)$.

(v) Let $\mathcal{C}$ be an exact category and $\text{Nil}(\mathcal{C})$ the category whose objects are pairs $(M, v)$ with $M \in \mathcal{C}$ and $v$ is a nilpotent endomorphism of $M$. Let $\mathcal{C}_0 \subset \mathcal{C}$ be an exact subcategory of $\mathcal{C}$ such that every object of $\mathcal{C}$ has a finite $\mathcal{C}_0$-resolution. Then every object of $\text{Nil}(\mathcal{C})$ has a finite $\text{Nil}(\mathcal{C}_0)$ resolution and so, by 4.1.1,

$$K_n(\text{Nil}(\mathcal{C}_0)) \cong K_n(\text{Nil}(\mathcal{C})).$$

4.2. Additivity theorem (for exact and Waldhausen categories)

4.2.1. Let $\mathcal{A}, \mathcal{B}$ be exact categories. A sequence of functors $F' \rightarrow F \rightarrow F''$ from $\mathcal{A}$ to $\mathcal{B}$ is called an exact sequence of exact functors if $0 \rightarrow F'(A) \rightarrow F(A) \rightarrow F''(A) \rightarrow 0$ is an exact sequence in $\mathcal{B}$ for every $A \in \mathcal{A}$.

Let $\mathcal{A}, \mathcal{B}$ be Waldhausen categories. If $F'(A) \rightarrow F(A) \rightarrow F''(A)$ is a cofibration sequence in $\mathcal{B}$ and for every cofibration $A \rightarrow A'$ in $\mathcal{A}$, $F(A) \cup_{F(F)} F'(A') \rightarrow F(A')$ is a
cofibration in \( \mathcal{B} \) say that \( F' \hookrightarrow F \twoheadrightarrow F'' \) a short exact sequence or a cofibration sequence of exact functors.

**4.2.2. Additivity Theorem.** Let \( F' \hookrightarrow F \twoheadrightarrow F'' \) be a short exact sequence of exact functors from \( \mathcal{A} \) to \( \mathcal{B} \) where both \( \mathcal{A} \) and \( \mathcal{B} \) are either exact categories or Waldhausen categories. Then \( F_* \cong F'_* + F''_* : K_*(\mathcal{A}) \to K_*(\mathcal{B}) \).

**4.2.3. Remarks and Examples.**

(i) It follows from 4.2.2 that if \( 0 \to F_1 \to F_2 \to \cdots \to F_s \to 0 \) is an exact sequence of exact functors \( \mathcal{A} \to \mathcal{B} \) then

\[
\sum_{k=0}^{s} (-1)^k F_k = 0 : K_n(\mathcal{A}) \to K_n(\mathcal{B})
\]

for all \( n \geq 0 \) (see [167]).

(ii) Let \( X \) be a scheme, \( E \in \mathcal{P}(X) \) (see 3.1.4(iii)). Then we have an exact functor \((E \otimes ?) : \mathcal{P}(X) \to \mathcal{P}(X)\) which induces homomorphisms \( K_n(\mathcal{X}) \to K_n(\mathcal{Y}) \).

If \( 0 \to E' \to E \to E'' \to 0 \) is an exact sequence in \( \mathcal{P}(X) \), then by 4.2.2

\[
(E \otimes ?)_s = (E' \otimes ?)_s + (E'' \otimes ?)_s : K_n(\mathcal{X}) \to K_n(\mathcal{Y})
\]

Hence we obtain a homomorphism

\[
K_0(\mathcal{X}) \otimes K_n(\mathcal{Y}) \to K_n(\mathcal{Y}) : (E) \otimes y \to (E \otimes ?)_s y, \quad y \in K_n(\mathcal{Y}),
\]

making each \( K_n(\mathcal{Y}) \) a \( K_0(\mathcal{X}) \)-module.

(iii) *Flasque categories.* An exact (or Waldhausen) category is called flasque if there is an exact functor \( \infty : \mathcal{A} \to \mathcal{A} \) and a natural isomorphism \( \infty(A) \cong A \amalg \infty(A) \); i.e. \( \infty \cong 1 \amalg \infty \) where \( 1 \) is the identity functor. By 4.2.2, \( \infty_s = 1_s \amalg \infty_s \) and hence the identity map \( 1_s : K(\mathcal{A}) \to K(\mathcal{A}) \) is null homotopic. Hence \( K(\mathcal{A}) \) is contractible and so \( \pi_n(K(\mathcal{A})) = K_n(\mathcal{A}) = 0 \) for all \( n \).

**4.3. Devissage**

**4.3.1. Devissage Theorem [114].** Let \( \mathcal{A} \) be an Abelian category, \( \mathcal{B} \) a non-empty full subcategory closed under subobjects, quotient objects and finite products in \( \mathcal{A} \). Suppose that every object \( M \) of \( \mathcal{A} \) has a finite filtration \( 0 = M_0 \subset M_1 \subset \cdots \subset M_n = M \) such that \( M_i/M_{i-1} \in \mathcal{B} \) for each \( i \), then the inclusion \( Q \mathcal{B} \to Q \mathcal{A} \) is a homotopy equivalence. Hence \( K_1(\mathcal{B}) \cong K_1(\mathcal{A}) \).

**4.3.2. Corollary [114].** Let \( \mathfrak{a} \) be a nilpotent two-sided ideal of a Noetherian ring \( R \). Then for all \( n \geq 0 \), \( G_n(\mathfrak{a}/\mathfrak{a}) \cong G_n(R) \).
4.3.3. Examples.

(i) Let $R$ be an Artinian ring with maximal ideal $m$ such that $m^r = 0$ for some $r$. Let $k = R/m$ (e.g., $R \equiv \mathbb{Z}/p^r, k \equiv \mathbb{F}_p$). In 4.3.1 put $B = \text{category of finite-dimensional } k$-vector spaces and $A = \mathcal{M}(R)$. Then we have a filtration $0 = m^r M \subset m^{r-1} M \subset \cdots \subset mM \subset M$ for any $M \in \mathcal{M}(R)$. Hence by 4.3.1, $G_n(R) \approx K_n(k)$.

(ii) Let $X$ be a Noetherian scheme, $i: Z \subset X$ the inclusion of a closed subscheme. Then $\mathcal{M}(Z)$ is an Abelian subcategory of $\mathcal{M}(X)$ via the direct image $i: \mathcal{M}(Z) \subset \mathcal{M}(X)$. Let $\mathcal{M}_Z(X)$ be the Abelian category of $\mathcal{O}_X$-modules supported on $Z$, an ideal sheaf in $\mathcal{O}_X$ such that $\mathcal{O}_X/a = \mathcal{O}_Z$. Then every $M \in \mathcal{M}_Z(X)$ has a finite filtration $M \supset Ma \supset Ma^2 \supset \cdots$ and so, by devissage, $K_n(\mathcal{M}_Z(X)) \approx K_n(\mathcal{M}(Z)) \approx G_n(Z)$.

4.4. Localization

4.4.1. A full subcategory $\mathcal{B}$ of an Abelian category $\mathcal{A}$ called a Serre subcategory if whenever:

$$0 \to M' \to M \to M'' \to 0$$

is an exact sequence in $\mathcal{A}$, then $M \in \mathcal{B}$ if and only if $M', M'' \in \mathcal{B}$. Given such a $\mathcal{B}$, construct a quotient Abelian category $\mathcal{A}/\mathcal{B}$ as follows:

$$\text{ob}(\mathcal{A}/\mathcal{B}) = \text{ob} \mathcal{A}.$$  

$\mathcal{A}/\mathcal{B}(M, N)$ is defined as follows: If $M' \subseteq M, N' \subseteq N$ are subobjects such that $M/M' \in \text{ob} \mathcal{B}, N' \in \text{ob} \mathcal{B}$, then there exists a natural isomorphism $\mathcal{B}(M, N) \to \mathcal{B}(M', N/N')$.

As $M', N'$ range over such pairs of objects, the group $\mathcal{B}(M', N/N')$ forms a direct system of Abelian groups and we define

$$\mathcal{A}/\mathcal{B}(M, N) = \lim_{\mathcal{B}(M', N/N')}.$$ 

NOTE. Let $T: \mathcal{A} \to \mathcal{A}/\mathcal{B}$ be the quotient functor: $M \mapsto T(M) = M$.

(i) $T: \mathcal{A} \to \mathcal{A}/\mathcal{B}$ is additive functor.

(ii) If $\mu \in \mathcal{A}(M, N)$ then $T(\mu)$ is null if and only if $\text{Ker} \mu \in \text{ob} \mathcal{B}$ and $T(\mu)$ is an epimorphism if and only if $\text{Coker} \mu \in \text{ob} \mathcal{B}$.

(iii) $\mathcal{A}/\mathcal{B}$ is an additive category such that $T: \mathcal{A} \to \mathcal{A}/\mathcal{B}$ is an additive functor.

4.4.2. Localization Theorem [114]. If $\mathcal{B}$ is a Serre subcategory of an Abelian category $\mathcal{A}$, then there exists a long exact sequence:

$$\cdots \to K_n(\mathcal{B}) \to K_n(\mathcal{A}) \to K_n(\mathcal{A}/\mathcal{B}) \to K_{n-1}(\mathcal{B}) \to \cdots$$

$$\cdots \to K_0(\mathcal{B}) \to K_0(\mathcal{A}) \to K_0(\mathcal{A}/\mathcal{B}) \to 0.$$ (I)
4.4.3. **Examples.**

(i) Let $A$ be a Noetherian ring, $S \subset A$ a central multiplicative system; $\mathcal{A} = \mathcal{M}(A)$, $\mathcal{B} = \mathcal{M}_S(A)$, the category of finitely generated $S$-torsion $A$-modules, $\mathcal{A}/\mathcal{B} \simeq \mathcal{M}(A_S)$ = category of finitely generated $A_S$-modules.

Let $T$ be the quotient functor $\mathcal{M}(A) \to \mathcal{M}(A)/\mathcal{M}_S(A)$, $u: \mathcal{M}(A)/\mathcal{M}_S(A) \to \mathcal{M}(A_S)$ is an equivalence of categories such the $uT \simeq L$, where $L: \mathcal{M}(A) \to \mathcal{M}(A_S)$. We thus have an exact sequence $K_{n+1}(\mathcal{M}(A_S)) \to K_n(\mathcal{M}_S(A)) \to K_n(\mathcal{M}(A)) \to K_{n-1}(\mathcal{M}_S(A))$, that is:

$$\cdots \to K_n(\mathcal{M}_S(A)) \to G_n(A) \to G_n(A_S) \to K_{n-1}(\mathcal{M}_S(A)) \to \cdots.$$

(ii) Let $A = R$ in (i) be a Dedekind domain with quotient field $F$, $S = R \setminus \{0\}$. Then, one can show that

$$\mathcal{M}_S(R) = \bigcup_m \mathcal{M}(R/m^k)$$

as $m$ runs through all maximal ideals of $R$.

So,

$$K_n(\mathcal{M}_S(R)) \simeq \bigoplus_m \lim_{k \to \infty} G_n(R/m^k)$$

$$= \bigoplus_m G_n(R/m) = \bigoplus_m K_n(R/m).$$

So, (I) gives

$$\to K_{n+1}(F) \to \bigoplus_m K_n(R/m) \to K_n(R) \to K_n(F) \to \bigoplus_m K_{n-1}(R/m)$$

$$\to \cdots \bigoplus_m K_2(R/m) \to K_2(R) \to K_2(F) \to \bigoplus_m K_1(R/m)$$

$$\to K_1(R) \to K_1(F) \to \bigoplus_m K_0(R/m) \to K_0(R) \to K_0(F),$$

that is

$$\cdots \to \cdots \to \bigoplus_m K_2(R/m) \to K_2(R) \to K_2(F) \to \bigoplus_m (R/m)^*$$

$$\to R^* \to F^* \to \bigoplus (\mathbb{Z}) \to \mathbb{Z} \oplus \text{Cl}(R) \to \mathbb{Z} \to 0.$$

(iii) Let $R$ in (i) be a discrete valuation ring (e.g., the ring of integers in a $p$-adic field) with unique maximal ideal $m = s R$. Let $F = \text{quotient field of } R$. Then $F = R[1/s]$, with residue field $= R/m = k$. Hence, we obtain the following exact sequence

$$\to K_n(k) \to K_n(R) \to K_n(F) \to K_{n-1}(k) \to \cdots \to K_2(k) \to K_2(R)$$

$$\to K_2(F) \to K_1(k) \to \cdots \to K_0(F) \to 0.$$

(II)
Gersten’s conjecture says that the sequence (II) breaks up into split short exact sequences

\[ 0 \to K_n(R) \xrightarrow{\alpha_n} K_n(F) \xrightarrow{\beta_n} K_{n-1}(k) \to 0. \]

For this to happen, one must have that for all \( n \geq 1 \), \( K_n(k) \to K_n(R) \) is the zero map and that there exists a map \( K_{n-1}(k) \to K_n(F) \) such that \( K_n(F) \simeq K_n(R) \oplus K_{n-1}(k) \), i.e. \( \beta_n \eta_n = 1_{K_{n-1}(k)} \).

True for \( n = 0 \), \( K_0(R) \simeq K_0(F) \simeq \mathbb{Z} \).

True for \( n = 1 \), \( K_1 F \simeq F^*, K_1(R) = R^*, F^* = R^* \times \{ s^n \} \).

True for \( n = 2 \).

\[ 0 \to K_2(R) \xrightarrow{\alpha_2} K_2(F) \xrightarrow{\beta_2} K_1(k) \to 0. \]

Here \( \beta_2 \) is the tame symbol. If characteristic of \( F = \) characteristic of \( k \), then Gersten’s conjecture is also known to be true. When \( k \) is algebraic over \( F_p \), then Gersten’s conjecture is also true. It is not known in the case when \( \text{char}(F) = 0 \) or \( \text{char}(k) = p \).

(iv) Let \( R \) be a Noetherian ring, \( S = \{ s^n \} \) a central multiplicative system \( B = \mathcal{M}_S(R) \), \( A = \mathcal{M}(R) \).

\[ \mathcal{A}/\mathcal{B} = \mathcal{M}(R_S) = \bigcup_{n=1}^{\infty} \mathcal{M}(R/s^n R). \]

Then (I) gives

\[ \cdots \to G_{n+1}(R_S) \to K_n(\mathcal{M}_S(R)) \to G_n(R) \to G_n(R_S) \to K_{n-1}(\mathcal{M}_S(R)). \]

Note that \( K_n(\mathcal{M}_S(R)) = K_n(\bigcup_{n=1}^{\infty} \mathcal{M}(R/s^n R)) \).

Now, by devissage \( G_n(R/s^n R) \simeq G_n(R/s R) \).

Hence \( K_n(\bigcup_{n=1}^{\infty} \mathcal{M}(R/s^n R)) = \lim_{n \to \infty} G_n(R/s^n R) = G_n(R/s R) \). So, we have

\[ \cdots G_{n+1} \left( R \left( \frac{1}{s} \right) \right) \to G_n(R/s R) \to G_n(R) \to G_n \left( R \left( \frac{1}{s} \right) \right) \to G_{n-1}(R/s R) \to \cdots. \]

(v) Let \( R \) be the ring of integers in a \( p \)-adic field \( F \), \( \Gamma \) a maximal \( R \)-order in a semi-simple \( F \)-algebra \( \Sigma \), if \( S = R \setminus \{ 0 \} \), then \( F = R_S \)

\[ B = \mathcal{M}_S(\Gamma), \quad A = \mathcal{M}(\Gamma), \quad A/B = \mathcal{M}(\Sigma). \]
Higher algebraic $K$-theory

Then sequence (I) yields an exact sequence

$$\cdots \to K_n(\Gamma) \to K_n(\Sigma) \to K_{n-1}(\mathcal{M}_S(\Gamma)) \to K_{n-1}(\Gamma)$$

$$\to K_{n-1}(\Sigma).$$

(\text{I})

One can see from (iv) that if $m \pi R$ is the unique maximal ideal of $R$, then $K_n(\mathcal{M}_S(\Gamma)) = \lim_{n \to \infty} G_n(\Gamma/\pi^n \Gamma) \simeq K_n(\Gamma/\text{rad} \Gamma)$ (see [25]). Here $\Sigma = \Gamma_S$ where $S = \{\pi^i\}$. We have also used above the corollary to devissage which says that if $a$ is a nilpotent ideal in a Noetherian ring $R$, then $G_n(R) \simeq G_n(R/a)$ (see 4.3.2).

(vi) Let $R$ be the ring of integers in an algebraic number field $F$, $A$ any $R$-order in a semi-simple $F$-algebra $\Sigma$. Let $S = R = 0$. Then we have the following exact sequence

$$\cdots \to K_n(\mathcal{M}_S(A)) \to G_n(A) \to G_n(\Sigma) \to K_{n-1}(\mathcal{M}_S(A)) \to \cdots.$$

One can show that $K_n(\mathcal{M}_S(A)) \simeq \bigoplus G_n(A/pA)$ where $p$ runs through all the prime ideals of $R$. See [69] for further details about how to use this sequence to obtain finite generation of $G_n(A)$, and the fact that $SG_n(A)$ is finite (see [69,71]).

(vii) Let $X$ be a Noetherian scheme, $U$ an open subscheme of $X$, $Z = X \setminus U$, the closed complement of $U$ in $X$. Put $\mathcal{A} = \mathcal{M}(X) = \text{category of coherent } 0_X\text{-modules}$, and let $\mathcal{B}$ be the category of coherent $O_X$-modules whose restriction to $U$ is zero (i.e. the category of coherent modules supported by $Z$). Let $\mathcal{A}/\mathcal{B}$ be the category of coherent $O_U$-modules. Then we have the following exact sequence

$$\cdots G_n(Z) \to G_n(X) \to G_n(U) \to G_{n-1}(Z) \to \cdots \to G_0(Z)$$

$$\to G_0(X) \to G_0(U) \to 0.$$

So far, our localization results have involved mainly the $G_n$-theory which translates into $K_n$-theory when the rings involved are regular. We now obtain localization for $K_n$-theory.

4.4.4. \textsc{Theorem.} Let $S$ be a central multiplicative system for a ring $R$, $\mathcal{H}_S(R)$ the category of $S$-torsion finitely generated $R$-modules of finite projective dimension. If $S$ consists of non-zero divisors, then there exists an exact sequence

$$\cdots \to K_{n+1}(R_S) \to K_n(\mathcal{H}_S(R)) \xrightarrow{\eta} K_n(R) \xrightarrow{\alpha} K_n(R_S) \to \cdots.$$

For a proof see [39].

4.4.5. \textsc{Remarks.} It is still an open problem to understand $K_n(\mathcal{H}_S(R))$ for various rings $R$. 

If $R$ is regular (e.g., $R = \mathbb{Z}$, the integers in a number field, a Dedekind domain, a maximal order), then $\mathcal{M}_S(R) = \mathcal{H}_S(R)$ and

$$K_n(\mathcal{H}_S(R)) = K_n(\mathcal{M}_S(R)); \quad G_n(R) = K_n(R).$$

So we recover $G$-theory. If $R$ is not regular, then $K_n(\mathcal{H}_S(R))$ is not known in general.

### 4.4.6. Definition

Let $\alpha : A \to B$ be a homomorphism of rings $A, B$. Suppose that $s$ is a central non-zero divisor in $B$. Call $\alpha$ an analytic isomorphism along $s$ if $A/sA \cong B/\alpha(s)B$.

### 4.4.7. Theorem

If $\alpha : A \to B$ is an analytic isomorphism along $s \in S = \{s^i\}$ where $s$ is a central non-zero divisor, then $H_S(A) = H_S(B)$.

**Proof** follows by comparing the localization sequences for $A \to A[\frac{1}{s}]$ and $B \to B[\frac{1}{s}]$.

### 4.5. Fundamental theorem for higher $K$-theory

#### 4.5.1. Let $C$ be an exact category, $\text{Nil}(C)$ the category of nilpotent endomorphisms in $C$, i.e. $\text{Nil}(C) = \{(M, \nu) \mid M \in C, \nu$ a nilpotent endomorphism of $M\}$. Then we have two functors $Z : C \to \text{Nil}(C) = Z(M) = (M, 0)$ (where ‘0’ = zero endomorphism) and $F : \text{Nil}(C) \to C : F(M, \nu) = M$ satisfying $FZ = 1_C$. Hence we have a split exact sequence $0 \to K_n(C) \xrightarrow{Z} K_n(\text{Nil}(C)) \to \text{Nil}(C) \to 0$ which defines $\text{Nil}_n(C)$ as the cokernel of $Z$.

Hence $K_n(\text{Nil}(C)) \simeq K_n(C) \oplus \text{Nil}_n(C)$.

#### 4.5.2. Let $R$ be a ring with identity, $H(R)$ the category of $R$-modules of finite homological dimension, $H_S(R)$ the category of $S$-torsion objects of $H(R)$, $\mathcal{M}_S(R)$ the category of finitely generated $S$-torsion $R$-modules. One can show (see [167]) that if $S = T_+ = \{t^i\}$, the free Abelian monoid on one generator $t$, then there exist isomorphisms $\mathcal{M}_{T_+}(R[t]) \simeq \text{Nil}(\mathcal{M}(R)), H_{T_+}(R[t]) \simeq \text{Nil}(H(R))$ and $K_n(H_{T_+}(R[t])) \simeq K_n(R) \oplus \text{Nil}_n(R)$ where we write $\text{Nil}_n(R)$ for $\text{Nil}_n(\mathcal{M}(R))$.

Moreover, the localization sequence 4.4.4 breaks up into short exact sequences

$$0 \to K_n(R[t]) \to K_n(R[t, t^{-1}]) \xrightarrow{\partial} K_{n-1}(\text{Nil}(R)) \to 0.$$

#### 4.5.3. Fundamental theorem of higher $K$-theory [114]. Let $R$ be a ring with identity. Define for all $n \geq 0$ $NK_n(R) := \text{Ker}(K_n(R[t]) \xrightarrow{i_+^*} K_n(R))$ where $i_+^*$ is induced by the augmentation $t \mapsto 1$.

Then there are canonical decompositions for all $n \geq 0$

(i) $K_n(R[t]) \cong K_n(R) \oplus NK_n(R)$.

(ii) $K_n(R[t, t^{-1}]) \cong K_n(R) \oplus NK_n(R) \oplus NK_n(R) \oplus K_{n-1}(R)$.

(iii) $K_n(\text{Nil}(R)) \cong K_n(R) \oplus NK_{n+1}(R)$.

The above decompositions are compatible with a split exact sequence

$$0 \to K_n(R) \to K_n(R[t]) \oplus K_n(R[t, t^{-1}]) \to K_n(R[t, t^{-1}]) \to K_{n-1}(R) \to 0.$$
Higher algebraic K-theory

We close this subsection with the fundamental theorem for G-theory.

**4.5.4. Theorem.** Let $R$ be a Noetherian ring. Then

(i) $G_n(R[t]) \cong G_n(R)$. \\
(ii) $G_n(R[t, t^{-1}]) \cong G_n(R) \oplus G_{n-1}(R)$.

**4.6. Some exact sequences in the K-theory of Waldhausen categories**

**4.6.1. Cylinder functors.** A Waldhausen category has a cylinder functor if there exists a functor $T : \mathcal{A} \rightarrow \mathcal{A}$ together with three natural transformations $p, j_1, j_2$ such that to each morphism $f : A \rightarrow B$, $T$ assigns an object $Tf$ and $j_1 : A \rightarrow Tf$, $j_2 : B \rightarrow Tf$, $p : Tf \rightarrow B$ satisfying certain properties (see [154]).

**Cylinder Axiom.** For all $f, p : Tf \rightarrow B$ is in $w(A)$.

**4.6.2.** Let $\mathcal{A}$ be a Waldhausen category. Suppose that $\mathcal{A}$ has two classes of weak equivalences $\nu(\mathcal{A}), w(\mathcal{A})$ such that $\nu(\mathcal{A}) \subset w(\mathcal{A})$. Assume that $w(\mathcal{A})$ satisfies the saturation and extension axioms and has a cylinder functor $T$ which satisfies the cylinder axiom. Let $\mathcal{A}^w$ be the full subcategory of $\mathcal{A}$ whose objects are those $A \in \mathcal{A}$ such that $0 \rightarrow A$ is in $w(\mathcal{A})$. Then $\mathcal{A}^w$ becomes a Waldhausen category with $\text{co}(\mathcal{A}^w) = \text{co}(\mathcal{A}) \cap \mathcal{A}^w$ and $\nu(\mathcal{A}^w) = \nu(\mathcal{A}) \cap \mathcal{A}^w$.

**4.6.3. Theorem** (Waldhausen fibration sequence, [154]). With the notations and hypothesis of 4.6.2, suppose that $\mathcal{A}$ has a cylinder functor $T$ which is a cylinder functor for both $\nu(\mathcal{A})$ and $\omega(\mathcal{A})$. Then the exact inclusion functors $(\mathcal{A}^w, \nu) \rightarrow (\mathcal{A}, \omega)$ induce a homotopy fibre sequence of spectra

$$K(\mathcal{A}^w, \nu) \rightarrow K(\mathcal{A}, \nu) \rightarrow K(\mathcal{A}, \omega)$$

and hence a long exact sequence

$$K_{n+1}(\mathcal{A}, \omega) \rightarrow K_n(\mathcal{A}^w) \rightarrow K_n(\mathcal{A}, \nu) \rightarrow K_n(\mathcal{A}, \omega) \rightarrow \cdots$$

The next result is a long exact sequence realizing the cofibre of the Cartan map as $K$-theory of a Waldhausen category, see [33].

**4.6.4. Theorem** [33]. Let $R$ be a commutative ring with identity. The natural map $K(\mathcal{P}(R)) \rightarrow K(\mathcal{M}'(R))$ induced by $\mathcal{P}(R) \rightarrow \mathcal{M}'(R)$ fits into a cofibre sequence of spectra $K(R) \rightarrow K(\mathcal{M}'(R)) \rightarrow K(\mathcal{A}, \omega)$ where $(\mathcal{A}, \omega)$ is the Waldhausen category of bounded chain complexes over $\mathcal{M}'(R)$ with weak equivalences being quasi-isomorphisms. In particular, we have a long exact sequence

$$\cdots \rightarrow K_{n+1}(\mathcal{A}, \omega) \rightarrow K_n(R) \rightarrow G'_n(R) \rightarrow K_{n-1}(\mathcal{A}, \omega) \rightarrow \cdots,$$
where

\[ G'_n(R) = K_n(M'(R)). \]

(See 7.1.16 for applications to orders.)

We close this subsection with a generalization of the localization sequence 4.4.3. In 4.6.5 below, the requirement that \( S \) contains no zero divisors is removed.

**4.6.5. Theorem [144]**. Let \( S \) be a central multiplicatively closed subset of a ring \( R \) with identity, \( \text{Perf}(R, S) \) the Waldhausen subcategory of \( \text{Perf}(R) \) consisting of perfect complexes \( M \) such that \( S^{-1}M \) is an exact complex. Then \( K(\text{Perf}(R, S)) \to K(R) \to K(S^{-1}R) \) is a homotopy fibration. Hence there is a long exact sequence.

\[
\cdots K_{n+1}(S^{-1}R) \xrightarrow{\delta} K_n(\text{Perf}(R, S)) \to K_n(R) \to K_n(S^{-1}R) \to \cdots.
\]

**4.7. Excision; relative and Mayer–Vietoris sequences**

**4.7.1.** Let \( \Lambda \) be a ring with identity, \( \underline{a} \) a 2-sided ideal of \( \Lambda \). Define \( F_{\Lambda, \underline{a}} \) as the homotopy fibre of \( B\text{GL}(\Lambda)^+ \to B\text{GL}(\Lambda/\underline{a})^+ \) where \( \text{GL}(\Lambda/\underline{a}) = \text{image}(\text{GL}(\Lambda) \to \text{GL}(\Lambda/\underline{a})) \). Then \( F_{\Lambda, \underline{a}} \) depends not only on \( \underline{a} \) but also on \( \Lambda \).

If we denote \( \pi_n(F_{\Lambda, \underline{a}}) \) by \( K_n(\Lambda, \underline{a}) \), then we have a long exact sequence

\[
\cdots K_{n+1}(\Lambda, \underline{a}) \xrightarrow{\delta} K_n(\text{Perf}(\Lambda, \underline{a})) \to K_n(\Lambda) \to K_n(\Lambda/\underline{a}) \to K_{n-1}(\Lambda, \underline{a}) \to \cdots
\]

from the fibration \( F_{\Lambda, \underline{a}} \to B\text{GL}(\Lambda)^+ \to B\text{GL}(\Lambda/\underline{a})^+ \).

**4.7.2. Definition.** Let \( B \) be any ring without unit and \( \tilde{B} \) the ring with unit obtained by formally adjoining a unit to \( B \), i.e., \( \tilde{B} = \text{set of all } (b, s) \in B \times \mathbb{Z} \) with multiplication defined by \( (b, s)(b', s') = (bb' + sb' + s'b, ss') \).

Define \( K_n(B) \) as \( K_n(\tilde{B}, B) \). If \( \Lambda \) is an arbitrary ring with identity containing \( B \) as a 2-sided ideal, then \( B \) is said to satisfy excision for \( K_n(\Lambda, \underline{a}) \) if the canonical map \( K_n(B) \to K_n(\Lambda, \underline{a}) \) is an isomorphism for any ring \( \Lambda \) containing \( B \). Hence, if in 4.7.1 \( \underline{a} \) satisfies excision, then we can replace \( K_n(\Lambda, \underline{a}) \) by \( K_n(\underline{a}) \) in the long exact sequence (I).

We denote \( F_{\Lambda, \underline{a}} \) by \( F_\underline{a} \).

**4.7.3. We now present another way to understand \( F_\underline{a} \) (see [17]). Let \( \Gamma_n(\underline{a}) := \text{Ker}(\text{GL}_n(\underline{a} \oplus \mathbb{Z})) \to \text{GL}_n(\mathbb{Z}) \) and write \( \Gamma(\underline{a}) = \text{lim } \Gamma_n(\underline{a}) \). Let \( \Sigma_n \) denote the \( n \times n \) permutation matrices. Then \( \Sigma_n \) can be identified with the \( n \)-th symmetric group. Put \( \Sigma = \text{lim } \Sigma_n \).

Then \( \Sigma \) acts on \( \Gamma(\underline{a}) \) by conjugation and so, we can form \( \tilde{\Gamma}(\underline{a}) = \Gamma(\underline{a}) \rtimes \Sigma \). One could think of \( \tilde{\Gamma}(\underline{a}) \) as the group of matrices in \( \text{GL}_n(\underline{a} \oplus \mathbb{Z}) \) whose image in \( \text{GL}_n(\mathbb{Z}) \) is a permutation matrix. Consider the fibration \( B\Gamma(\underline{a}) \to B\tilde{\Gamma}(\underline{a}) \to B(\Sigma) \). Note that \( B(\Sigma), B\tilde{\Gamma}(\underline{a}) \) has an associated +-construction which are infinite loop spaces. Define \( F_\underline{a} \) as the homotopy fibre

\[
F_\underline{a} \to B\tilde{\Gamma}(\underline{a})^+ \to B\Sigma^+.
\]
Higher algebraic $K$-theory

Then, for any ring $\Lambda$ (with identity) containing $\underline{a}$ as a two-sided ideal, we have a map of fibrations

\[
\begin{array}{ccc}
F_{\underline{a}} & \xrightarrow{f_{\Lambda,\underline{a}}} & F_{\Lambda,\underline{a}} \\
\downarrow & & \downarrow \\
B\tilde{\Gamma}(\underline{a})^+ & \rightarrow & B\text{GL}(\Lambda)^+ \\
\downarrow & & \downarrow \\
B\Sigma^+ & \rightarrow & B\text{GL}(\Lambda/\underline{a})^+
\end{array}
\]

4.7.4. Definition. Let $\underline{a}$ be a ring without unit, $S \subseteq \mathbb{Z}$ a multiplicative subset. Say that $\underline{a}$ is an $S$-excision ideal if for any ring $\Lambda$ with unit containing $\underline{a}$ as a two-sided ideal, $f_{\Lambda,\underline{a}}$ induces an isomorphism $\pi_*(F_{\underline{a}}) \otimes S^{-1}\mathbb{Z} \cong \pi_*(F_{\Lambda,\underline{a}}) \otimes S^{-1}\mathbb{Z}$.

4.7.5. Theorem [17]. Let $\underline{a}$ be a ring without unit and $S \subseteq \mathbb{Z}$ a multiplicative set such that $\underline{a} \otimes S^{-1}\mathbb{Z} = 0$ or $\underline{a} \otimes S^{-1}\mathbb{Z}$ has a unit. Then $\underline{a}$ is an $S$-excision ideal and

\[H_n(F_{\underline{a}}, S^{-1}\mathbb{Z}) \cong H_n(\Gamma(\underline{a}); S^{-1}\mathbb{Z}).\]

4.7.6. Examples/Applications.

(i) If $\underline{a}$ is a 2-sided ideal in a ring $\Lambda$ with identity such that $\Lambda/\underline{a}$ is annihilated by some $s \in \mathbb{Z}$, then the hypothesis of 4.7.5 is satisfied by $S = \{s^i\}$ and $\underline{a}$ is an $S$-excision ideal.

(ii) Let $R$ be the ring of integers in a number field $F$, $\Lambda$ an $R$-order in a semi-simple $F$-algebra $\Sigma$, $\Gamma$ a maximal $R$-order containing $\Lambda$. Then there exists an $s \in \mathbb{Z}$, $s > 0$, such that $s\Gamma \subset \Lambda$ and so $\underline{a} = s\Gamma$ is a 2-sided ideal in both $\Lambda$ and $\Gamma$. Since $s$ annihilates $\Lambda/\underline{a}$ (also $\Gamma/\underline{a}$), $\underline{a}$ is an $S$-excision ideal and so, we have a long exact Mayer–Vietoris sequence

\[
\begin{align*}
&\rightarrow K_{n+1}(\Gamma/\underline{a})\left(\frac{1}{s}\right) \rightarrow K_n(\Lambda)\left(\frac{1}{s}\right) \rightarrow K_n(\Lambda/\underline{a})\left(\frac{1}{s}\right) \oplus K_n(\Gamma)\left(\frac{1}{s}\right) \\
&\rightarrow K_n(\Gamma/\underline{a})\left(\frac{1}{s}\right) \rightarrow,
\end{align*}
\]

where we have written $A(\frac{1}{s})$ for $A \otimes \mathbb{Z}(\frac{1}{s})$ for any Abelian group $A$.

(iii) Let $\Lambda$ be a ring with unit and $K_n(\Lambda, \mathbb{Z}/r)$ the $K$-theory with mod-$r$ coefficients (see [16]). Let $S = \{s \in \mathbb{Z} \mid (r, s) = 1\}$. Then multiplication by $s \in S$ is invertible on $K_n(\Lambda, \mathbb{Z}/r)$. Hence for an $S$-excision ideal $\underline{a} \subset \Lambda$, $\pi_* (F_{\Lambda,\underline{a}}) \otimes S^{-1}\mathbb{Z} \cong \pi_* (F_{\underline{a}}) \otimes S^{-1}\mathbb{Z}$ implies that $\pi_* (F_{\Lambda,\underline{a}}; \mathbb{Z}/r) \cong \pi_* (F_{\underline{a}}; \mathbb{Z}/r)$.

If we write $\mathbb{Z}(r)$ for $S^{-1}\mathbb{Z}$ in this situation, we have that $K_n(\Lambda, \mathbb{Z}/r)$ satisfies excision on the class of ideals $\underline{a}$ such that $\underline{a} \otimes \mathbb{Z}(r) = 0$ or $\underline{a} \otimes \mathbb{Z}(r)$ has a unit.
5. Higher $K$-theory and connections to Galois, étale and motivic cohomology theories

5.1. Higher $K$-theory of fields

The importance of $K$-theory of fields lies not only in its connections with such areas as Brauer groups; Galois, étale, motivic cohomologies; symbols in arithmetic; zeta functions; Bernoulli numbers, etc., but also because understanding $K$-theory of fields also helps to understand $K$ theory of other rings, e.g., $K$-theory of rings of integers in number fields, $p$-adic fields, algebras over such rings (e.g., orders) as well as $K$-theory of varieties and schemes.

First, we present some well known calculations of the $K$-theory of some special fields. Interested readers may see [41,132] for a comprehensive survey of this topic.

5.1.1. THEOREM [112].
(a) Let $F_q$ be a finite field of order $q$. Then $K_{2n}(F_q) = 0$ and $K_{2n-1}(F_q)$ is a cyclic group of order $q^n - 1$ for all $n \geq 1$.
(b) If $E$ is finite extension of $F_q$, then the natural map $K_{2n-1}(F_q) \to K_{2n-1}(E)$ is injective. Moreover the automorphisms of $E$ over $F_q$ act on $K_{2n-1}(E)$ by multiplication by $q^n$. If $E/F_q$ is Galois then the natural map $K_{2n-1}(F_q) \to K_{2n-1}(E)$ is an isomorphism.
(c) If $E$ is the algebraic closure of $F_q$, then for all $n \geq 1$, $K_{2n}(E) = 0$ and $K_{2n-1}(E) \cong (\mathbb{Q}/\mathbb{Z}) = \bigoplus (\mathbb{Q}_l/\mathbb{Z}_l)$.

On $K_{2n}(F)$, the Frobenius automorphism of $F_q$ acts through multiplication by $q^n$.

For more general fields we have

5.1.2. THEOREM [132]. Let $F$ be an algebraically closed field. Then for $n \geq 1$, $K_{2n}(F)$ is uniquely divisible and $K_{2n-1}(F)$ is the direct sum of a uniquely divisible group and a group isomorphic to $\mathbb{Q}/\mathbb{Z}$.

5.1.3. THEOREM [34]. Let $F$ be a field of positive characteristic $p$. Then $K_n(F)$ has no $p$-torsion.

5.1.4. For any commutative ring $A$, we have a composition $\otimes : \text{GL}_n(A) \times \text{GL}_p(A) \to \text{GL}_{np}(A)$ which induces $\gamma_{n,p} : B\text{GL}_n(A)^+ \times B\text{GL}_p(A)^+ \to B\text{GL}_{n,p}(A)^+ \to B\text{GL}(A)^+$. Also $\gamma_{n,p}$ induces $B\text{GL}(A)^+ \times B\text{GL}(A)^+ \to B\text{GL}(A)^+$ which induces a product: $K_n(A) \times K_m(A) \to K_{n+m}(A)$. The product $\ast$ endows $\bigoplus_{n=1}^\infty K_n(A)$ with the structure of a skew-commutative graded ring. If $m = n = 1$, then the product coincides with the Steinberg symbol (up to sign) discussed in [79]. This is in particular true for fields.
5.1.5. **Theorem [144].** Let $F$ be a number field with $r_1$ real places and $r_2$ complex places. Then

$$
\text{rank } K_n(F) = \begin{cases} 
1, & n = 0, \\
\infty, & n = 1, \\
0, & n = 2k, k > 0, \\
r_1 + r_2, & n = 4k + 1, k > 0, \\
r_2 & n = 4k + 3.
\end{cases}
$$

5.1.6. **Remarks.** The above is also true for $K_n(R)$ if $R$ is the ring of integers in $F$ for $n \geq 2$ and $n = 0$. It is classical that $\text{rank } K_1(R) = r_1 + r_2 - 1$ (see [9]). It is a result of Quillen that $K_n(R)$ is finitely generated for all $n \geq 1$ (see [115]).

5.1.7. **Theorem [41].** $K_2^\mathbb{Q} \simeq \mathbb{Z}/2 \oplus (\bigoplus_p U(\mathbb{Z}/p))$.

5.1.8. **Theorem [89].** $K_3^\mathbb{Q} \simeq K_3(\mathbb{Z}) \simeq \mathbb{Z}/48$.

We now recall the definition of Milnor $K$-theory in the present context.

5.1.9. **Definition.** Let $F$ be any field, $F^*$ the group of non-zero elements of $F$. Define $K_n^M(F) = T(F^*)/J$ where $T(F)$ is the tensor algebra over $F^*$ and $J$ the ideal generated by all $x \otimes (1 - x)$. Thus $K_n^M(F) = (F^* \otimes F^* \otimes \cdots \otimes F^*)/J_n$ where $J_n$ is the subgroup of $F^* \otimes F^* \otimes \cdots \otimes F^*$ generated by all $a_1 \otimes a_2 \otimes \cdots \otimes a_n$ such that $a_i + a_j = 1$ for some $i \neq j$. The image of $x_1 \otimes x_2 \otimes \cdots \otimes x_n$ in $K_n^M(F)$ is denoted by $\{x_1, x_2, \ldots, x_n\}$. Note that by the Matsumoto theorem, $K_2(F) = K_2^M(F) = F^* \times F^*/(x \otimes (1 - x))$. Also, $K_1(F) = K_1^M(F) = F^*$.

We also have a map $\varphi: K_n^M(F) \to K_n(F)$ defined by: $\{x_1, x_2, \ldots, x_n\} \to x_1 \ast x_2 \cdots \ast x_n$ where $\ast$ is the product defined in 5.1.4.

5.1.10. **Theorem [132].** The kernel of $\varphi: K_n^M(F) \to K_n(F)$ is annihilated by $(n - 1)!$.

5.1.11. **Theorem [134].** Let $F$ be an infinite field. Then, there exists an isomorphism $H_n(\text{GL}_n(F)) \simeq H_n(\text{GL}_{n+1}(F), \mathbb{Z}) \simeq \cdots \simeq H_n(\text{GL}(F), \mathbb{Z})$ and an exact sequence

$$
H_n(\text{GL}_{n-1}(F), \mathbb{Z}) \to H_n(\text{GL}_n(F), \mathbb{Z}) \to K_n^M(F) \to 0.
$$

We also have a map (cup product)

$$
H_1(\text{GL}_1(F), \mathbb{Z}) \otimes \cdots \otimes H_1(\text{GL}_1(F), \mathbb{Z}) \to H_n(\text{GL}_n(F), \mathbb{Z}),
$$

that is

$$
F^* \otimes \cdots \otimes F^* \to H_n(\text{GL}_n(F), \mathbb{Z})$$

$\text{Higher algebraic K-theory}$
as well as a map $\delta : K_n^M(F) \to H_n(\text{GL}_n(F), \mathbb{Z})$ which makes the following diagram commutative

$$
\begin{array}{ccc}
F^* \otimes \cdots \otimes F^* & \rightarrow & H_n(\text{GL}_n(F), \mathbb{Z}) \\
\downarrow^\delta & & \\
K_n(F) & & \\
\end{array}
$$

5.1.12. THEOREM [134]. There exists a map $\psi : K_nF \to K_n^M(F)$ defined by

$$
\begin{align*}
K_n(F) &= \pi_n(B\text{GL}(F)^+) \to H_n(\text{GL}(F), \mathbb{Z}) \simeq H_n(\text{GL}_n(F), \mathbb{Z}) \\
&\to H_n(\text{GL}_n(F), \mathbb{Z})/\text{Im}(H_n(\text{GL}_{n-1}(F), \mathbb{Z})) \simeq K_n^M(F)
\end{align*}
$$

such that

(i) $\varphi \circ \psi = c_\cdot$, 
(ii) $\psi \circ \varphi = \text{multiplication by } (-1)^{n-1}(n-1)!$.

5.1.13. THEOREM [132]. Let $F$ be a number field with $r_1$ real places. Then $K_n^M(F) \cong (\mathbb{Z}/2)^{r_1}$ for $n \geq 3$.

5.1.14. THEOREM [132]. Let $F$ be a real number field. Then the map $K_4^M(F) \to K_4(F)$ is not injective and the map $K_n^M(F) \to K_n(F)$ is zero for $n \geq 5$.

5.1.15. DEFINITION.

$$
K_n^{\text{ind}}(F) := \text{Coker}(K_n^M(F) \xrightarrow{\psi} K_n(F)).
$$

NOTE. It is well known that $K_3^{\text{ind}}(F) \neq 0$ for all fields, i.e. $\varphi$ is never surjective. Also the map is not injective in general.

5.1.16. THEOREM [132]. Let $L/F$ be a field extension such that $F$ is algebraically closed in $L$. Then the map $K_3^{\text{ind}}F \to K_3^{\text{ind}}L$ in an isomorphism on the torsion and cotorsion. Hence $K_3^{\text{ind}}L/K_3^{\text{ind}}F$ is uniquely divisible.

5.2. Galois cohomology

5.2.1. Let $G$ be a profinite group. A discrete $G$-module is a $G$-module $A$ such that if $A$ is given the discrete topology the multiplication map $G \times A \to A$ is continuous. If $A$ is a discrete $G$-module then for every $a \in A$, the stabilizer $U$ of $a$ is an open subgroup of $G$ and

$$
a \in A^U = \{a \in A \mid gA = A \text{ for all } g \in U\}.
$$
Note that $A$ is a discrete $G$-module iff $\bigcup A^U = A$. Note also that if $G$ is a finite group, then every $G$-module is discrete.

The category $C_G$ of discrete $G$-modules is an Abelian subcategory of $G\text{-mod}$ with enough injectives. The right derived functor of the left exact functor $C_G \to \text{Ab}; A \to A^G$ where $A^G = \{ a \in A | ga = a \ \forall g \in G \}$ are the cohomology groups of $G$ with coefficients in $A$ and denoted by $H^*(G, A)$, see [165].

5.2.2. Let $A$ be a discrete $G$-module ($G$ a profinite group), $C^n(G, A)$ the Abelian group of continuous maps $G^n \to A$ and $C^n(G, A) = \lim_{\leftarrow} (C^n(G/U, A^U))$, where $U$ runs through all open normal subgroups of $G$.

5.2.3. THEOREM. If $G$ is a profinite group and $A$ a discrete $G$-module

$$H^n(G, A) \cong H^n(C^n(G, A))$$

$$\cong \lim_{U} H^n(G/U, A^U),$$

where $U$ runs through all open normal subgroups of $G$.

5.2.4. EXAMPLE. Let $F_s$ be the separable closure of a field $F$, i.e. $F_s$ is the subfield of the algebraic closure $\overline{F}$ consisting of all elements separable over $F$ and $F_s = \overline{F}$ if $\text{char}(F) = 0$.

Krull’s theorem (see [165]) says that the group $\text{Gal}(F_s/F) \cong \lim_{\leftarrow} \text{Gal}(F_i/F)$ where $F_i$ runs through all finite Galois extensions of $F$. As such $\text{Gal}(F_s/F)$ is a profinite group and $F_s$ is a discrete $\text{Gal}(F_s/F)$-module.

We shall denote $H^n(\text{Gal}(F_s/F), F_s)$ by $H^n(F, F_s)$.

5.2.5. Let $F$ be field and $B_r(F)$ the Brauer group of $F$, i.e. the group of stable isomorphism classes of central simple $F$-algebras with multiplication given by the tensor product of algebras. See [119].

A central simple $F$-algebra $A$ is said to be split by an extension $E$ of $F$ if $E \otimes A$ is $E$-isomorphic to $M_r(E)$, the algebra of $r \times r$ matrices over $E$, for some positive integer $r$. It is well known (see [60]) that such an $E$ can be taken as some finite Galois extension of $F$. Let $B_r(F, E)$ be the group of stable isomorphism classes of $E$ split central simple algebras. Then $B_r(F) := B_r(F, F_s)$ where $F_s$ is the separable closure of $F$.

5.2.6. THEOREM [119]. Let $E$ be a Galois extension of a field $F$, $G = \text{Gal}(E/F)$. Then there exists an isomorphism $H^2(G, E^*) \cong B_r(F, E)$. In particular $B_r(F) \cong H^2(G, F_s^*)$ where $G = \text{Gal}(F_s/F) = \lim_{\rightarrow} \text{Gal}(E_i/F)$, where $E_i$ runs through the finite Galois extensions of $F$.

5.2.7. Now, for any $m > 0$, let $\mu_m$ be the group of $m$-th roots of 1, $G = \text{Gal}(F_s/F)$, then we have the Kummer sequence of $G$-modules

$$0 \to \mu_m \to F_s^* \to F_s^* \to 0$$
from which we obtain an exact sequence of Galois cohomology groups

\[ F^* \xrightarrow{m} F^* \rightarrow H^1(F, \mu_m) \rightarrow H^1(F, F^*_s) \rightarrow \cdots, \]

where \( H^1(F, F^*_s) = 0 \) by the Hilbert theorem 90. So we obtain an isomorphism \( \chi_m : F^*/mF^* \cong F^* \otimes \mathbb{Z}/m \rightarrow H^1(F, \mu_m) \).

Now, the composite

\[ F^* \otimes \mathbb{Z} F^* \rightarrow (F^* \otimes \mathbb{Z} F^*) \otimes \mathbb{Z}/m \rightarrow H^1(F, \mu_m) \otimes H^1(F, \mu_m) \]

\[ \rightarrow H^2(F, \mu_m^\otimes 2) \]

is given by \( a \otimes b \rightarrow \chi_m(a) \cup \chi_m(b) \) (where \( \cup \) is the cup product) which can be shown to be a Steinberg symbol inducing a homomorphism \( g_{2,m} : K_2(F) \otimes \mathbb{Z}/m \otimes H^2(F, \mu_m^\otimes 2). \)

We then have the following result due to A.S. Merkurjev and A.A. Suslin, see [100].

5.2.8. Theorem [100]. Let \( F \) be a field, \( m \) an integer \( > 0 \) such that the characteristic of \( F \) is prime to \( m \). The map

\[ g_{2,m} : K_2(F)/mK_2(F) \rightarrow H^2(F, \mu_m^\otimes 2) \]

is an isomorphism where \( H^2(F, \mu_m^\otimes 2) \) can be identified with the \( m \)-torsion subgroup of \( Br(F) \).

5.2.9. Remarks. By generalizing the process outlined in 5.2.7 above, we obtain a map

\[ g_{n,m} : K_n^M(F)/mK_n^M(F) \rightarrow H^n(F, \mu_m^\otimes n). \]

(1) It is a conjecture of Bloch and Kato that \( g_{n,m} \) is an isomorphism for all \( F, m, n \). So, 5.2.8 is the \( g_{2,m} \) case of the Bloch–Kato conjecture when \( m \) is prime to the characteristic of \( F \). Furthermore, A. Merkurjev proved that 5.2.8 holds without any restriction of \( F \) with respect to \( m \).

It is also a conjecture of Milnor that \( g_{n,2} \) is an isomorphism. In 1996, V. Voevodsky proved that \( g_{n,2} \) is an isomorphism for any \( r \), see [106].
(3) If \( \{ T_\alpha \to T \}_\alpha \in \text{Cov}(T) \) and \( g: V \to T \) is an arbitrary morphism, then for all \( \alpha \), the fibre product \( T_\alpha \times_T V \) exists and \( \{ T_\alpha \times_T V \} \) is a cover.

A sheaf \( F \) on a site \( S \) with values in a suitable category \( \mathcal{A} \) is a contravariant functor (presheaf) \( F: S \to \mathcal{A} \) such that for every covering \( \{ T_\alpha \to T \}_\alpha \), the sequence

\[
F(T) \to \prod_\alpha F(T_\alpha) \Rightarrow \prod_\beta \prod_\alpha F(T_\alpha \times T_\beta)
\]

is a difference kernel or equalizer diagram in \( \mathcal{A} \).

If, for example, the category \( \mathcal{A} \) is Abelian, then the category of sheaves on the site \( S \) has enough injectives; i.e. every object \( A \) or complex \( A^* \) has an injective resolution \( A \to I \) or \( A^* \to I^* \). The cohomology (resp. hypercohomology) of \( X \) is then the cohomology of the complex \( I(X) \) (resp. \( I^*(X) \)).

5.3.2. Zariski cohomology. Let \( \text{Sch} \) be a suitable category of schemes, \( S \in \text{Sch} \). The big Zariski site \( S_{\text{zar}} \) (on \( S \)) is the category such that \( \text{ob}(S_{\text{zar}}) = \text{schemes over } S \) and for \( X \in \text{Sch} \), a covering family is a Zariski open cover by open subschemes of \( X \). If \( \mathcal{A} = \text{Sets} \) or \( \mathbb{Z}\text{-mod} \), etc., then a sheaf on \( \text{Sch} \) is a presheaf which when restricted to any scheme \( X \in \text{Sch} \) is actually a sheaf in the usual Zariski topology of \( X \).

A cohomology theory on \( \text{Sch} \) consists of a bounded below complex of sheaves \( \Gamma^* \) (which we could assume to be a complex of injectives) on \( S_{\text{zar}} \). If \( X \) is a scheme over \( S \), then the cohomology of \( X \) with coefficient in \( \Gamma^* \) is defined as the hypercohomology of \( X \) with coefficient in the restriction of \( \Gamma^* \) to \( X \) and is denoted by \( H^*(X, T^*) \).

5.3.3. Étale cohomology. Let \( \text{Sch} \) be a suitable category of schemes, \( S \in \text{Sch} \). The big étale site \( S_{\text{ét}} \) of \( S \) is defined as follows: \( \text{ob}(S_{\text{ét}}) = \text{schemes } X \in \text{Sch} \) over \( S \). Coverings are étale coverings; i.e. families \( \{ U_i \stackrel{p_i}{\to} X \} \) of maps where each \( p_i \) is an étale mapping, i.e. each \( p_i \) is flat and unramified (see [101]).

For \( X \in \text{Sch} \), we shall write \( H^n_{\text{ét}}(X, M) \) for the étale hypercohomology of \( X \) where \( M \) is a complex of sheaves of the étale topology. If \( R \) is a commutative ring with identity, we shall write \( H^n_{\text{ét}}(R, M) \) for \( H^n_{\text{ét}}(\text{Spec}(R), M) \). (Here \( M \) is a complex of sheaves on étale site of \( \text{Spec}(R) \).)

5.3.4. Remarks and notations. In 5.3.3, \( M \) could be finite, torsion or profinite and could take any of the following forms.

In what follows \( \ell \) is a rational prime, \( i \) a positive integer

(i) \( G_m = G_m(R) = R^* = \text{units of } R \).
(ii) \( \mu(R) = \text{torsion subgroup of } R^* = \text{roots of unity in } R \).
(iii) \( \mu_{\ell^i}(R) = \ell^i\text{-th roots of unity in } R \).
(iv) \( \mu_{\ell^\infty}(R) = \text{all } \ell\text{-th power roots of unity in } R \).
(v) \( \mu_{\ell^i} = \mathbb{Z}/\ell^i \) (in additive notation).
(vi) \( \mu_{\ell^i}(i) = \mathbb{Z}/\ell^i(i) \) where \( \mu_{\ell^i}(i) = \mu_{\ell^i} = \mu_{\ell^i} \otimes \cdots \otimes \mu_{\ell^i} \).
\( i \text{ times} \).
(vii) \( W(i) = \lim Z/\ell^v(i) \), the direct limit taken over the injections \( Z/\ell^v(i) \to Z/\ell^{v+1}(i) \). Note that \( W(i) \) is a discrete torsion group.

(viii) \( Z_\ell(i) = \lim Z/\ell^v(i) \), the inverse limit taken over projections \( \pi: Z/\ell^{v+1}(i) \to Z/\ell^v(i) \) is a profinite group.

5.3.5. Examples.

(i) For any commutative ring \( R \) with identity, \( \text{Pic}(R) \simeq H^1_q(R,\mathbb{G}_m) \). If \( R \) is a field, \( \text{Pic}(F) \simeq H^1(F,\mathbb{G}_m) = 0 \) (see [166]).

(ii) Let \( F \) be a number field with ring of integers \( O_F \), \( \ell \) a prime, \( O'_F = O_F(1/\ell) \). Then \( H^n_q(O'_F,\mathbb{Z}_\ell(n)) \) is finite and \( = 0 \) for almost all primes \( \ell \). Also

\[
\text{rk}_2 H^1_q(O'_F,\mathbb{Z}_\ell(n)) = \begin{cases} 
1 + r_2 & \text{if } n \text{ is odd} > 1, \\
r_2 & \text{if } n \text{ is even}.
\end{cases}
\]

(iii) Let \( E \) be a \( p \)-adic field, then each \( H^n_q(E, W(i)) \) is a discrete torsion group, \( H^n_q(E, \mathbb{Z}_\ell(i)) \) is a profinite group.

In particular for \( n \neq 0, 1, 2 \), \( H^n_q(E, W(i)) = 0 \) and \( H^n_q(E, \mathbb{Z}_\ell(i)) = 0 \), see [166]. If \( E \) has residue field \( \mathbb{F}_q \) of characteristic \( p \neq \ell \), then for all \( i > 1 \),

\[
H^n_q(E, W(i)) \simeq \begin{cases} 
\mathbb{Z}/w_i(\mathbb{F}_q), & n = 0, \\
\mathbb{Z}/w_{i+1}(\mathbb{F}_q), & n = 1, \\
0, & \text{otherwise}
\end{cases}
\]

(see [166]). Also see 6.2.9 for the definition of \( \omega_0(F) \).

(iv) Let \( F \) be a number field with separable closure \( F_s \), \( \ell \) a prime. Then \( H^n_q(\text{Gal}(F_s/F),\mu_\ell^{\otimes n}) \simeq H^n_q(\text{Spec}(F),\mu_\ell^{\otimes n}) \), see [166].

5.3.6. Let \( \ell \) be a rational prime, \( F \) a number field with ring of integers \( O_F \). An extension \( E \) of \( F \) is said to be unramified outside \( \ell \) when given any prime ideal \( p \) of \( O_F \) such that \( \ell \) does not divide \( p \), we have \( pO_F = p_1^{r_1} p_2^{r_2} \cdots p_n^{r_n} \) where \( p_1, p_2, \ldots, p_n \) are distinct prime ideals in \( O_F \).

Let \( \overline{F} \) be the algebraic closure of \( F \) and \( \Gamma \) the union of all finite extensions \( E \) of \( F \) which are unramified outside \( \ell \).

Let \( F(\mu_\ell^\infty) = \bigcup_{v \geq 1} F(\mu_\ell^v) \) be the maximal \( \ell \)-cyclotomic extension of \( F \). Then \( F(\mu_\ell^\infty) \subset \Gamma \).

For any \( i \in \mathbb{Z} \), the Galois \( \text{Gal}(F(\mu_\ell^\infty)/F) \)-module \( \mathbb{Z}_\ell(i) \) is defined as the additive group \( \mathbb{Z}_\ell \) equipped with Galois action \( g \cdot x = \epsilon(g)x \) where \( \epsilon: \text{Gal}(F(\mu_\ell^\infty)/F) \to \mathbb{Z}_\ell^* \) is defined by \( \epsilon(\xi) = \xi^\ell \) for all \( \xi \in \overline{F} \) such that \( \xi^\ell = 1 \) for some \( \ell \geq 1 \).

One can regard \( \mathbb{Z}_\ell(i) \) as a continuous Galois \( \Gamma/F \)-module via the projection \( \text{Gal}(\Gamma/F) \to \text{Gal}(F(\mu_\ell^\infty)/F) \). One can identify the Galois cohomology groups \( H^k(\text{Gal}(\Gamma/F),\mathbb{Z}_\ell(i)) \) with \( H^k_q(O'_F,\mathbb{Z}_\ell(i)) \); i.e. \( H^k_q(O'_F,\mathbb{Z}_\ell(i)) \simeq H^k(\text{Gal}(\Gamma/F),\mathbb{Z}_\ell(i)) \) where \( O'_F = O_F(1/\ell) \). Some authors use this Galois cohomology group to define \( H^k_q(O'_F,\mathbb{Z}_\ell(i)) \).
5.4. Motivic cohomology

5.4.1. Let $\text{Sm}(k)$ be the category of smooth quasi-projective varieties over a field $k$ (usually of char 0). For $X, Y \in \text{Sm}(k)$, let $c(X, Y) := \text{free Abelian group on the set of closed irreducible subvarieties } Z \subset X \times_k Y \text{ for which the projection } Z \to X \text{ is finite and surjective onto an irreducible component of } X$. Call $c(X, Y)$ the group of finite correspondences from $X$ to $Y$.

Define a category $\text{Sm Cor}(k)$ as follows

$$\text{ob}(\text{Sm Cor}(k)) = \text{ob}(\text{Sm}(k)),$$

$$\text{Hom}_{\text{Sm Cor}(k)}(X, Y) := c(X, Y).$$

For any smooth $k$-variety $X$, define a functor $L[X] : (\text{Sm Cor}(k))^\text{op} \to \text{Ab}$:

$$Y \mapsto c(Y, X).$$

5.4.2. Let $F : (\text{Sm}(k))^\text{op} \to \text{Ab}$ be a presheaf of Abelian groups. Define a singular chain complex $C^\bullet(F)$ as the presheaf of complexes $C^\bullet(F) : \text{Sm}(k)^\text{op} \to \text{chain complexes}$:

$$X \mapsto C^\bullet(F)(X) \text{ where } C^\bullet(X)(\text{spec}(k)) \text{ is the chain complex associated to the HS}^\bullet(X) \text{, the singular homology groups of } X := \text{homology groups of the chain complex } C^\bullet(X)(\text{spec}(k)).$$

Now define a presheaf of chain complexes $C^\bullet L[X] : (\text{Sm Cor}(k))^\text{op} \to \text{chain complexes}$:

$$Y \mapsto C^\bullet L(X)(Y).$$

Now let $G_m = \mathbb{A}^1 \setminus \{0\}$ and for $i \geq 0$, let $L_i = \text{cokernel of the morphism } \bigoplus L((G_m)^{i-1}) \to L((G_m)^i)$ induced by $(G_m)^{i-1} \subset (G_m)^i$ which puts the $j$-th co-ordinate equal to 1.

Now let $Z(i)$ be the presheaf of complexes on $\text{Sm}(k)$ given by $Z(i) = C^\bullet(L_i)[-1]$, the $i$-th desuspension of $C^\bullet(L_i)$. (See [106].)

5.4.3. Definition. Let $X$ be a smooth $k$-variety, $i \geq 0$. The Zariski hypercohomology of $X$ with coefficients in $Z(i)$ is denoted by $H^\bullet_B(X, Z(i))$. The bigraded Abelian group $H^\bullet_B(X, Z(-))$ is called the Beilinson motivic cohomology (or just motivic cohomology) with coefficients in $Z$. It is usual to denote $H^\bullet_B(X, Z(\ast))$ by $H^\bullet_M(X, Z(\ast))$.

For any Abelian group $A$, $X \in \text{Sm}(k)$, define the motivic cohomology of $X$ with coefficients in $A$ by $H^\bullet_M(X, A(\ast))$ where $A(\ast) = \mathbb{Z}(\ast) \otimes A$.

5.4.4. Examples.

(i) For any $X \in \text{Sm}(k)$,

$$H^n_M(X, Z(\ast)) = \begin{cases} 0 & \text{if } n \neq 0, \\ \mathbb{Z}\pi_0(X) & \text{if } n = 0, \end{cases}$$

where $\pi_0(X)$ = set of connected components of $X$.

(ii)

$$H^n_M(X, Z(i)) = \begin{cases} 0 & \text{if } n \neq 1, 2, \\ O_X^\ast & \text{for } n = 1, \\ \text{Pic}(X) & \text{for } n = 2. \end{cases}$$
(iii) \( k^* \equiv H^1_M(\text{Spec}(k), Z(1)) \).

(iv) Let \( H^*_{AM}(k) := \text{graded commutative ring equal in degree } n \text{ to } H^n_{AM}(\text{Spec}(k), Z(n)) \).

Then for any field \( k \), \( K^M_n(k) \simeq H^*_{AM}(k) \). (See [106].)

(v) For any field \( k \) and any integers \( n \geq 0, m > 0 \), we have

\[
K^n_M(k)/m \simeq H^n_{AM}(\text{Spec}(k), Z/m(n))
\]

(see [106]).

(vi) Let \( \ell \) be a rational prime, \( F \) a global field with ring of integers \( O_F, O'_F = O_F(\frac{1}{\ell}) \).

Then there are isomorphisms \( H^i_{AM}(O_F, Z(n)) \otimes \mathbb{Z}_\ell \simeq H^i_{et}(O'_F, Z(n)) \) for all primes \( \ell, n \geq Z \) and \( i = 1, 2 \) (see [106]).

(vii) Let \( F \) be a field of characteristic \( \neq p \). If the Bloch–Kato conjecture holds, then there is an isomorphism

\[
H^i_{AM}(F, Z/\ell^n(n)) \simeq \left\{ \begin{array}{cl} H^i_{et}(F, Z/\ell^n(n)), & 0 \leq i \leq n, \\ 0, & \text{otherwise} \end{array} \right. 
\]

(see [106]).

5.4.5. Let \( A \) be an Abelian group and \( X \) any smooth \( k \)-variety. Let \( H^*_{LM}(X, A(i)) \) be the hypercohomology groups of \( X \) of the presheaves \( A(i) \) in the étale topology. Call \( H^*_{LM}(X, A(i)) \) the Lichtenbaum (or étale) cohomology groups of \( X \) with coefficients in \( A \).

5.4.6. Theorem [106]. Let \( k \) be a field, \( X \) a smooth \( k \)-variety, \( A \) a \( Q \)-vector space. Then

\[
H^*_{LM}(X, A(i)) \to H^*_{LM}(X, A(i^*)).
\]

5.4.7. Theorem [106]. Let \( A \) be a torsion Abelian group of torsion prime to \( \text{char}(k) \).

Let \( A_{\tilde{a}} \) be the constant étale sheaf associated to \( A \). For \( i \geq 0 \), let \( A_{\tilde{a}}(i) = i\text{-th Tate twist} \).

Then, for all \( i \geq 0 \) and any smooth \( k \)-variety \( X \), we have

\[
H^i_{LM}(X, A(i)) \simeq H^i_{et}(X, A_{\tilde{a}}(i)).
\]

We close this subsection with the following result

5.4.8. Theorem [106]. Let \( k \) be a field which admits resolution of singularities (e.g., \( \text{char}(k) = 0 \)), \( \ell \) a prime number not equal to \( \text{char}(k) \). Then for any integer \( n \geq 0 \), the following are equivalent

(i) For any field extension \( k \subset F \) of finite type,

\[
g_{n,\ell}: K^n_M(F)/\ell K^n_M(F) \simeq H^n(\text{Gal}(F_s/F), \mu_{\ell^n}).
\]

(ii) For any smooth \( K \)-variety \( X \), and any integer \( m \in \{0, 1, \ldots, n + 1\} \), we have an isomorphism

\[
H^m_{AM}(X, \mathbb{Z}_\ell(n)) \simeq H^m_{LM}(X, \mathbb{Z}_\ell(n)).
\]
(iii) For any field extension $k \subset F$ of finite type we have

$$H^{n+1}_L(\text{Spec}(F); \mathbb{Z}_\ell(n)) = 0.$$ 

5.4.9. Remarks. In 5.4.8, (i) is a special case of the Bloch–Kato conjecture which is believed to have now been proved by Rost and Voevodsky. The case $g_{2,\ell}$ was proved by Voevodsky in 1997 (see [106]). (ii) was a conjecture of S. Lichtenbaum.

5.5. Connections to Bloch’s higher Chow groups

5.5.1. Let $k$ be a field and $\Delta^n = \text{Spec}(R)$ where $R = k[t_0 \ldots t_n]/(\Sigma t_j - 1)$. Let $X$ be a quasi-projective variety, $z^i(X, n)$ free Abelian group on the set of codim $i$ subvarieties $Z$ of $X \times_k \Delta^n$ which meet all faces properly (see [106]). Call $z^i(X, *)$ Bloch’s cycle group.

Define $CH^i(X) = (\pi_j z^i(X, *)) = H_j$ of the associated chain complex.

5.5.2. Remarks/Examples.

(i) For $j = 0$, $CH^0(X)$ is the classical Chow group of codim $i$ cycles in $X$ modulo rational equivalence.

(ii) $CH^i_j(X) = 0$ for $i > j + \text{dim}(X)$. See [106].

(iii) $CH^i_j(X) \simeq CH^i_j(X[t])$ where $X(t) = X \times_k \text{Spec}(k[t])$, i.e. the higher Chow groups are homotopy invariant.

(iv) If $X$ is smooth

$$CH^j(X) = \begin{cases} \text{Pic}(X), & j = 0, \\ H^0(X, \mathcal{O}_X), & j = 1, \\ 0, & \text{otherwise}. \end{cases}$$

5.5.3. Theorem [106]. For every field $k$ and all $n$,

$$K^M_n(k) \cong CH^0_n(k).$$

Here $X = \text{Spec}(k)$.

5.5.4. Theorem [106]. Let $X$ be a smooth scheme of finite type over a field $k$. Then for $n \geq 0$, there is an isomorphism

$$K_n(X)_Q \cong \bigoplus_{d \geq 0} CH^d_n(X)_Q,$$

where $A_Q := A \otimes Q$ for any Abelian group $A$.

5.5.5. Theorem [106]. Let $X$ be a smooth algebraic variety. There exists a spectral sequence

$$E_2^{p,q} = CH^{-q,-q}(X) \Rightarrow K_{-p-q}(X).$$
5.5.6. Theorem [106]. Let $k$ be a field of characteristic zero (more generally a field which admits the resolution of singularities). Then, for each smooth scheme $X$ of finite type over $k$, there are natural isomorphisms

$$H^p_{\mathcal{M}}(X, \mathbb{Z}(q)) \simeq CH^p_{2q-p}(X).$$

6. Higher $K$-theory of rings of integers in local and global fields

6.1. Some earlier general results on the higher $K$-theory of ring of integers in global fields

The first result is due to D. Quillen.

6.1.1. Theorem [40,115]. Let $O_F$ be the ring of integers in a global field $F$. Then for all $n \geq 1$, the Abelian groups $K_n(O_F)$ are finitely generated.

The next result, due to A. Borel, computes the rank of $K_n(O_F)$, $K_n(F)$ for $n \geq 2$.

6.1.2. Theorem [13]. Let $F$ be a number field and let us write $[F : \mathbb{Q}] = r_1 + 2r_2$, where $r_1$ is the number of distinct embeddings of $F$ into $\mathbb{R}$ and $r_2$ the number of distinct conjugate pairs of embeddings of $F$ into $\mathbb{C}$ with image not contained in $\mathbb{R}$.

(i) If $R$ denotes either the number field $F$ or its ring of algebraic integers $O_F$, then the rational cohomology of the special linear group $SL(R)$ is given by

$$H^*(SL(R); \mathbb{Q}) \simeq \left( \bigotimes_{1 \leq j \leq r_1} A_j \right) \otimes \left( \bigotimes_{1 \leq k \leq r_2} B_k \right),$$

where $j$ runs over all distinct embeddings of $F$ into $\mathbb{R}$, $k$ over all distinct conjugate pairs of embeddings of $F$ into $\mathbb{C}$ with image not contained in $\mathbb{R}$, and where $A_j$ and $B_k$ are the following exterior algebras:

$$A_j = \Lambda_{\mathbb{Q}}(x_5, x_9, x_{13}, \ldots, x_{4l+1}, \ldots) \quad \text{and} \quad B_k = \Lambda_{\mathbb{Q}}(x_3, x_5, x_7, \ldots, x_{2l+1}, \ldots),$$

where $\deg(x_j) = j$.

(ii) If $R$ denotes either the number field $F$ or its ring of algebraic integers $O_F$, then for any integer $i \geq 2$,

$$K_i(R) \otimes \mathbb{Q} \simeq \begin{cases} 0 & \text{if } i \text{ is even,} \\ \mathbb{Q}^{r_1+r_2} & \text{if } i \equiv 1 \text{ mod } 4, \\ \mathbb{Q}^{r_2} & \text{if } i \equiv 3 \text{ mod } 4. \end{cases}$$

6.1.3. Remarks.

(i) As consequences of 6.1.2, we have
(a) For all $n \geq 1$, the $K_{2n}(O_F)$ are finite groups and

$$\text{rank } K_{2n-1}(O_F) = \begin{cases} r_1 + r_2 & \text{if } n \text{ is odd}, \\ r_2 & \text{if } n \text{ is even}. \end{cases}$$

(b) For all $n \geq 1$, $K_{2n-1}(F)$ is finitely generated and $K_{2n}(F)$ is torsion.

(ii) Let $R$ be a Dedekind domain with quotient field $F$ and finite residue field $R/p$ for each prime ideal $p$ of $R$. Then it follows from the localization sequence 4.4.3(ii) that we have an exact sequence (for all $n \geq 1$)

$$0 \to K_{2n}(R) \to K_{2n}(F) \to \bigoplus_p K_{2n-1}(R/p) \to K_{2n-1}(R) \to K_{2n-1}(F) \to 0.$$  \hfill (I)

The next result, due to C. Soulé shows that we do get short exact sequences out of the sequence (I).

6.1.4. THEOREM [129]. Let $O_F$ be the ring of integers in a number field $F$. Then

(i) $K_{2n-1}(O_F) \approx K_{2n-1}(F)$ for all $n \geq 1$.

(ii) For all $n \geq 1$, $K_{2n}(F)$ is an infinite torsion group and we have an exact sequence

$$0 \to K_{2n}(O_F) \to K_{2n}(F) \to \bigoplus_p K_{2n-1}(O_F/p),$$

where $p$ runs through the prime ideals of $O_F$.

6.1.5. EXAMPLES.

(i) If $F = \mathbb{Q}$, $O_F = \mathbb{Z}$ then $K_3(\mathbb{Z}) \simeq K_3(\mathbb{Q}) \simeq \mathbb{Z}/48$. That $K_3(\mathbb{Z}) \simeq \mathbb{Z}/48$ is a result of Lee and Szcarba, see [89].

(ii) $F = \mathbb{Q}(i)$, then $K_3(Q(i)) \simeq \mathbb{Z} \oplus \mathbb{Z}/24$.

6.1.6. Let $F$ be a global field of characteristic $p$, i.e. $F$ is finitely generated of finite transcendence degree over $\overline{F}_q$. It is well known (see [128]) that there is a unique smooth projective curve $X$ over $\overline{F}_q$ whose function field is $F$, i.e. $F = \overline{F}_q(X)$. If $S$ is a non-empty set of closed points of $X$, then $X \setminus S$ is affine and we shall refer to the coordinate ring of $X \setminus S$ as the ring of $S$-integers in $F$.

6.1.7. THEOREM [166]. Let $X$ be a smooth projective curve over a finite field of characteristic $p$, then

(i) The groups $K_n(X)$ are finite groups of order prime to $p$.

(ii) If $R$ is the ring of $S$-integers in $F = \overline{F}_q(X)$ ($S \neq \emptyset$ as in 6.1.6) then

(a) $K_1(R) \cong R^* \cong \mathbb{F}_q^* \times \mathbb{Z}^1$, $|S| = s + 1$.

(b) For $n \geq 2$, $K_n(R)$ is a finite group of order prime to $p$. 
For further information on $K_n(X)$, see [166].

Next, we provide some general results on the higher $K$-theory of integers in local fields.

6.1.8. **Theorem** [166]. Let $E$ be a local field with discrete valuation ring $A$ and residue class field $\mathbb{F}_q$ of characteristic $p$. Assume that $\text{char}(E) = p$, then $E = \mathbb{F}_q[[\pi]]$ where $\pi$ is a uniformizing parameter of $A$ and we have for $n \geq 2$

(i) $K_n(E) \cong K_n(A) \oplus K_{n-1}(\mathbb{F}_q)$,

(ii) $K_n(A) \cong K_n(\mathbb{F}_q) \oplus U_n$ where the $U_n$ are uncountable uniquely divisible Abelian groups.

**Remarks.** 6.1.8 is a generalization of Moore’s theorem which states that $K_2(E) \cong \mu(E) \times U_2$ and $K_2(A) \cong \mu_p \times U_2$ (see [166]).

6.2. Étale and motivic Chern characters

6.2.1. A lot of the computations of $K$-theory of integers in number fields and $p$-adic fields have been through mapping $K$-theory into “étale cohomology via étale Chern characters” initially defined by C. Soulé, see [129]. We now briefly review this construction with the observation that since Soulé’s definition, there have been other approaches, e.g., maps from étale $K$-theory to étale cohomology due to Dwyer and Friedlander, [26], and “anti-Chern” characters, i.e. maps from étale cohomology to $K$-theory due to B. Kahn, [55]. Moreover the various interconnections already outlined in Section 5 between $K$-theory, Galois, étale and motivic cohomologies have also made computations of the $K$-theory of $\mathbb{Z}$ more accessible.

6.2.2. Let $X$ be an $H$-space, $M^n_m$ the $n$-dimensional mod-$m$ Moore space. We write $\pi_n(X, \mathbb{Z}/m)$ for $\pi_n(X, \mathbb{Z}/m)$ (see 3.1.4(viii)). Note then $\pi_n(X, \mathbb{Z}/m)$ is a group for $n \geq 2$.

Let $h_n : \pi_n(X, \mathbb{Z}/m) \to H_n(X, \mathbb{Z}/m)$ be the mod-$m$ Hurewitz map defined by $\alpha \in [M^n_m, X] \mapsto \alpha_*(\varepsilon_n)$ where $\varepsilon_n$ is the generator of $\mathbb{Z}/m \cong H_n(S^n, \mathbb{Z}/m) \cong H_n(M^n_m, \mathbb{Z}/m)$ and $\alpha_*$ is the homomorphism $H_n(M^n_m, \mathbb{Z}/m) \to H_n(X, \mathbb{Z}/m) : \varepsilon_n \mapsto \alpha_*(\varepsilon_n)$.

Then, $h_n$ is a group homomorphism for $n \geq 3$. If $X = B\text{GL}(A)^+$, then we have a mod-$m$ Hurewitz map $K_n(A, \mathbb{Z}/m) \to H_n(\text{GL}(A), \mathbb{Z}/m)$ (I) which is a group homomorphism.

6.2.3. Soulé’s étale Chern character. Let $l$ be a rational prime, $A$ a commutative ring with identity such that $1/l \in A$. Let $G$ be a group acting on $\text{Spec}(A)$ and $\rho : G \to \text{GL}_d(A)$ ($d$ an integer $> 1$) a representation of $G$, $c_i(\rho) \in H^2_{\text{et}}(A, G, \mu^{\otimes i})$ where $H^2_{\text{et}}(A, G, \mu^{\otimes i})$ are étale cohomology of $G$-sheaves on the étale site of $\text{Spec}(A)$. (See [101] and 5.3.3.) Assume that $G$ acts trivially on $\text{Spec}(A)$ then there exists a homomorphism

$$\phi : H^2_{\text{et}}(A, G, \mu^{\otimes i}) \to \prod_{k=0}^{2i} \text{Hom}(H^{2i-k}_{\text{et}}(G, \mathbb{Z}/(l^k)), H^k_{\text{et}}(A, \mu^{\otimes i}))$$
Higher algebraic $K$-theory

mapping $c_i(\rho)$ to $c_{ik}(\rho): H_{2i-k}(G, \mathbb{Z}/\ell^v) \to H_{k}^k(A, \mathbb{Z}/\ell^v)$. Put $G = \text{GL}_n(A)$, we obtain a map

$$c_{ik}(\text{id}): H^k_n(\text{GL}(A), \mathbb{Z}/\ell^v) \to H^k_{\text{et}}(A, \mathbb{Z}/\ell^v),$$

where $n + k = 2i$, $n \geq 2$.

Composing (I) in 6.2.2 with (II) above, yields the Soulé Chern characters

$$c_{ik}: K_n(A, \mathbb{Z}/\ell^v) \to H^k_{\text{et}}(A, \mathbb{Z}/\ell^v), \quad n + k = 2i, \ n \geq 2.$$  

(III)

6.2.4. Recall that if $X$ is an $H$-space, then the quotient map $M^n_{\ell^v} \to S^n$ (see 3.1.4(ix)) induces a map

$$\left[ S^n, X \right] \to \left[ M^n_{\ell^v}, X \right], \quad \text{i.e.} \quad \pi_n(X) \to \pi_n \left( X, \mathbb{Z}/\ell^v \right),$$

where $n + k = 2i$. If $X = B\text{GL}(A)^+$, $A$ as in 6.2.3, then we have a homomorphism $K_n(A) \to K_n(A, \mathbb{Z}/\ell^v)$, and by composing this map with (III) of 6.2.3, we obtain a homomorphism

$$c_{ik}: K_n(A) \to K_n(A, \mathbb{Z}/\ell^v) \to H^k_{\text{et}}(A, \mathbb{Z}/\ell^v), \quad n + k = 2i.$$  

(V)

6.2.5. Remarks/Examples.

(i) Let $A$ be a Dedekind domain with field of fractions $F$, $\ell$ a prime. Assume that $A$ contains $1/\ell$. Then we have localization sequences

$$\cdots \to K_n(A, \mathbb{Z}/\ell^v) \to K_n(F, \mathbb{Z}/\ell^v) \to \bigoplus_v K_{n-1}(k_v, \mathbb{Z}/\ell^v)$$

$$\to K_{n-1}(A, \mathbb{Z}/\ell^v) \to \cdots$$

and

$$0 \to H^1_{\text{et}}(A, \mathbb{Z}/\ell^v) \to H^1_{\text{et}}(F, \mathbb{Z}/\ell^v) \to \bigoplus_v H^0_{\text{et}}(k_v, \mathbb{Z}/\ell^{v-1})$$

$$\to H^2_{\text{et}}(A, \mathbb{Z}/\ell^v) \to \cdots.$$  

(ii) If in (i) $A$ contains $1/\ell$ and $\xi_m$, $m = \ell^v, \ m \neq 2$ (4), and if all residue fields $k_v$ are finite, then there is a map of localization sequences.
(iii) If $A$ is the ring of integers in a $p$-adic field $F$ with residue field $k$, and $1/\ell \in A$ ($\ell$ a rational prime, $\ell \neq p$), then $K_n(A, \mathbb{Z}/\ell^v) \simeq K_n(k, \mathbb{Z}/\ell^v)$ and $H^i_{et}(A, \mu^i_{\mathbb{Q}}) \simeq H^i_{et}(k, \mu^i_{\mathbb{Q}})$ and $c_{ij}$ can be identified with multiplication by $\pm (i-1)!$ (see [166]).

6.2.6. Let $F$ be a global field with ring of integers $O_F$, $\ell$ a prime, $\ell \neq \text{char}(F)$, $O'_F = O_F(1/\ell)$, then by tensoring (V) ($A = O'_F$) with $\mathbb{Z}_\ell$ we obtain map

$$
\text{ch}_{n,k} : K_{2n-k}(O_F) \otimes \mathbb{Z}_\ell \to H^k_{et}(O'_F, \mathbb{Z}_\ell(n))
$$

6.2.7. QUILLEN–LICHTENBAUM (Q–L) CONJECTURE. Let $F$ be a global field with ring of integers $O_F$, $\ell$ a prime, $\ell \neq \text{char}(F)$, $O'_F = O_F(1/\ell)$. Then the Chern characters

$$
\text{ch}^\ell_{i,n} : K_{2n-k}(O_F) \otimes \mathbb{Z}_\ell \to H^k_{et}(O'_F, \mathbb{Z}_\ell(n))
$$

are isomorphisms for $n \geq 2$, $i = 1, 2$ unless $\ell = 2$ and $F$ is a number field with a real embedding.

The following result 6.2.8 establishes the Q–L conjecture in a special case $k \leq 2(n-1)$.

6.2.8. THEOREM. In the notation of 6.2.6, when $k = 1$ or 2, the mapping

$$
\text{ch}_{n,k} : K_{2n-k}(O_F) \otimes \mathbb{Z}_\ell \to H^k_{et}(O'_F, \mathbb{Z}_\ell(n))
$$

satisfies

(i) If $2n - k \geq 2$ and $\ell \neq 2$, then $\text{ch}_{n,k}$ is split surjective.

(ii) If $2n - k = 2$, $2n - k = 3$, $\ell \neq 2$ or when $\sqrt{-1} \in F$, $2n - k \geq 2$ and $\ell = 2$, then $\text{ch}_{n,k}$ is an isomorphism.

6.2.9. DEFINITION. Let $\ell$ be a fixed rational prime. For any field $F$, define integers $\omega^{(i)}_\ell (F)$ by $\omega^{(i)}_\ell (F) := \max \{ \ell^v \mid \text{Gal}(F(\mu_{\ell^v})/F) \text{ has exponent dividing } i \}$.
If there is no maximum \( v \), put \( \omega_i^{(\ell)}(F) = \ell^\infty \). Suppose that \( \omega_i^{(\ell)}(F) = 1 \) for almost all \( \ell \) and that \( \omega_i^{(\ell)}(F) \) is finite otherwise, write \( \omega_i(F) = \prod \omega_i^{(\ell)}(F) \).

Note that if \( F(\mu_\ell) \) has only finitely many \( \ell \)-primary roots of unity for all primes \( \ell \) and \( [F(\mu_\ell) : F] \to \infty \) as \( \ell \to \infty \), then the \( \omega_i(F) \) are finite for all \( i \neq 0 \). This is true for all local and global fields. See [166].

Also note that \( \omega_i(F) = |H^0(F, \mathbb{Q}/\mathbb{Z}(n))| \).

### 6.2.10. Examples.

(i) \( \omega_i(\mathbb{F}_p) = q^i - 1 \) \( \forall \) positive integers \( i \). Hence \( K_{2n-1}(\mathbb{F}_p) = q^i - 1 \) unless \( (p - 1)|i \).

(ii) \( \omega_i(\overline{\mathbb{Q}}_p) = q^i - 1 \) unless \( (p - 1)|i \). If \( i = (p - 1)p^b m, (p \nmid m) \), then \( \omega_i(\overline{\mathbb{Q}}_p) = (q^i - 1)p^{1+b} \).

For further information and results on the \( \omega_i(F) \), see [166].

### 6.2.11. Let \( F \) be a number field. Then each real embedding \( \delta_i : F \to \mathbb{R} \) induces a map \( F^* \to \mathbb{R}^* \to \mathbb{Z}/2 \) which detects the sign of \( F^* \) under \( \delta_i \). The sign map \( \delta : F^* \to (\mathbb{Z}/2)^{r_1} \) is the sum of the \( \delta_i \) and \( \delta \) is surjective. \( \ker \delta \) is the group of totally positive units in \( F \) and denoted by \( F^*_+ \).

Now let \( S \) be a finite set of places of \( F \), \( O_S \) the ring of \( S \)-integers of \( F \). Then the kernel of \( \delta|_{O_S} : O_S \to F^* \to (\mathbb{Z}/2)^{r_1} \) is a subgroup of \( O_S^* \) called the subgroup of totally positive units in \( O_S \) and denoted by \( O_S^*_+ \). The sign map \( \delta|_{O_S} \) factors through \( F^*/2 = H^1(F, \mathbb{Z}/2) \) and hence factors through \( \delta^1 : H^1(O_S, \mathbb{Z}/2) \to (\mathbb{Z}/2)^{r_1} \). The signature defect \( j(O_S) \) of \( O_F \) is defined as the dimension of the cokernel of \( \delta^1 \). Note that \( j(O_S) \leq j(O_F) \) and that \( j(F) = 0 \) if \( 0 \leq j(O_S) < r_1 \).

The narrow Picard group, \( \text{Pic}_+(O_S) \) is the cokernel of the restricted divisor map \( F^* \to \bigoplus_{\ell \in \mathbb{Z}} \mathbb{Z} \).

### 6.2.12. Theorem [118]. Let \( \ell \) be an odd prime, \( F \) a number field, \( O_S \) a ring of \( S \)-integers in \( F \) containing \( \frac{1}{2} \), \( j \) the signature defect of \( O_S \). Denote \( \omega_i^{(2)}(F) \) by \( \omega_i \). Then, there exists an integer \( m \), \( j \leq m < r \), such that for all \( n \geq 2 \), the 2-primary subgroup \( K_n(O_S)(2) \) of \( K_n(O_S) \) is given by

\[
K_n(O_S)(2) \cong \begin{cases} 
H_2^2(O_S, \mathbb{Z}_2(4a + 1)) & \text{for } n = 8a, \\
\mathbb{Z}/2 & \text{for } n = 8a + 1, \\
H_2^2(O_S, \mathbb{Z}_2(4a + 2)) & \text{for } n = 8a + 2, \\
(\mathbb{Z}/2)^{r_1-1} \oplus \mathbb{Z}/2^a_{4b+2} & \text{for } n = 8a + 3, \\
(\mathbb{Z}/2)^m \times H_2^2(O_S, \mathbb{Z}_2(4a + 3)) & \text{for } n = 8a + 4, \\
0 & \text{for } n = 8a + 6, \\
H_2^2(O_S, \mathbb{Z}_2(4a + 4)) & \text{for } n = 8a + 6, \\
\mathbb{Z}/(2a_{4a+4}) & \text{for } n = 8a + 7.
\end{cases}
\]

The next result is on the odd prime torsion subgroup of \( K_n(O_S) \).
6.2.13. **Theorem [166]**. Let \( \ell \) be an odd prime, \( F \) a number field, \( O_S \) a ring of \( S \)-integers of \( F \), \( O'_S = O_S(\frac{1}{\ell}) \). Then for all \( n \geq 2 \), we have

\[
K_n(O_S)(\ell) \cong \begin{cases} 
H^2_{\ell}(O'_S, \mathbb{Z}_\ell(i+1)) & \text{for } n = 2i > 0, \\
\mathbb{Z}_\ell^r \oplus \mathbb{Z}/\omega_i(F) & \text{for } n = 2i - 1, \ i \text{ even}, \\
\mathbb{Z}_\ell^{r_1+r_2} \oplus \mathbb{Z}/\omega_i^{(\ell)}(F) & \text{for } n = 2i - 1, \ i \text{ odd}.
\end{cases}
\]

The following result gives a picture for \( K_n(O_S) \) when \( F \) is totally imaginary.

6.2.14. **Theorem [166]**. In the notation of 6.2.13, let \( F \) be totally imaginary. Then for all \( n \geq 2 \)

\[
K_n(O_S) \cong \begin{cases} 
\mathbb{Z} \oplus \text{Pic}(O_S) & \text{for } n = 0, \\
\mathbb{Z}^{n_2 + |S| - 1} \oplus \mathbb{Z}/\omega_i & \text{for } n = 1, \\
\bigoplus_i H^2_{\ell}(O'_S, \mathbb{Z}_\ell(i+1)) & \text{for } n = 2i \geq 2, \\
\mathbb{Z}_\ell^r \oplus \mathbb{Z}/\omega_i(F) & \text{for } n = 2i - 1 \geq 3.
\end{cases}
\]

The next result which gives the picture of \( K_n(O_F) \) for each odd \( n \geq 3 \) is a consequence of 6.2.13 and 6.2.14 and some other results which can be found in [166]. This reference also contains details of the proof.

6.2.15. **Theorem [166]**. Let \( F \) be a number field, \( O_S \) a ring of \( S \)-integers. Then

(i) \( K_n(O_S) \cong K_n(F) \) for each odd \( n \geq 3 \).

(ii) If \( F \) is totally imaginary then

\[
K_n(F) \cong \mathbb{Z}^{r_2} \oplus \mathbb{Z}/\omega_i(F).
\]

(iii) If \( i = \frac{n+1}{2} \), then

\[
K_n(O_S) \simeq K_n(F) \cong \begin{cases} 
\mathbb{Z}^{r_1+r_2} \oplus \mathbb{Z}/\omega_i(F), & n \equiv 1 \pmod{8}, \\
\mathbb{Z}^{r_1} \oplus \mathbb{Z}/2\omega_i(F) \oplus (\mathbb{Z}/2)^{r_1-1}, & n \equiv 3 \pmod{8}, \\
\mathbb{Z}^{r_1+r_2} \oplus \mathbb{Z}/2\omega_i(F), & n \equiv 5 \pmod{8}, \\
\mathbb{Z}^{r_2} \oplus \mathbb{Z}/\omega_i(F), & n \equiv 7 \pmod{8}.
\end{cases}
\]

6.2.16. It follows from 6.2.15 that \( K_n(Q) \cong \mathbb{Z} \) for all \( n \equiv 5 \pmod{8} \) (since \( w_1(Q) = 2 \)); more generally, if \( F \) has a real embedding, and \( n \equiv 5 \pmod{8} \) then \( K_n(F) \) has no 2-primary torsion. (Since \( \frac{1}{2}\omega_i(F) \) is an odd integer.)

6.2.17. **Theorem [166]**. Let \( E \) be a \( p \)-adic field of degree \( d \) over \( \widehat{\mathbb{Q}}_p \), with ring of integers \( A \). Then for all \( n \geq 2 \), we have

(i) \( K_n(A, \widehat{\mathbb{Q}}_p) \cong K_n(E, \widehat{\mathbb{Q}}_p) \cong \begin{cases} 
\mathbb{Z}/\omega_i^{(p)}(E), & n = 2i, \\
(\widehat{\mathbb{Q}}_p)^d \oplus \mathbb{Z}/\omega_i^{(p)}(E), & n = 2i - 1.
\end{cases} \)
Higher algebraic \( K \)-theory

(ii) \( K_{2i-1}(E, \mathbb{Z}/p^v) \cong H^1(E, \mu^{p^{2i}}_{p^v}) \) for all \( i \) and \( v \) and for all \( i \) and \( \ell^v > \omega_1 \), \( H^1(E, \mu^{p^{2i}}_{p^v}) \cong (\mathbb{Z}/p^v)^d \oplus \mathbb{Z}/\omega_{i-1} \).

(iii) \( K_n(E, \hat{\mathbb{Z}}_p) \) are finitely generated \( \hat{\mathbb{Z}}_p \)-modules.

6.2.18. Remarks.

(i) Pushin, [110], has constructed motivic Chern characters \( ch^M_{i,n}: K_{2n-i}(F) \to H^i_M(F, \mathbb{Z}(n)) \) for \( n \geq 2, \ i = 1, 2 \), which induce étale Chern characters after tensoring by \( \mathbb{Z}_l \).

(ii) For all \( n \geq 2 \), there exist isomorphisms \( K_{2n-2}(F) \cong H^2_M(F, \mathbb{Z}(n)) \) and \( K_{2n-1}(F) \cong H^1_M(F, \mathbb{Z}(n)) \) up to 2-torsion, see [166].

(iii) In view of 5.4.4(v), we have up to 2-torsion isomorphisms \( K_{2n-2}(OF) \cong H^2_M(OF, \mathbb{Z}(n)) \) and \( K_{2n-1}(OF) \cong H^1_M(OF, \mathbb{Z}(n)) \).

6.3. Higher \( K \)-theory and zeta functions

6.3.1. Definition. Let \( s \) be a complex number with \( \text{Re}(s) > 1 \). The Riemann zeta function is defined as the convergent series \( \zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} \).

Note that \( \zeta(s) \) admits a meromorphic continuation to the whole complex plane with a simple pole at \( s = 1 \) and no pole anywhere else.

For connections with \( K \)-theory we shall be interested in the values of \( \zeta(s) \) at negative integers.

6.3.2. The Bernoulli numbers \( B_n, n \geq 0 \), are the rational numbers which arise in the power series expansion of \( \frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!} \).

6.3.3. Theorem.

(i) For \( n \) odd \( \geq 3 \), \( \zeta(s) \) has a simple zero at \( s = 1 - n \). For \( n \) even \( > 0 \), \( \zeta(1-n) \neq 0 \).

(ii) For \( n \geq 1 \), \( \zeta(1-n) = (-1)^{n-1} \frac{B_n}{n} \).

6.3.4. Let \( F \) be a number field with ring of integers \( OF, a \), an ideal of \( OF, N(a) = |OF/a| \).

Then the Dedekind generalization of Riemann zeta function is defined by \( \xi_F(s) := \sum_{0 \neq a \subset OF} \frac{1}{N(a)^s} \). Note that \( \xi_F(s) \) is convergent for \( \text{Re}(s) > 1 \) and can be extended to a meromorphic function on \( \mathbb{C} \) with a single pole at \( s = 1 \).

6.3.5. Theorem. Let \( F \) be a number field and let \( r_1 = \text{number of real embeddings of } F, r_2 = \text{number of pairs of complex embeddings of } F \). Let \( d_n \) be the order of vanishing of \( \zeta(s) \) at \( s = 1 - n \).

Then

\[
d_n = \begin{cases} 
  r_1 + r_2 - 1 & \text{if } n = 1, \\
  r_1 + r_2 & \text{if } n \geq 3 \text{ is odd}, \\
  r_2 & \text{if } n \text{ is even}. 
\end{cases} \]
In view of 6.3.5, we now have a reformulation of 6.1.3(i)(a) in the following.

**6.3.6. Theorem.** Let $F$ be a number field and $O_F$ the ring of integers of $F$. Then for all $n \geq 2$

$$K_{2n-1}(O_F) \simeq \mathbb{Z}^{dn} \oplus \text{finite},$$

i.e. the rank of $K_{2n-1}(O_F)$ is the order of vanishing of $\xi_F$ at $s = 1 - n$.

**6.3.7. Definition.** Define $\xi^*(-2m)$ as the first non-vanishing coefficient in the Taylor expansion around $s = -2m$ and call this the special value of the zeta function at $s = -2m$.

**6.3.8.** Let $F$ be a number field. Recall from [14] that there exist higher regulator maps $X^B_n(F) : K_{2n-1}(O_F) \to \mathbb{R}^{d_n}$ with finite kernel and image a lattice of rank $d_n$. The covolume of the lattice is called the Borel regulator and denoted by $R^B_n(F)$. Borel proved that $\xi_F^*(1-n) = q_n \cdot R^B_n(F)$ where $q_n$ is a rational number.

**6.3.9. Lichtenbaum Conjecture [91,92].** For all $n \geq 2$

$$\xi^*_F(1-n) = \pm \frac{|K_{2n-2}(O_F)|}{|K_{2n-1}(O_F)_{\text{tors}}|} R^B_n(F)$$

up to powers of 2.

**6.3.10. Remarks.**

(i) Birch and Tate had earlier conjectured a special case of 6.3.10 above viz. for a totally real number field $F$, $\xi_F(-1) = \pm \frac{|K_2(O_F)|}{|O_F^2|}$.

(ii) The next result is due to A. Wiles, [170].

**6.3.11. Theorem [170].** Let $F$ be a totally real number field. If $\ell$ is odd, and $O'_F = O_F(\frac{1}{\ell})$, then for all even $n = 2i$

$$\xi_F(1-n) = \frac{|H^2_{\text{et}}(O'_F, \mathbb{Z}/(n))|}{|H^1_{\text{et}}(O'_F, \mathbb{Z}/(n))|} u_n,$$

where $u_n$ is a rational number prime to $\ell$.

**6.3.12. Theorem [166].** Let $F$ be a totally real number field. Then

$$\xi_F(1-2k) = (-1)^{kr} \frac{|K_{4k-2}(O_F)|}{|K_{4k-1}(O_F)|}$$

up to a factor of 2.
Higher algebraic $K$-theory

6.3.13. Theorem [166]. Let $F$ be a totally real number field, $O_F$ the ring of integers of $F$, $O^\prime_F = O_F(\frac{1}{\ell})$. Then for all even $i > 0$

$$2^i \left| \frac{K_{2i-2}(O_F)}{K_{2i-1}(O_F)} \right| = \prod_\ell \left(\frac{H^2_{\text{et}}(O^\prime_F, F_{2\ell}(\ell))}{H^2_{\text{et}}(O^\prime_F, \mathbb{Z}_2(0))} \right).$$

6.3.14. Definition. Let $F$ be a global field of characteristic $p$ and $X$ the associated smooth projective curve over $\mathbb{F}_q$, i.e. $F = \mathbb{F}_q(X)$, $O_F$ the integral closure of $\mathbb{F}_q[t]$ in $F$. The maximal ideals of $O_F$ are the finite primes in $F$ and the associated zeta function is defined by

$$\zeta_F(s) = \prod_{p \text{ finite}} \frac{1}{1 - N(p)^{-s}}.$$

6.3.15. Theorem. Let $F$ be a global field of characteristic $p > 0$. Then for all $n \geq 2$, we have

$$\zeta_F(1 - n) = \pm \frac{H^2(O_F, \mathbb{Z}(n))}{\omega_n(F)}.$$

6.4. Higher $K$-theory of $\mathbb{Z}$

In this subsection, we briefly survey the current situation with the computations of $K_n(\mathbb{Z})$, $n \geq 0$. Again, details of arguments leading to the computations can be found in [166]. First we list specific information for $K_n(\mathbb{Z})$, $n = 0, 1, 2, 3, 4, 5, 6, 7, 9, 10$.

6.4.1. Theorem. $K_0(\mathbb{Z}) \simeq \mathbb{Z}$, $K_1(\mathbb{Z}) \simeq \mathbb{Z}/2$, $K_2(\mathbb{Z}) \simeq \mathbb{Z}/2$, $K_3(\mathbb{Z}) \simeq \mathbb{Z}/48$, see [89]. $K_4(\mathbb{Z}) = 0$, see [117], $K_5(\mathbb{Z}) = \mathbb{Z}$, see [166], $K_6(\mathbb{Z}) = 0$, see [166], $K_7(\mathbb{Z}) = \mathbb{Z}/240$, see [166], $K_9(\mathbb{Z}) = \mathbb{Z} \oplus \mathbb{Z}/2$, $K_{10}(\mathbb{Z}) = \mathbb{Z}/2$, [166].

More generally we have.

6.4.2. Theorem.

$$K_n(\mathbb{Z}) = \begin{cases} \text{finite} & \forall n > 0, \quad n \not\equiv 1 \pmod{4}, \\ \mathbb{Z} + \text{finite} & \text{if } n \equiv 1 \pmod{4}. \end{cases}$$

Using 6.2.10 and 6.2.12, one can prove.

6.4.3. Theorem [166]. Let $\ell$ be an odd rational prime

(i) $K_{2n-1}(\mathbb{Z})$ has $\ell$-torsion exactly when $n \equiv 0 \pmod{(\ell - 1)}$.
(ii) The $\ell$-primary subgroups of $K_{2n}(\mathbb{Z})$ are $H^2_{\text{et}}(\mathbb{Z}(\frac{1}{\ell}); \hat{\mathbb{Z}}_{\ell}(i + 1)).$
The following result is a consequence of 6.3.13 and shows that the Lichtenbaum conjecture holds up to a factor of 2.

6.4.4. Theorem [166]. For all $n \geq 1$, we have

\[
\frac{|K_{4n-2}(\mathbb{Z})|}{|K_{4n-1}(\mathbb{Z})|} = \frac{b_n}{4n} = \frac{-1^n}{2} \xi(1 - 2n).
\]

Hence if $c_n$ denotes the numerator of $\frac{b_n}{4n}$, then $|K_{4n-2}| = \begin{cases} c_n, & \text{n even,} \\ 2c_n, & \text{n odd.} \end{cases}$

6.4.5. Definition. A prime $p$ is said to be regular if Pic$(\mathbb{Z}[\mu_p])$ has no element of exponent $p$, i.e., $p$ does not divide the order $h_p$ of Pic$(\mathbb{Z}[\mu_p])$. Note that $p$ is regular iff Pic$(\mathbb{Z}[\mu_p\nu])$ has no $p$-torsion for all $\nu$.

6.4.6. Theorem [166]. If $\ell$ is an odd regular prime, then $K_{2n}(\mathbb{Z})$ has no $\ell$-torsion.

The following result shows that the 2-primary subgroups of $K_n(\mathbb{Z})$ are essentially periodic of period 8.

6.4.7. Theorem [166].

\[
K_n(\mathbb{Z})(2) = \begin{cases} \mathbb{Z}/2 & \text{if } n \equiv 1 \pmod{8}, \\ \mathbb{Z}/2 & \text{if } n \equiv 2 \pmod{8}, \\ \mathbb{Z}/16 & \text{if } n \equiv 3 \pmod{8}, \\ 0 & \text{if } n \equiv 4 \pmod{8}, \\ 0 & \text{if } n \equiv 5 \pmod{8}, \\ 0 & \text{if } n \equiv 6 \pmod{8}, \\ \mathbb{Z}/6a & \text{if } n \equiv 7 \pmod{8}, \\ 0 & \text{if } n \equiv 8 \pmod{8}. \end{cases}
\]

6.4.8. Vandiver’s Conjecture. If $\ell$ is an irregular prime, then Pic$(\mathbb{Z}[\xi_{\ell} + \xi_{\ell}^{-1}])$ has no $\ell$-torsion.

This conjecture is equivalent to the expression of the natural representation of $G = \text{Gal}(\mathbb{Q}(\xi_{\ell})/\mathbb{Q})$ on Pic$(\mathbb{Z}[\xi_{\ell}])/\ell$ as a sum of $G$-modules $\mu_{\ell}^\otimes_i$ when $i$ is odd.

Note that Vandiver’s conjecture for $\ell$ is equivalent to the assertion that $K_{4n}(\mathbb{Z})$ has no $\ell$-torsion for all $n < \frac{\ell^2}{2}$, see [166].

The following result is due independently to S. Mitchel and M. Kurihara.
6.4.9. Theorem [84]. If Vandiver’s conjecture holds, then

\[ K_n(\mathbb{Z}) = \begin{cases} 
\mathbb{Z} \oplus \mathbb{Z}/2 & \text{if } n \equiv 1 \pmod{8}, \\
\mathbb{Z}/2c_k & \text{if } n \equiv 2 \pmod{8}, \\
\mathbb{Z}/2\omega_{2k} & \text{if } n \equiv 3 \pmod{8}, \\
0 & \text{if } n \equiv 4 \pmod{8}, \\
\mathbb{Z} & \text{if } n \equiv 5 \pmod{8}, \\
\mathbb{Z}/c_k & \text{if } n \equiv 6 \pmod{8}, \\
\mathbb{Z}/\omega_{2k} & \text{if } n \equiv 7 \pmod{8}, \\
0 & \text{if } n \equiv 8 \pmod{8}. 
\end{cases} \]

7. Higher \(K\)-theory of orders, group-rings and modules over ‘EI’-categories

7.1. Higher \(K\)-theory of orders and group-rings

7.1.1. We recall that if \(R\) is a Dedekind domain with quotient field \(F\), and \(\Lambda\) is any \(R\)-order in a semi-simple \(F\)-algebra \(\Sigma\), then \(SK_n(\Lambda) := \ker(K_n(\Lambda) \to K_n(\Sigma))\) and \(SG_n(\Lambda) := \ker(G_n(\Lambda) \to G_n(\Sigma))\). Also, any \(R\)-order in a semi-simple \(F\)-algebra \(\Sigma\) can be embedded in a maximal \(R\)-order \(\Gamma\) which has well-understood arithmetic properties relative to \(\Sigma\). More precisely if \(\Sigma = \prod_{i=1}^{r} M_{n_i}(D_i)\) then \(\Gamma\) is Morita equivalent to \(\prod_{i=1}^{r} M_{n_i}(\Gamma_i)\) where the \(\Gamma_i\) are maximal orders in the division algebra \(D_i\) and so \(K_n(\Gamma) \approx \bigoplus K_n(\Gamma_i)\) while \(K_n(\Sigma) \approx \bigoplus K_n(D_i)\). So, the study of \(K\)-theory of maximal orders in semi-simple algebras can be reduced to the \(K\)-theory of maximal orders in division algebras.

As remarked in [79] the study of \(SK_n(\Lambda)\) facilitates the understanding of \(K_n(\Lambda)\) apart from the various known topological applications known for \(n = 0, 1, 2\), where \(\Lambda = \mathbb{Z}G\) for some groups \(G\) that are usually fundamental groups of some spaces (see [163,150,103]). Also \(SK_n(\Lambda)\) is connected to the definition of higher class groups which generalizes to higher \(K\)-groups the notion of class groups of orders and group-rings. (See [48,65].)

First, we have the following result 7.1.2 on \(SK_n\) of maximal orders in \(p\)-adic semi-simple algebras. This result as well as the succeeding ones, 7.1.4 to 7.1.6, are due to Kuku.

7.1.2. Theorem [68]. Let \(R\) be the ring of integers in a \(p\)-adic field \(F\), \(\Gamma\) a maximal \(R\)-order in a semi-simple \(F\)-algebra \(\Sigma\). Then for all \(n \geq 1\),

(i) \(SK_{2n}(\Gamma) = 0\).

(ii) \(SK_{2n-1}(\Gamma) = 0\) if and only if \(\Sigma\) is unramified over its centre (i.e. \(\Sigma\) is a direct product of matrix algebras over fields).

7.1.3. Remark.

(i) The result above is a generalization to higher \(K\)-groups of an earlier result for \(SK_1\).

(ii) Before discussing arbitrary orders, we record the following consequences of 7.1.2

(a) 7.1.2 holds for group rings \(\Gamma = RG\) where \(G\) is a finite group of order relatively prime to \(p\) and \(FG\) splits, see [19].

(b) If \(m = \text{rad} \Gamma\), then for all \(n \geq 1\), the transfer map: \(K_{2n-1}(\Gamma/m) \to K_{2n-1}(\Gamma)\) is non-zero unless \(\Sigma\) is a product of matrix algebras over fields. Hence, the Gersten conjecture does not hold in the non-commutative case.
(c) For all $n > 0$, we have an exact sequence $0 \to K_{2n}(\Gamma) \to K_{2n-1}(\Sigma) \to K_{2n-1}(\Gamma/m) \to 0$ if and only if $\Sigma$ is a direct product of matrix algebras over fields.

We now discuss arbitrary orders and the global situations.

7.1.4. **Theorem** [74]. Let $R$ be the ring of integers in a number field $F$, $\Lambda$ any $R$-order in a semi-simple $F$-algebra $\Sigma$. Then for all $n \geq 1$
(i) $K_n(\Lambda)$ is a finitely generated Abelian group.
(ii) $SK_n(\Lambda)$ is a finite group.
(iii) If $\hat{\Lambda}_p$ denotes the completion of $\Lambda$ at a prime $p$ of $R$, then $SK_n(\hat{\Lambda}_p)$ is a finite group.

7.1.5. **Remarks**.
(i) The above results 7.1.4 hold for $\Lambda = RG, \hat{\Lambda}_p = \hat{R}_pG$ where $G$ is a finite group and $\hat{R}_p$ is the completion of $R$ at a prime $p$ of $R$.
(ii) 7.1.4(i) was proved in [73], Theorem 2.1.1, 7.1.4 (ii) and (iii) were proved for group-rings $RG, \hat{R}_pG$ in [73], Theorem 3.2 and later for $R$-orders in [74], Theorem 1.1 (ii) and (iii).

We next discuss similar finiteness results for $G$-theory.

7.1.6. **Theorem** [69,74]. Let $R$ be the ring of integers in a number field $F$, $\Lambda$ any $R$-order in a semi-simple $F$-algebra $\Sigma$, $p$ any prime ideal of $R$. Then for all $n \geq 1$
(i) $G_n(\Lambda)$ is a finitely generated Abelian group.
(ii) $G_{2n-1}(\Lambda_p)$ is a finitely generated Abelian group.
(iii) $SG_{2n}(\Lambda) = SG_{2n}(\Lambda_p) = SG_{2n}(\hat{\Lambda}_p) = 0$.
(iv) $SG_{2n-1}(\Lambda)$ is a finite group.
(v) $SG_{2n-1}(\Lambda_p), SG_{2n-1}(\hat{\Lambda}_p)$ are finite groups of order relatively prime to the rational primes lying below $p$.

7.1.7. **Remarks**.
(i) The results above hold for $\Lambda = RG, \hat{\Lambda}_p = \hat{R}_pG$.
(ii) One can show from 7.1.6(i) that if $A$ is any $R$-algebra finitely generated as an $R$-module, then $G_n(A)$ is finitely generated (see [73, Theorem 2.3]).

We also have the following result on the vanishing of $SG_n(\Lambda)$ (see [87]) due to Laubenbacher and Webb.

7.1.8. **Theorem** [87]. Let $R$ be a Dedekind domain with quotient field $F$, $\Lambda$ any $R$-order in a semi-simple $F$-algebra. Assume that
(i) $SG_1(\Lambda) = 0$.
(ii) $G_n(\Lambda)$ is finitely generated for all $n \geq 1$.
(iii) $R/p$ is finite for all primes $p$ or $R$. 

Higher algebraic $K$-theory

(iv) If $\xi$ is an $\ell^s$-th root of unit for any rational prime $\ell$ and positive integer $s$, $R$ the integral closure of $R$ in $F(\xi)$, then $SG_1(\tilde{R} \otimes_R A) = 0$.

Then $SG_n(A) = 0$ for all $n \geq 1$.

7.1.9. REMARKS. It follows from 7.1.8 that if $R$ is the ring of integers in a number field $F$ then (i) $SG_n(RG) = 0$, (ii) $SG_n(\hat{A}_p) = 0$ for all prime ideals $p$ of $R$ and all $n \geq 1$ (see [78, 1.9]), if $\Lambda$ satisfies the hypotheses of 7.1.8.

The following extension of finiteness results for $G$-theory to group-rings of infinite groups is significant. This result is due to Kuku and Tang, see [82].

7.1.10. THEOREM [82]. Let $V = G \rtimes T$ be the semi-direct product of a finite group $G$ of order $p$ with an infinite cyclic group $T = \langle t \rangle$ with respect to the automorphism $\alpha : G \to G : g \mapsto tgt^{-1}$. Then for all $n \geq 0$, $G_n(RV)$ is a finitely generated Abelian group, where $R$ is the ring of integers in a number field.

We also have the following result on the ranks of $K_n(\Lambda)$, $G_n(\Lambda)$, due to Kuku, see [76].

7.1.11. THEOREM [76]. Let $R$ be the ring of integers in a number field $F$, $\Lambda$ an $R$-order in a semi-simple $F$-algebra $\Sigma$, $\Gamma$ a maximal $R$-order containing $\Lambda$.

Then for all $n \geq 2$

$$\text{rank } K_n(\Lambda) = \text{rank } G_n(\Lambda) = \text{rank } K_n(\Sigma) = \text{rank } K_n(\Gamma).$$

7.1.12. REMARKS.

(i) The above results 7.1.11 hold for $\Lambda = RG$ where $G$ is any finite group.

(ii) The ranks of $K_n(R)$ and $K_n(F)$ have been discussed in Section 6.

It then means that if $\Sigma$ is a direct product of matrix algebras over fields and $\Gamma$ is a maximal order in $\Gamma$, then $\text{rank } K_n(\Gamma) = \text{rank } K_n(\Sigma)$ is completely determined since $\Sigma = \prod M_{n_i}(F_i)$ and $\Gamma = \prod M_{n_i}(R_i)$ where $R_i$ is the ring of integers in $F_i$. Also by theorem 7.1.11 this is equal to $\text{rank } G_n(\Lambda)$ as well as $\text{rank } K_n(\Lambda)$ if $\Lambda$ is any $R$-order contained in $\Gamma$.

However, if $\Sigma$ does not split, there exists a Galois extension $E$ of $F$ which splits $\Sigma$, in which case we can reduce the problem to that of computation of the ranks of $K_n$ of fields.

We next obtain some results connecting $K_n$ and $G_n$ through the Cartan map. Recall that if $A$ is any Noetherian ring, the inclusions $P(A) \to M(A)$ induces a homomorphism $K_n(A) \to G_n(A)$ for all $n \geq 0$. First we have the following result, whose proof uses methods of equivariant higher $K$-theory, is due to Dress and Kuku (see [25]).

7.1.13. THEOREM [25]. Let $k$ be a field of characteristic $p$, $G$ a finite group. Then the inclusion functor $P(kG) \to M(kG)$ induces an isomorphism $\mathbb{Z}(\frac{1}{p}) \otimes K_n(kG) \cong \mathbb{Z}(\frac{1}{p}) \otimes G_n(kG)$ for all $n \geq 0$.

As a consequence of 7.1.13 one can prove the following result 7.1.14 and 7.1.15 due to Kuku, [73].
7.1.14. **Corollary [73]**. Let $G$ be a finite group, $k$ a finite field of characteristic $p \neq 0$. Then for all $n \geq 1$,

1. $K_{2n}(kG)$ is a finite $p$-group.
2. The Cartan map $K_{2n-1}(kG) \to G_{2n-1}(kG)$ is surjective and $\text{Ker} \varphi_{2n-1}$ is the Sylow-$p$-subgroup of $K_{2n-1}(kG)$.

Finally we have the following:

7.1.15. **Theorem [74]**. Let $R$ be the ring of integers in a number field $F$, $G$ a finite group, $p$ a prime ideal of $R$. Then for all $n \geq 1$.

1. The Cartan homomorphisms $K_n((R/p)G) \to G_n((R/p)G)$ are surjective.
2. The Cartan homomorphism $K_n(RG) \to G_n(RG)$ induces a surjection $SK_n(RG) \to SG_n(RG)$.
3. $SK_{2n}(RG) = \text{Ker}(K_{2n}(RG) \to G_{2n}(RG))$.

7.1.16. **Remarks**. Recall from 4.6.4 that if $R$ is the ring of integers in a number field $F$ and $A$ is any $R$-order in a semi-simple $F$-algebra $\Sigma$, then we have a long exact sequence

$$\cdots \to K_{n+1}(A, \omega) \to K_n(A) \to G_n(A) \to K_{n-1}(A, \omega) \to \cdots,$$

where $(A, \omega)$ is the Waldhausen category of bounded chain complexes over $\mathcal{M}(A)$ with weak equivalences as quasi-isomorphisms.

It follows from 7.1.4 and 7.1.6 that for all $n \geq 1$, $K_{n+1}(A, \omega)$ is a finitely generated Abelian group.

7.1.17. For the rest of this subsection, we shall focus briefly on the higher $G$-theory of Abelian group-rings and some ramifications of this theory to non-commutative cases, viz. dihedral and quaternion groups, nilpotent groups, and groups of square free order. The work reported here is due to D. Webb (see [159–162]).

The following results generalize to higher $G$-theory of results in 2.2.5 of [79]. The notations are those of 2.2.4, 2.2.5 of [28].

7.1.18. **Theorem [159]**. Let $\pi$ be an Abelian group and $R$ a Noetherian ring (not necessarily commutative). Then

$$G_n(R\pi) \simeq \bigoplus_{\rho \in X(\pi)} G_n(R(\rho)) \quad \text{for all } n \geq 0,$$

where $X(\pi)$ is the set of cyclic quotients of $\pi$.

Next we have the following ramifications of 2.5.20 of [28].

7.1.19. **Theorem [159]**. Let $G = \pi \rtimes \Gamma$ be the semidirect product of an Abelian group $\pi$ and a finite group $\Gamma$, $R$ a Noetherian ring. Assume that the $\Gamma$-action stabilizes every
cyclic subgroup of $\pi$ so that $\Gamma$ acts on each cyclic quotient $\rho$ of $\pi$. Let $R(\rho)\#\Gamma$ be the twisted group-ring. Then for all $n > 0$

$$G_n(RG) \simeq \bigoplus_{\rho \in X(\pi)} G_n(R(\rho)\#\Gamma).$$

**7.1.20. Theorem [159].** Let $G$ be a non-Abelian group of order $pq$, $p|(q-1)$. Let $\Gamma$ be the unique subgroup of $\text{Gal}(\mathbb{Q}(\xi_q)/\mathbb{Q})$ of order $p$. Then for all $n \geq 0$

$$G_n(\mathbb{Z}G) \simeq G_n(\mathbb{Z}) \oplus G_n\left(\mathbb{Z}\left[\xi_p, \frac{1}{p}\right]\right) \oplus G_n\left(\mathbb{Z}\left[\xi_q, \frac{1}{q}\right]\right)^\Gamma.$$ 

**7.1.21. Remarks.** Note that the dihedral group $D_{2s}$ of order $2s$ has the form $\pi \rtimes \Gamma$ where $\pi$ is a cyclic group of order $s$ and $\Gamma$ has order 2. So, by 7.1.19

$$G_n(\mathbb{Z}D_{2s}) \simeq \bigoplus_{\rho \in X(\pi)} G_n(\mathbb{Z}G)\#\Gamma.$$

We now have a more explicit form of $G_n(\mathbb{Z}D_{2s})$.

**7.1.22. Theorem.** For all $n \geq 0$,

$$G_n(\mathbb{Z}D_{2s}) \simeq \bigoplus_{d|s, d \neq 1} G_n\left(\mathbb{Z}\left[\xi_d, \frac{1}{d}\right]\right) + G_n\left(\mathbb{Z}\left[\frac{1}{2}\right]\right)^\epsilon + G_n(\mathbb{Z}),$$

where $\epsilon = \begin{cases} 1 & \text{if } s \text{ is odd,} \\ 2 & \text{if } s \text{ is even} \end{cases}$

and $\mathbb{Z}[\xi_d, \frac{1}{d}]_+$ is the complex conjugation invariant subring of $\mathbb{Z}[\xi_d, \frac{1}{d}]$.

**7.1.23.** Now, let $G$ be the generalized quaternion group of order $4, 2^4$, i.e.

$$G = \{x, y \mid x^2 = y^2, y^4 = 1, yxy^{-1} = x^{-1}\}.$$  

For $s \geq 0$, let $\Gamma = \{1, \gamma\}$ be a two-element group acting on $\mathbb{Q}[\xi_{2^s}]$ by complex conjugation with fixed field $\mathbb{Q}[\xi_{2^s}]_+$, the maximal real subfield.

Let $c: \Gamma \times \Gamma \to \mathbb{Q}[\xi_{2^{s+1}}]$ be the normalized 2-cocycle given by $c(\gamma, \gamma) = -1$ and let $H = \mathbb{Q}[\xi_{2^{s+1}}]$ be the cross-product algebra, $L$ a maximal $\mathbb{Z}$-order in $H$ (see [58]).

**7.1.24. Theorem [161].** In the notation of 7.1.23

$$G_n(\mathbb{Z}G) \simeq \bigoplus_{j=0}^{s} G_n\left(\mathbb{Z}\left[\xi_{2^j}, \frac{1}{2^j}\right]\right) \oplus G_n\left(\mathbb{Z}\left[\frac{1}{2^{s+1}}\right]\right) \oplus G_n\left(\mathbb{Z}\left[\frac{1}{2}\right]\right)^2.$$
Let \( \pi \) be a finite nilpotent group whose 2-Sylow subgroup has no subquotients isomorphic to the quaternion group of order 8, \( R \) a left Noetherian ring. The rational group algebra \( Q\pi \sim \prod_{\rho \in X(\pi)} Q(\rho) \) where each \( Q(\rho) \) is a simple algebra and \( \rho : \pi \to \text{GL}(V) \) is an irreducible rational representation of \( \pi \) and \( X(\pi) \) is a complete set of inequivalent irreducible rational representations of \( \pi \) and \( Q(\rho) \simeq M_{n(\rho)}(K_{\rho}) \) is a full matrix algebra over a field \( K_{\rho} \) (see [161]).

Let \( Z(\rho) = M_{n(\rho)}(O_{\rho}) \) be a maximal \( A \)-order in \( Q(\rho) \) where \( O_{\rho} \) is the ring of integers in \( K_{\rho} \). Put \( Z(\rho) = Z(\rho)[\frac{1}{|\rho|}] \) where \(|\rho| = \text{index}[\pi : \ker \rho] \). Then we have the following.

7.1.26. Theorem [161]. In the notation of 7.1.25
\[
G_n(R\pi) \simeq \bigoplus_{\rho \in X(\pi)} G_n\big(R(\rho)\big), \quad \text{where } R(\rho) = R \otimes Z(\rho).
\]
Hence
\[
G_n(Z\pi) \simeq \bigoplus_{\rho \in X(\pi)} K_n\big(O_{\rho}\big[\frac{1}{|\rho|}\big]\big).
\]

7.1.27. Now let \( G \) be a finite group of square free order. Then \( G \) is metacyclic and hence can be written as \( G = \pi \rtimes \Gamma \) where \( \pi, \Gamma \) are cyclic of square free order. Then \( QG \simeq Q\pi \# \Gamma \simeq \prod_{\rho \in X(\pi)} Q(\rho) \# \Gamma \) where \( X(\pi) \) is the set of cyclic quotients of \( \pi \).

7.1.28. Theorem [162]. In the notation of 7.1.27 we have
\[
G_n(RG) \simeq \bigoplus_{\rho \in X(\pi)} G_n\big(R(\rho) \# \Gamma\big).
\]

7.1.29. Remarks. For further simplification of each \( G_n(R(\rho) \# \Gamma) \) in the context of a conjecture of Hambleton, Taylor and Williams, see [162].

7.2. Higher class groups of orders and group-rings

7.2.1. Let \( F \) be a number field with ring of integers \( O_F \), \( \Lambda \) any \( R \)-order in a semi-simple \( F \)-algebra \( \Sigma \). In this subsection, we review briefly some results on the higher class groups \( Cl_n(\Lambda) \), \( n \geq 0 \), of \( \Lambda \) which constitute natural generalization to higher dimensions of the classical notion of class groups \( Cl(\Lambda) \) of an order (see [19] or [79]).

The class groups \( Cl(\Lambda) \), apart from generalizing the classical notion of class groups of integers in number fields, is also intimately connected with representation theory while \( Cl(\mathbb{Z}G) \) is also known to house several topological/geometric invariants, see [156]. For various computations of \( Cl(\Lambda) \), see [19].
7.2.2. DEFINITION. Let $F$ be a number field with ring of integers $O_F$, $\Lambda$ an $O_F$-order in a semi-simple algebra. The higher class groups are defined for all $n \geq 0$ by

$$Cl_n(\Lambda) = \text{Ker} \left( SK_n(\Lambda) \to \bigoplus_p SK_n(\hat{\Lambda}_p) \right),$$

where $p$ runs through all prime ideals of $O_F$.

NOTE. One can show that $Cl_0(\Lambda) = Cl(\Lambda)$. Moreover, $Cl_1(O_F G)$ is intimately connected with the Whitehead groups of $G$.

The result below, due to Kuku, [73], says that $\forall n \geq 1$, $Cl_n(\Lambda)$ are finite groups for any $O_F$-orders $\Lambda$.

7.2.3. THEOREM [74]. Let $F$ be a number field with ring of integers $O_F$, $\Lambda$ any $O_F$-order in a semi-simple $F$-algebra $\Sigma$. Then the groups $Cl_n(\Lambda)$ are finite.

Note that $Cl_0(\Lambda) = Cl(\Lambda)$ is finite is a classical result. See [19].

The next result due to M.E. Keating says that $Cl_n(\Gamma)$ vanishes for maximal orders $\Gamma$.

7.2.4. THEOREM [60]. Let $F$, $O_F$ be as in 7.2.3, $\Gamma$ a maximal $O_F$-order in a semi-simple algebra. Then for all $n \geq 1$, $Cl_n(\Gamma) = 0$.

The next result 7.2.6 due to Kolster and Laubenbacher, [65], gives information on possible torsion in odd-dimensional class groups. First we make some observations.

7.2.5. REMARKS. Let $F$ be a number field with ring of integers $O_F$, $\Lambda$ any $O_F$-order in a semi-simple $F$-algebra $\Sigma$, and $\Gamma$ a maximal order containing $\Lambda$. It is well known that $\hat{\Lambda}_p$ is a maximal order for all except a finite number of prime ideals $p$ of $O_F$ at which $\Lambda_p$ is not maximal. We denote by $P_\Lambda$ the set of rational primes $q$ lying below the primes ideals $p$ in $S_\Lambda$, where $S_\Lambda$ denotes the finite set of prime ideals $p$ of $O_F$ for which $\Lambda_p$ is not maximal.

7.2.6. THEOREM [65]. Let $F$, $O_F$, $\Lambda$ be as in 7.2.3. Then for all $n \geq 1$, $Cl_{2n-1}(\Lambda)(q) = 0$ for $q \notin P_\Lambda$.

7.2.7. COROLLARY. Let $G$ be a finite group. Then for all $n \geq 1$, the only $p$-torsion that can possibly occur in $Cl_{2n-1}(O_F G)$ are for those $p$ dividing the order of $G$.

7.2.8. REMARKS. Finding out what $p$-torsion could occur in even-dimensional class groups of arbitrary orders in semi-simple algebras is still open. However, we have the following result 7.2.9 due to Guo and Kuku, [48], analogous to 7.2.6 above for Eichler and hereditary orders.

7.2.9. THEOREM [48]. Let $F$ be a number field with ring of integers $O_F$, $\Lambda$ an Eichler order in a quaternion $F$-algebra or a hereditary order in a semi-simple $F$-algebra. Then, in the notation of 7.2.3, $Cl_{2n}(\Lambda)(q) = 0$ for $q \notin P_\Lambda$. 
7.3. Profinite higher $K$-theory of orders and group-rings

7.3.1. Let $R$ be a Dedekind domain with quotient field $F$, $A$ any $R$-order in a semi-simple $F$-algebra. In the notation of 3.1.6(xi) and [78], we shall write $K^n_{pr}(\mathcal{P}(A), \hat{\mathbb{Z}}_{\ell})$ for $K^n_{pr}(\mathcal{P}(A), \mathbb{Z}/\ell)$ and $G^n_{pr}(\mathcal{M}(A), \hat{\mathbb{Z}}_{\ell})$ for $K^n_{pr}(\mathcal{M}(A), \mathbb{Z}/\ell)$ respectively for the profinite $K$-theory and $G$-theory of $A$. Also, following the notations in 3.1.6(ix) and as in [78], we shall write

$$K_n(A, \hat{\mathbb{Z}}_{\ell}) = \lim_{\leftarrow} K_n(P(A), \mathbb{Z}/\ell), \quad G_n(A, \hat{\mathbb{Z}}_{\ell}) = \lim_{\leftarrow} G_n(M(A), \mathbb{Z}/\ell).$$

Note that this situation applies to $R$ being ring of integers in number fields and $p$-adic fields. Note that in the statements above and throughout this subsection $\ell$ is a prime.

Although several of the computations in [78] do apply to general exact categories we shall restrict ourselves in this section to results on orders and group rings. The results in this subsection are all due to Kuku, [78].

7.3.2. Theorem [78]. Let $R$ be the ring of integers in a number field $F$, $A$ an $R$-order in a semi-simple $F$-algebra. Then we have

1. For all $n \geq 1$
   \[ K_n(A)[\ell^s] \cong K^n_{pr}(A, \hat{\mathbb{Z}}_{\ell})[\ell^s]; \quad G_n(A)[\ell^s] \cong G^n_{pr}(A, \hat{\mathbb{Z}}_{\ell})[\ell^s]. \]

2. For all $n \geq 2$
   \[ K_n(A)/\ell^s \cong K^n_{pr}(A, \hat{\mathbb{Z}}_{\ell})/\ell^s; \quad G_n(A)/\ell^s \cong G^n_{pr}(A, \hat{\mathbb{Z}}_{\ell})/\ell^s. \]

3. (a) $K_n(A) \otimes \hat{\mathbb{Z}}_{\ell} \cong K^n_{pr}(A, \hat{\mathbb{Z}}_{\ell}) \cong K_n(A, \hat{\mathbb{Z}}_{\ell})$ are $\ell$-complete profinite Abelian groups for all $n \geq 2$.
   (b) $G_n(A) \otimes \hat{\mathbb{Z}}_{\ell} \cong G^n_{pr}(A, \hat{\mathbb{Z}}_{\ell}) \cong G_n(A, \hat{\mathbb{Z}}_{\ell})$ are $\ell$-complete profinite Abelian groups for all $n \geq 2$.
   (c) $K_{2n+1}(\Sigma) \otimes \hat{\mathbb{Z}}_{\ell} \cong K^n_{pr}(\Sigma, \hat{\mathbb{Z}}_{\ell}) \cong K_{2n-1}(\Sigma, \hat{\mathbb{Z}}_{\ell})$ are $\ell$-complete profinite Abelian groups for all $n \geq 2$.

Note. An Abelian group $G$ is said to be $\ell$-complete if $G = \lim(G/\ell^s)$.

We also have the following results in the local situation.

7.3.3. Theorem [78]. Let $p$ be a rational prime, $F$ a $p$-adic field (i.e. any finite extension of $\hat{\mathbb{Q}}_p$), $R$ the ring of integers of $F$, $\Gamma$ a maximal $R$-order in a semi-simple $F$-algebra $\Sigma$. Then for all $n \geq 2$ we have

1. $K^n_{pr}(\Sigma, \hat{\mathbb{Z}}_{\ell}) \cong K_n(\Sigma, \hat{\mathbb{Z}}_{\ell})$ is an $\ell$-complete profinite Abelian group.

7.3.4. Theorem [78]. Let $R$ be the ring of integers in a $p$-adic field $F$, $\ell$ a rational prime such that $l \neq p$, $A$ an $R$-order in a semi-simple $F$-algebra $\Sigma$. Then, for all $n \geq 2$, $G^n_{pr}(A, \hat{\mathbb{Z}}_{\ell})$ is an $\ell$-complete profinite Abelian group.
7.3.5. **Theorem** [78]. Let \( R \) be the ring of integers in a \( p \)-adic field \( F \), \( \Lambda \) an \( R \)-order in a semi-simple \( F \)-algebra \( \Sigma \), \( \ell \) a rational prime such that \( \ell \neq p \). Then for all \( n \geq 1 \),

(i) \( G_n(\Lambda)_{\ell} \) are finite groups.

(ii) \( K_n(\Sigma)_{\ell} \) are finite groups.

(iii) The kernel and cokernel of \( G_n(\Lambda) \to G_n^p(\Lambda, \hat{\mathbb{Z}}_{\ell}) \) are uniquely \( \ell \)-divisible.

(iv) The kernel and cokernel of \( K_n(\Sigma) \to K_n^p(\Sigma, \hat{\mathbb{Z}}_{\ell}) \) are uniquely \( \ell \)-divisible.

7.3.6. **Theorem** [78]. Let \( R \) be the ring of integers in a number field \( F \), \( \Lambda \) an \( R \)-order in a semi-simple \( F \)-algebra \( \Sigma \) satisfying the hypothesis of 7.1.8, \( p \) any prime ideal of \( R \) and \( \ell \) a rational prime such that \( \ell \neq \text{char}(R/p) \). Then

(i) \( K_n(\hat{\Lambda}_p)_{\ell} \) is a finite group.

(ii) The map \( \varphi : K_n(\hat{\Lambda}_p)_{\ell} \to K_n^p(\hat{\Lambda}_p, \hat{\mathbb{Z}}_{\ell})_{\ell} \) is an isomorphism.

7.4. **Higher \( K \)-theory of modules over EI categories**

7.4.1. An EI category \( \mathcal{C} \) is a small category in which every endomorphism is an isomorphism.

\( \mathcal{C} \) is said to be finite if the set \( \text{Is}(\mathcal{C}) \) of isomorphism classes of \( \mathcal{C} \)-objects is finite and for any two \( \mathcal{C} \)-objects \( X, Y \) the set \( \text{mor}_{\mathcal{C}}(X, Y) \) of \( \mathcal{C} \)-morphisms from \( X \) to \( Y \) is also finite.

Let \( R \) be a commutative ring with identity. An \( R\mathcal{C} \) module is a contravariant functor from \( \mathcal{C} \) to the category \( R\text{-mod} \) of \( R \)-modules. There is a notion of finitely generated projective \( R\mathcal{C} \)-modules as well as a notion of finitely generated \( R\mathcal{C} \)-modules (see 7.4.3 below). So, for all \( n > 0 \), let \( K_n(R\mathcal{C}) \) be the (Quillen) \( K_n \) of the category \( P(R\mathcal{C}) \) and \( G_n(R\mathcal{C}) \) the \( K_n \) of the category \( M(R\mathcal{C}) \) when \( R \) is Noetherian.

The significance of the study of \( K \)-theory of \( R\mathcal{C} \) modules lies mainly in the fact that several geometric invariants take values in the \( K \)-groups associated with \( R\mathcal{C} \) where \( \mathcal{C} \) is an appropriately defined EI-category and \( R \) could be \( \mathbb{C} \), \( \mathbb{Q} \), \( \mathbb{R} \), etc. For example, if \( G \) is a finite group, \( \mathcal{C} = \text{orb}(G) \) (a finite EI-category), \( X \) a \( G \)-CW complex with round structure (see [96]), then the equivariant Reidemeister torsion takes values in \( \text{Wh}(\mathcal{Q}\text{orb}(G)) \) where \( \text{Wh}(\mathcal{Q}\text{orb}(G)) \) is the quotient of \( K_1(\mathcal{Q}\text{orb}(G)) \) by the subgroups of “trivial units” — see [96]. For more invariants in the lower \( K \)-theory of modules over suitable EI-categories, see [96].

The study of modules over EI-categories is a natural generalization of study of modules over group rings.

7.4.2. **Examples**.

(i) Let \( G \) be a finite group, \( \text{ob}(\mathcal{C}) = \{G/H \mid H \leq G\} \) and let the morphisms be \( G \)-maps. Then \( \mathcal{C} \) is a finite EI-category called the orbit category of \( G \) and denoted by \( \text{orb}(G) \). Note that here, \( \mathcal{C}(G/H, G/H) \simeq \text{Aut}(G/H) \simeq N_G(H)/H \) where \( N_G(H) \) is the normalizer of \( H \) in \( G \) (see [75]). We shall denote \( N_G(H)/H \) by \( W_G(H) \).

(ii) Let \( G \) be a Lie group, \( \text{ob}(\mathcal{C}) = \{G/H \mid H \text{ a compact subgroup of } G\} \). Again the morphisms are \( G \)-maps and \( \mathcal{C} \) is also called the orbit category of \( G \) and denoted by \( \text{orb}(G) \).
(iii) Let $G$ be a Lie group, $\text{ob}(\mathcal{C}) = \{G/H \mid H \text{ compact subgroup of } G\}$. For $G/H$, $G/H' \in \text{ob}(\mathcal{C})$ let $\mathcal{C}(G/H, G/H')$ be the set of homotopy classes of $G$-maps. Then $\mathcal{C}$ is an EI-category called the discrete orbit category of $G$ and denoted by $\text{orb} / (G)$.

(iv) Let $G$ be a discrete group, $\mathcal{F}$ the family of all finite subgroups of $G$. Consider the category $\text{Or}_{\mathcal{F}}(G)$ such that $\text{ob}(\text{Or}_{\mathcal{F}}(G)) = \{G/H \mid H \in \mathcal{F}\}$. $\text{Or}_{\mathcal{F}}(G)$ is an EI category.

7.4.3. Let $R$ be a commutative ring with identity, $\mathcal{C}$ an EI-category. A $\mathcal{C}$-module is a contravariant functor $\mathcal{C} \to R\text{-mod}$.

An $\text{ob}\mathcal{C}$-set is a functor $N$ from $\mathcal{C}$ to the category of sets. Alternatively, an $\text{ob}\mathcal{C}$-set could be visualized as a pair $(N, \beta)$ where $N$ is a set and $\beta : N \to \text{ob}\mathcal{C}$ is a set map. Then $N[\beta^{-1}(X) \mid X \in \text{ob}\mathcal{C}]$.

An $\text{ob}\mathcal{C}$-set $(N, \beta)$ is said to be finite if $N$ is a finite set. If $S$ is an $(N, \beta)$-subset of an $\mathcal{C}$-module $M$, define span $S$ as the smallest $\mathcal{C}$-submodule of $M$ containing $S$. Say that $M$ is finitely generated if $M = \text{span} S$ for some finite $\text{ob}\mathcal{C}$-subset $S$ of $M$.

If $R$ is a Noetherian ring and $\mathcal{C}$ an EI-category, let $M(R\mathcal{C})$ be the category of finitely generated $R\mathcal{C}$-modules. Then $M(R\mathcal{C})$ is an exact category in the sense of Quillen, see [114].

An $\mathcal{C}$-module $P$ is said to be projective if any exact sequence of $\mathcal{C}$-modules $0 \to M' \to M \to P \to 0$ splits or equivalently if $\text{Hom}_{\mathcal{C}}(P, -)$ is exact.

Let $P(R\mathcal{C})$ be the category of finitely generated projective ($\mathcal{C}$)-modules. The $P(R\mathcal{C})$ is also exact. We write $K_n(R\mathcal{C})$ for $K_n(P(R\mathcal{C}))$.

Now, let $R$ be a commutative ring with identity, $\mathcal{C}$ an EI-category, $P(R\mathcal{C})$ the category of finitely generated $R\mathcal{C}$-modules $M$ such that for each $X \in \text{ob}\mathcal{C}$ $M(X)$ is projective as $R$-module. Then $P(R\mathcal{C})$ is an exact category and we write $G_n(R, \mathcal{C})$ for $K_n(P(R\mathcal{C}))$. Note that if $R$ is regular, then $G_n(R, \mathcal{C}) \simeq G_n(R\mathcal{C})$.

Finally, if $R$ is a Dedekind domain with quotient field $F$, the inclusion functor $\mathcal{P}(R\mathcal{C}) \hookrightarrow \mathcal{M}(R\mathcal{C})$ ($\mathcal{C}$ an EI-category) induces the Cartan maps $K_n(R\mathcal{C}) \to G_n(R\mathcal{C})$. Define $SK_n(R\mathcal{C}) := \text{Ker}(K_n(R\mathcal{C}) \to K_n(F\mathcal{C}))$, $SG_n(R\mathcal{C}) = \text{Ker}(G_n(R\mathcal{C}) \to G_n(F\mathcal{C}))$.

First we have the following splitting result:

7.4.4. Theorem.

(i) Let $R$ be a commutative ring with identity, $\mathcal{C}$ an EI-category. Then

$$K_n(R\mathcal{C}) \simeq \bigoplus_{X \in I_1(\mathcal{C})} K_n\left(R(\text{Aut}(X))\right).$$

(ii) If $R$ is commutative Noetherian ring and $\mathcal{C}$ a finite EI-category then

$$G_n(R\mathcal{C}) \simeq \bigoplus_{X \in I_1(\mathcal{C})} G_n\left(R(\text{Aut}(X))\right).$$

We now record the following.

7.4.5. Theorem. Let $R$ be the ring of integers in a number field $F$, $\mathcal{C}$ any finite EI-category. Then for all $n \geq 1$, we have
Higher algebraic K-theory

(i) $K_n(RC)$ is finitely generated Abelian group.
(ii) $G_n(RC)$ is a finitely generated Abelian group.
(iii) $SK_n(RC), SK_n(\widehat{R}_pC)$ are finite groups.
(iv) $SG_n(RC), SG_n(\widehat{R}_pC)$ are finite groups.

Finally, we present the following result on the Cartan map.

7.4.6. THEOREM. Let $k$ be a field of characteristic $p, C$ a finite EI-category. Then for all
$n \geq 0$, the Cartan homomorphism $K_n(kC) \to G_n(kC)$ induce an isomorphism
\[
\mathbb{Z}\left(\frac{1}{p}\right) \otimes K_n(kC) \simeq \mathbb{Z}\left(\frac{1}{p}\right) \otimes G_n(kC).
\]

8. Equivariant higher algebraic $K$-theory together with relative generalizations

In order to economize on time and space, we restrict our discussion of equivariant higher
$K$-theory in this section to finite group actions, with the remark that there are analogous theories for profinite group actions (see [70,80]) and compact Lie group actions (see
[77,80]).

8.1. Equivariant higher algebraic $K$-theory

8.1.1. Let $G$ be a finite group, $S$ a $G$-set and $S$ the category associated with $S$ (or translation category of $S$), see [25,72].

If $D$ is an exact category in the sense of Quillen, [114], then the category of $[S,D]$ of covariant functors from $S$ to $D$ is also exact (see [25]).

8.1.2. DEFINITION. Let $K^G_n(S,D)$ be the $n$-th algebraic $K$-group associated to the category $[S,D]$ with respect to fibre-wise exact sequences.

We now have the following:

8.1.3. THEOREM. $K^G_n(\dashv D): GSet \to Ab$ is a Mackey functor.

PROOF. See [24] and [25].

We now want to turn $K^G_0(\dashv D)$ into a Green functor. We first recall the definition of a pairing of exact categories (see [110]).

8.1.4. DEFINITION. Let $D_1, D_2, D_3$ be exact categories. An exact pairing $\langle \cdot, \cdot \rangle : D_1 \times D_2 \to D_3$ given by $(X_1, X_2) \to \langle X_1 \circ X_2 \rangle$ is a covariant functor such that
Relative equivariant higher algebraic $K$-theory

In this section, we discuss the relative version of the theory in 8.1.

8.2.1. **Definition.** Let $S$, $T$ be $G$-sets. Then the projection map $S \times T \to S$ gives rise to a functor $S \times T \rightarrow S$. Suppose that $\mathcal{D}$ is an exact category in the sense of Quillen, [78]. Then, a sequence $\xi_1 \rightarrow \xi_2 \rightarrow \xi_3$ of functors in $[S, \mathcal{D}]$ is said to be $T$-exact if the sequence $\xi'_1 \rightarrow \xi'_2 \rightarrow \xi'_3$ of restricted functors $S \times T \rightarrow S \rightarrow \mathcal{D}$ is split exact.

If $\psi : S_1 \rightarrow S_2$ is a $G$-map, and $\xi_1 \rightarrow \xi_2 \rightarrow \xi_3$ is a $T$-exact sequence in $[S_2, \mathcal{Q}]$, then $\xi'_1 \rightarrow \xi'_2 \rightarrow \xi'_3$ is a $T$-exact sequence in $[S_1, \mathcal{D}]$ where $\xi'_1 : S_1 \rightarrow S_2 \xrightarrow{\psi} S_2 \xrightarrow{\xi_1} \mathcal{D}$.

Let $K_n^G(S, \mathcal{D}, T)$ be the $n$-th algebraic $K$-group associated to the exact category $[S, \mathcal{D}]$ with respect to $T$-exact sequences.

8.2.2. **Definition.** Let $S$, $T$ be $G$-sets. A functor $\xi \in [S, \mathcal{D}]$ is said to be $T$-projective if any $T$-exact sequence $\xi_1 \rightarrow \xi_2 \rightarrow \xi$ is split exact. Let $[S, \mathcal{D}]_T$ be the additive category of $T$-projective functors in $[S, \mathcal{D}]$ considered as an exact category with respect to split exact sequences. Note that the restriction functor associated to $S_1 \rightarrow S_2$ carries $T$-projective functors $\xi \in [S_2, \mathcal{D}]$ into $T$-projective functors $\xi \circ \psi \in [S_1, \mathcal{D}]$. Define $P_n^G(S, \mathcal{D}, T)$ as the $n$-th algebraic $K$-group associated to the exact category $[S, \mathcal{D}]_T$, with respect to split exact sequences.

8.2.3. **Theorem [35].** $K_n^G(\mathcal{D}, T)$ and $P_n^G(\mathcal{D}, T)$ are Mackey functors from $G$-Set to $Ab$ for all $n \geq 0$. If the pairing $\mathcal{D} \times \mathcal{D} \rightarrow \mathcal{D}$ is naturally associative and commutative and $\mathcal{D}$ contains a natural unit, then $K_0^G(\mathcal{D}, T) : GSet \rightarrow Ab$ is a Green functor and $K_n^G(\mathcal{D}, T)$ and $P_n^G(\mathcal{D}, T)$ are $K_0^G(\mathcal{D}, T)$ modules.

**Proof.** See [24] or [25].
8.3. Interpretation in terms of group-rings

In this section, we discuss how to interpret the theories in the two previous sections in terms of group-rings.

8.3.1. Recall that any $G$-set $S$ can be written as a finite disjoint union of transitive $G$-sets each of which is isomorphic to a quotient set $G/H$ for some subgroup $H$ of $G$. Since Mackey functors, by definition, take finite disjoint unions into finite direct sums, it will be enough to consider exact categories $[G/H, D]$ where $D$ is an exact category in the sense of Quillen, [78].

For any ring $A$, let $M(A)$ be the category of finitely generated $A$-modules and $P(A)$ the category of finitely generated projective $A$-modules.

8.3.2. **Theorem [72].** Let $G$ be a finite group, $H$ a subgroup of $G$, $A$ a commutative ring with identity, then there exists an equivalence of exact categories $[G/H, M(A)] \to M(AH)$. Under this equivalence, $[G/H, P(A)]$ is identified with the category of finitely generated $A$-projective $AH$-modules.

We also observe that a sequence of functors $\varphi_1 \to \varphi_2 \to \varphi_3$ in $[G/H, P(A)]$ or $[G/H, M(A)]$ is exact if the corresponding sequence $\varphi_1(H) \to \varphi_2(H) \to \varphi_3(H)$ of $AH$-modules is exact.

8.3.3. **Remarks.**

(i) It follows that for every $n \geq 0$, $K^n(G/H, P(A))$ can be identified with the $n$-th algebraic $K$-group of the category of finitely generated $A$-projective $AH$-modules while $K^n(G/H, M(A)) = G^n(AH)$ if $A$ is Noetherian. It is well known that $K^n(G/H, P(A)) = K^n(G/H, M(A))$ is an isomorphism when $A$ is regular.

(ii) Let $\varphi: G/H_1 \to G/H_2$ be a $G$-map for $H_1 \leq H_2 \leq G$. We may restrict ourselves to the case $H_2 = G$ and so we have $\varphi_*: [G/G, M(A)] \to [G/H, M(A)]$ corresponding to the restriction functor $M(AG) \to M(AH)$, while $\varphi^*: [G/H, M(A)] \to [G/G, M(A)]$ corresponds to the induction functor $M(AH) \to M(AG)$ given by $N \to AG \otimes_{AH} N$. Similar situations hold for functor categories involving $P(A)$.

(iii) If $D = P(A)$ and $A$ is commutative, then the tensor product defines a naturally associative and commutative pairing $P(A) \times P(A) \to P(A)$ with a natural unit and so $K^n_D(\mathcal{E}, P(A))$ are $K^n_G(\mathcal{E}, P(A))$-modules.

8.3.4. We now interpret the relative situation. So let $T$ be a $G$-set. Note that a sequence $\xi_1 \to \xi_2 \to \xi_3$ of functors in $[G/H, P(A)]$ or $[G/H, M(A)]$ is said to be $T$-exact if $\xi_1(H) \to \xi_2(H) \to \xi_3(H)$ is $AH'$-split exact for all $H' \leq H$ such that $T^{H'} \neq \emptyset$ where $T^{H'} = \{ t \in T \mid g t = t \forall g \in H' \}$. In particular, the sequence is $G/H$-exact (resp. $G/G$-exact) if the corresponding sequence of $AH$-modules (resp. $AG$-modules) is split exact. If $\mathcal{E}$ is the trivial subgroup of $G$, it is $G/\mathcal{E}$-exact if it is split exact as a sequence of $A$-modules.
So, $K_n^G(G/H, P(A), T)$ (resp. $K_n^G(G/H, M(A), T)$) is the $n$-th algebraic $K$-group of the category of finitely generated $A$-projective $AH$-modules (resp. category of finitely generated $AH$-modules) with respect to exact sequences which split when restricted to the various subgroups $H'$ of $H$ such that $T^{H'} \neq \emptyset$.

Moreover, observe that $P_n^G(G/H, p(A), T)$ (resp. $P_n^G(G/H, M(A), T)$) is an algebraic $K$-group of the category of finitely generated $A$-projective $AH$-modules (resp. finitely generated $AH$-modules) which are relatively $H'$-projective for subgroups $H'$ of $H$ such that $T^{H'} \neq \emptyset$ with respect to split exact sequences. In particular, $P_0^G(G/H, P(A), G/\varepsilon) = K_0(AH)$. If $A$ is commutative, the $K_0^G(-, P(A), T)$ is a Green functor and $K_0^G(-, P(A), T)$ and $P_0^G(-, P(A), T)$ are $K_0^G(-, P(A), T)$-modules.

Now, let us interpret the maps associated to $G$-maps $S_1 \to S_2$. We may specialize to maps $\varphi: G/H_1 \to G/H_2$ for $H_1 \subseteq H_2 \subseteq G$ and for convenience we may restrict ourselves to the case $H_2 = G$, in which case we write $H_1 = H$. In this case $\varphi: [G/H, M_0(A)] \to [G/H, M_0(A)]$ corresponds to the restriction of $AG$-modules to $AH$-modules and $\varphi^*: [G/H, M(A)] \to [G/H, M(A)]$ corresponds to the induction of $AH$-modules to $AG$-modules (see [21]).

We hope that this wealth of equivariant higher algebraic $K$-groups will satisfy a lot of future needs, and moreover, that the way they have been produced systematically, will help to keep them in some order, and to produce new variants of them, whenever desired.

Since any $G$-set $S$ can be written as a disjoint union of transitive $G$-sets, isomorphic to some coset-set $G/H$, and since all the above $K$-functors satisfy the additivity condition, the above identifications extend to $K$-groups, defined on an arbitrary $G$-set $S$.

### 8.4. Some applications

We are now in a position to draw various conclusions just by quoting well established induction theorems, concerning $K_0^G(-, P(A))$ and $K_0^G(-, P(A); T)$ and, more generally $R \otimes \mathbb{Z} K_0^G(-, P(A))$ and $R \otimes \mathbb{Z} K_0^G(-, P(A); T)$ for $R$ a subring of $\mathbb{Q}$ or just any commutative ring (see [23,20–22]). Since any exact sequence in $P(A)$ is split exact, we have a canonical identification $K_0^G(-, P(A)) = K_0^G(-, P(A); G/\varepsilon)$ ($\varepsilon \subseteq G$ the trivial subgroup) and, thus, may direct our attention to the relative case, only.

So let $T$ be a $G$-set. For $p$ a prime and $q$ a prime or 0, let $D(p, T, q)$ denote the set of subgroups $H \leq G$ such that the smallest normal subgroup $H_1$ of $H$ with a $q$-factor group has a normal Sylow-subgroup $H_2$ with $T^{H_2} \neq \emptyset$ and a cyclic factor group $H_1/H_2$. Let $\mathcal{H}_q$ denote the set of subgroups $H \leq G$ which are $q$-hyperelementary, i.e. have a cyclic normal subgroup with a $q$-factor group (or are cyclic for $q = 0$).

For $A$ and $R$ commutative rings, let $D(A, T, R)$ denote the union of all $D(p, T, q)$ with $pA \neq A$ and $qR \neq R$ and let $\mathcal{H}_R$ denote the union of all $\mathcal{H}_q$ with $qR \neq R$.

Then it has been proved (see [20,21,23]), that $R \otimes_A K_0^G(-, P(A); T)$ is $S$-projective for some $G$-set $S$ if $S^H \neq \emptyset$ for all $H \in D(A, T, R) \cup \mathcal{H}_R$. Moreover (see [21]), if $A$ is a field of characteristic $p \neq 0$, then $K_0^G(-, P(A); T)$ is $S$-projective, already if $S^H \neq \emptyset$ for all $H \in D(A, T, R)$.
8.4.1. Among the many possible applications of these results, we discuss just one special case. Let $A = k$ be a field of characteristic $p \neq 0$, let $R = \mathbb{Z}(\frac{1}{p})$ and let $S = \bigcup_{H \in D(k, T, R)} G/H$. Then $R \otimes K^n_G(\cdot, P(k); T)$ is an isomorphism for any $G$-set $X$, for which the Sylow-$p$-subgroups $H$ of the stabilizers of the elements in $X$ are $T$-exact. Moreover, the Cartan map $P^n_G(X, P(k); T) \to K^n(G, X, P(k); T)$ is an isomorphism for all $G$-sets $X$, for which the Sylow-$p$-subgroups $H$ of the stabilizers of the elements in $X$ are $T$-exact. This implies in particular that for all $G$-sets $X$ the Cartan map $P^n_G((X \times S), P(k); T) \to K^n_G((X \times S), P(k); T)$ is an isomorphism, since any stabilizer group of an element in $X \times S$ is a subgroup of a stabilizer group of an element in $S$ and thus, by the very definition of $S$ and $D(k, T, \mathbb{Z}(\frac{1}{p}))$, has a Sylow-$p$-subgroup $H$ with $T^H \neq \emptyset$.

This finally implies that $P^n_G((\cdot, P(k); T)_S \to K^n_G((\cdot, P(k); T)_S$ is an isomorphism, so by the general theory of Mackey functors,

$$\mathbb{Z}\left(\frac{1}{p}\right) \otimes P^n_G((\cdot, P_kT) \to \mathbb{Z}\left(\frac{1}{p}\right) \otimes K^n_G((\cdot, P_kT)$$

is an isomorphism. In the special case $T = G/\varepsilon$ this is just the $K$-theory of finitely generated projective $kG$-modules. $K^n_G((\cdot, P_kG; G/\varepsilon)$ the $K$-theory of finitely generated $G$-modules with respect to exact sequences.

Thus, we have proved:

8.4.2. **Theorem.** Let $k$ be a field of characteristic $p$, $G$ a finite group. Then for all $n \geq 0$ the Cartan map $K_n(kG) \to G_n(kG)$ induces isomorphisms

$$\mathbb{Z}\left(\frac{1}{p}\right) \otimes K_n(kG) \simeq \mathbb{Z}\left(\frac{1}{p}\right) \otimes G_n(kG).$$

Finally, with the identification of Mackey functors: $GSet \to Ab$ with Green’s G-functors $\delta G \to Ab$ as in [72], and the above interpretations of our equivariant theory in terms of group rings, we now have from the foregoing, the following result 8.4.3 which says that higher algebraic $K$-groups are hyper-elementary computable.

For the proof of this result, see [69].

8.4.3. **Theorem** [69]. Let $R$ be a Dedekind domain, $G$ a finite group, $M$ any of the functors

$$K_n(R-), G_n(R-), SG_n(R-), SK_n(R-); \delta G \to \mathbb{Z} \text{-mod}.$$
For any commutative ring \( A \) with identity, define \((A \otimes M)(H) = A \otimes M(H), H \in \delta G\).

Let \( P \) be a set of rational primes, \( \mathbb{Z}_p = \{ \frac{1}{q} \mid q \notin P \} \), \( C(G) \) the collection of all cyclic subgroups of \( G \), \( h \in C(G) \), \( A = \{ H \leq G \mid \exists H' \in C(G), H/H' a \text{ p-group for some } p \in P \} \).

Then \( \mathbb{Z}_p \otimes M(G) = \lim \leftarrow H \mathbb{Z}_p \otimes M(H) \) where \( \lim \leftarrow \mathbb{Z}_p \otimes M(H) \) is the subgroup of all \( (x_H) \in \prod_{H \in A} \mathbb{Z}_p \otimes M(H) \) such that for any \( H, H' \in A \) satisfying \( gH'g^{-1} \subset H, \varphi : H' \to H \) given by \( \varphi(h) = ghg^{-1} \), then \( \mathbb{Z}_p \otimes M(\varphi)(x_H) = x_{H'} \).

Acknowledgement

This chapter was written while I was a member of the Institute for Advanced Study, Princeton during the 2003–2004 Academic year. I am grateful to the Institute for hospitality and financial support.

References

Higher algebraic K-theory

[56] B. Kahn, K-theory of semi-local rings with finite coefficients and étale cohomology.
A. Kuku

Higher algebraic $K$-theory

Section 3B
Associative Rings and Algebras
This page intentionally left blank
# Filter Dimension

V. Bavula

*Department of Pure Mathematics, University of Sheffield, Hicks Building, Sheffield S3 7RH, UK*

*E-mail: v.bavula@sheffield.ac.uk*

## Contents

1. Introduction ........................................ 79
2. Filter dimension of algebras and modules ................. 80
   2.1. The Gelfand–Kirillov dimension ..................... 80
   2.2. The return functions and the (left) filter dimension 81
3. The first filter inequality ................................ 83
4. Krull, Gelfand–Kirillov and filter dimensions of simple finitely generated algebras .................. 85
5. Filter dimension of the ring of differential operators on a smooth irreducible affine algebraic variety (proof of Theorem 4.3) .................................. 89
6. Multiplicity for the filter dimension, holonomic modules over simple finitely generated algebras ............ 91
   6.1. Multiplicity in the commutative situation .......... 91
   6.2. Somewhat commutative algebras ........................ 92
   6.3. Multiplicity ........................................... 93
   6.4. Holonomic modules .................................... 94
7. Filter dimension and commutative subalgebras of simple finitely generated algebras and their division algebras ........................................................................................................... 97
   7.1. An upper bound for the Gelfand–Kirillov dimensions of commutative subalgebras of simple finitely generated algebras .......................................................... 97
   7.2. An upper bound for the transcendence degree of subfields of quotient division algebras of simple finitely generated algebras .................................................. 99
8. Filter dimension and isotropic subalgebras of Poisson algebras .................................................. 102
References .................................................. 105
1. Introduction

Throughout the chapter, $K$ is a field, a module $M$ over an algebra $A$ means a left module denoted $AM \cong \otimes K$.

Intuitively, the filter dimension of an algebra or a module measures how ‘close’ standard filtrations of the algebra or the module are. In particular, for a simple algebra it also measures the growth of how ‘fast’ one can prove that the algebra is simple.

The filter dimension appears naturally when one wants to generalize the Bernstein’s inequality for the Weyl algebras to the class of simple finitely generated algebras.

The $n$-th Weyl algebra $A_n$ over the field $K$ has $2n$ generators $X_1, \ldots, X_n, \partial_1, \ldots, \partial_n$ that satisfy the defining relations

$$\partial_i X_j - X_j \partial_i = \delta_{ij}, \quad \text{the Kronecker delta,} \quad X_i X_j - X_j X_i = \partial_i \partial_j - \partial_j \partial_i = 0,$$

for all $i, j = 1, \ldots, n$. When char $K = 0$ the Weyl algebra $A_n$ is a simple Noetherian finitely generated algebra canonically isomorphic to the ring of differential operators $K[X_1, \ldots, X_n, \frac{d}{dX_1}, \ldots, \frac{d}{dX_n}]$ with polynomial coefficients ($X_i \leftrightarrow X_i, \partial_i \leftrightarrow \frac{d}{dX_i}, \ i = 1, \ldots, n$).

Let $\text{Kdim}$ and $\text{GK}$ be the (left) Krull (in the sense of Rentschler and Gabriel, [22]) and the Gelfand–Kirillov dimension respectively.

**Theorem 1.1** (The Bernstein’s inequality, [8]). Let $A_n$ be the $n$-th Weyl algebra over a field of characteristic zero. Then $\text{GK}(M) \geq n$ for all non-zero finitely generated $A_n$-modules $M$.

Let $A$ be a simple finitely generated infinite-dimensional $K$-algebra. Then $\text{dim}_K(M) = \infty$ for all non-zero $A$-modules $M$ (the algebra $A$ is simple, so the $K$-linear map $A \to \text{Hom}_K(M, M)$, $a \mapsto (m \mapsto am)$, is injective, and so $\infty = \text{dim}_K(A) \leq \text{dim}_K(\text{Hom}_K(M, M))$ hence $\text{dim}_K(M) = \infty$). So, the Gelfand–Kirillov dimension (over $K$) $\text{GK}(M) \geq 1$ for all non-zero $A$-modules $M$.

**Definition.** $h_A := \inf \{\text{GK}(M) \mid M \text{ is a non-zero finitely generated } A\text{-module}\}$ is called the holonomic number for the algebra $A$.

**Problem.** For a simple finitely generated algebra find its holonomic number.

To find an approximation of the holonomic number for simple finitely generated algebras and to generalize the Bernstein inequality for these algebras was a main motivation for introducing the filter dimension, [4]. In this chapter $d$ stands for the filter dimension $fd$ or the left filter dimension $lfd$. The following two inequalities are central for the proofs of almost all results in this chapter.

**The first filter inequality** [4]. Let $A$ be a simple finitely generated algebra. Then

$$\text{GK}(M) \geq \frac{\text{GK}(A)}{d(A) + \max\{d(A), 1\}}$$

for all non-zero finitely generated $A$-modules $M$. 

THE SECOND FILTER INEQUALITY [5]. Under certain mild conditions (Theorem 4.2) the (left) Krull dimension of the algebra $A$ satisfies the following inequality

$$\text{K.dim}(A) \leq \text{GK}(A) \left(1 - \frac{1}{\text{d}(A) + \max\{\text{d}(A), 1\}}\right).$$

The chapter is organized as follows. Both filter dimensions are introduced in Section 2. In Sections 3 and 4 the first and the second filter inequalities are proved respectively. In Section 4 we use both filter inequalities for giving short proofs of some classical results about the rings $D(X)$ of differential operators on smooth irreducible affine algebraic varieties. The (left) filter dimension of $D(X)$ is 1 (Section 5). A concept of multiplicity for the filter dimension and a concept of holonomic module for (simple) finitely generated algebras appear in Section 6. Every holonomic module has finite length (Theorem 6.8). In Section 7 an upper bound is given (i) for the Gelfand–Kirillov dimension of commutative subalgebras of simple finitely generated infinite-dimensional algebras (Theorem 7.2), and (ii) for the transcendence degree of subfields of quotient rings of (certain) simple finitely generated infinite-dimensional algebras (Theorems 7.4 and 7.5). In Section 8 a similar upper bound is obtained for the Gelfand–Kirillov dimension of isotropic subalgebras of strongly simple Poisson algebras (Theorem 8.1).

2. Filter dimension of algebras and modules

In this section, the filter dimension of algebras and modules will be defined.

2.1. The Gelfand–Kirillov dimension

Let $\mathcal{F}$ be the set of all functions from the set of natural numbers $\mathbb{N} = \{0, 1, \ldots\}$ to itself. For each function $f \in \mathcal{F}$, the non-negative real number or $\infty$ defined as

$$\gamma(f) := \inf\{r \in \mathbb{R} \mid f(i) \leq i^r \text{ for } i \gg 0\}$$

is called the degree of $f$. The function $f$ has polynomial growth if $\gamma(f) < \infty$. Let $f, g, p \in \mathcal{F}$, and $p(i) = p^*(i)$ for $i \gg 0$ where $p^*(t) \in \mathbb{Q}[t]$ (the polynomial algebra with coefficients from the field of rational numbers). Then

$$\gamma(f + g) \leq \max\{\gamma(f), \gamma(g)\}, \quad \gamma(fg) \leq \gamma(f) + \gamma(g),$$

$$\gamma(p) = \deg(p^*(t)), \quad \gamma(pg) = \gamma(p) + \gamma(g).$$

Let $A = K\langle a_1, \ldots, a_s \rangle$ be a finitely generated $K$-algebra. The finite-dimensional filtration $F = \{A_i\}$ associated with algebra generators $a_1, \ldots, a_s$:

$$A_0 := K \subseteq A_1 := K + \sum_{i=1}^{s} Ka_i \subseteq \cdots \subseteq A_i := A_i \subseteq \cdots$$
is called the standard filtration for the algebra $A$. Let $M = AM_0$ be a finitely generated $A$-module where $M_0$ is a finite-dimensional generating subspace. The finite-dimensional filtration $[M_i := A_i M_0]$ is called the standard filtration for the $A$-module $M$.

**Definition.** $\text{GK}(A) := \gamma(i \mapsto \dim_K(A_i))$ and $\text{GK}(M) := \gamma(i \mapsto \dim_K(M_i))$ are called the Gelfand–Kirillov dimensions of the algebra $A$ and the $A$-module $M$ respectively.

It is easy to prove that the Gelfand–Kirillov dimension of the algebra (resp. the module) does not depend on the choice of the standard filtration of the algebra (resp. and the choice of the generating subspace of the module).

### 2.2. The return functions and the (left) filter dimension

**Definition [4].** The function $\nu_{F, M_0} : \mathbb{N} \to \mathbb{N} \cup \{\infty\}$,

$$\nu_{F, M_0}(i) := \min \{ j \in \mathbb{N} \cup \{\infty\} : A_j M_{i, \text{gen}} \supseteq M_0 \text{ for all } M_{i, \text{gen}} \}$$

is called the return function of the $A$-module $M$ associated with the filtration $F = \{A_i\}$ of the algebra $A$ and the generating subspace $M_0$ of the $A$-module $M$ where $M_{i, \text{gen}}$ runs through all generating subspaces for the $A$-module $M$ such that $M_{i, \text{gen}} \subseteq M_i$.

Suppose, in addition, that the finitely generated algebra $A$ is a simple algebra. The return function $\nu_F \in \mathcal{F}$ and the left return function $\lambda_F \in \mathcal{F}$ for the algebra $A$ with respect to the standard filtration $F := \{A_i\}$ for the algebra $A$ are defined by the rules:

$$\nu_F(i) := \min \{ j \in \mathbb{N} \cup \{\infty\} : 1 \in A_j a A_j \text{ for all } 0 \neq a \in A_i \},$$

$$\lambda_F(i) := \min \{ j \in \mathbb{N} \cup \{\infty\} : 1 \in A a A_j \text{ for all } 0 \neq a \in A_i \},$$

where $A_j a A_j$ is the vector subspace of the algebra $A$ spanned over the field $K$ by the elements $x a y$ for all $x, y \in A_j$; and $A a A_j$ is the left ideal of the algebra $A$ generated by the set $a A_j$. The next result shows that under a mild restriction the maps $\nu_F(i)$ and $\lambda_F(i)$ are finite.

Recall that the centre of a simple algebra is a field.

**Lemma 2.1.** Let $A$ be a simple finitely generated algebra such that its centre $Z(A)$ is an algebraic field extension of $K$. Then $\lambda_F(i) \leq \nu_F(i) < \infty$ for all $i \geq 0$.

**Proof.** The first inequality is evident.

The centre $Z = Z(A)$ of the simple algebra $A$ is a field that contains $K$. Let $\{ \omega_j \mid j \in J \}$ be a $K$-basis for the $K$-vector space $Z$. Since $\dim_K(A_i) < \infty$, one can find finitely many $Z$-linearly independent elements, say $a_1, \ldots, a_s$, of $A_i$ such that $A_i \subseteq Z a_1 + \cdots + Z a_s$. Next, one can find a finite subset, say $J'$, of $J$ such that $A_i \subseteq V a_1 + \cdots + V a_s$ where $V = \sum_{j \in J'} K \omega_j$. The field $K'$ generated over $K$ by the elements $\omega_j$, $j \in J'$, is a finite field extension of $K$ (i.e. $\dim_K(K') < \infty$) since $Z/K$ is algebraic, hence $K' \subseteq A_n$ for some $n \geq 0$. Clearly, $A_i \subseteq K'a_1 + \cdots + K'a_s$. 
The $A$-bimodule $AA$ is simple with ring of endomorphisms $\text{End}(AA) \cong Z$. By the Density theorem, [21, 12.2], for each integer $1 \leq j \leq s$, there exist elements of the algebra $A$, say $x_1^j, \ldots, x_m^j, y_1^j, \ldots, y_m^j$, $m = m(j)$, such that for all $1 \leq l \leq s$

$$\sum_{k=1}^{m} x_k^j a_j y_l^j = \delta_{j,l}, \quad \text{the Kronecker delta.}$$

Let us fix a natural number, say $d = d_i$, such that $A_d$ contains all the elements $x_k^j$, $y_l^j$, and the field $K'$. We claim that $v_F(i) \leq 2d$. Let $0 \neq a \in A_i$. Then $a = \lambda_1 a_1 + \cdots + \lambda_s a_s$ for some $\lambda_i \in K'$. There exists $\lambda_j \neq 0$. Then $\sum_{k=1}^{m} \lambda_j^{-1} x_k^j a_j y_l^j = 1$, and $\lambda_j^{-1} x_k^j, y_l^j \in A_{2d}$.

This proves the claim and the lemma. \hfill \Box

REMARK. If the field $K$ is uncountable then automatically the centre $Z(A)$ of a simple finitely generated algebra $A$ is algebraic over $K$ (since $A$ has a countable $K$-basis and the rational function field $K(x)$ has uncountable basis over $K$ since the elements $\frac{1}{x^k}$, $\lambda \in K$, are $K$-linearly independent).

It is easy to see that for a finitely generated algebra $A$ any two standard finite-dimensional filtrations $F = \{A_i\}$ and $G = \{B_i\}$ are equivalent, $(F \sim G)$, that is, there exist natural numbers $a, b, c, d$ such that

$$A_i \subseteq B_{ai+b} \quad \text{and} \quad B_i \subseteq A_{ci+d} \quad \text{for} \ i \gg 0.$$  

If one of the inclusions holds, say the first, we write $F \leq G$.

**LEMMA 2.2.** Let $A$ be a finitely generated algebra equipped with two standard finite-dimensional filtrations $F = \{A_i\}$ and $G = \{B_i\}$.

1. Let $M$ be a finitely generated $A$-module. Then $\gamma(v_F, M_0) = \gamma(v_G, N_0)$ for any finite-dimensional generating subspaces $M_0$ and $N_0$ of the $A$-module $M$.

2. If, in addition, $A$ is a simple algebra then $\gamma(v_F) = \gamma(v_G)$ and $\gamma(\lambda_F) = \gamma(\lambda_G)$.

**PROOF.** (1) The module $M$ has two standard finite-dimensional filtrations $\{M_i = A_iM_0\}$ and $\{N_i = B_iN_0\}$. Let $v = v_F, M_0$ and $\mu = v_G, N_0$.

Suppose that $F = G$. Choose a natural number $s$ such that $M_0 \subseteq N_s$ and $N_0 \subseteq M_s$, so $N_i \subseteq M_{i+s}$ and $M_i \subseteq N_{i+s}$ for all $i \geq 0$. Let $N_{i,\text{gen}}$ be any generating subspace for the $A$-module $M$ such that $N_{i,\text{gen}} \subseteq N_i$. Since $M_0 \subseteq A_{v(i+s)}N_{i,\text{gen}}$ for all $i \geq 0$ and $N_0 \subseteq A_{v(i+s)}N_{i,\text{gen}}$, hence, $\mu(i) \geq v(i) + s$ and finally $\gamma(\mu) \geq \gamma(v)$. By symmetry, the opposite inequality is true and so $\gamma(\mu) = \gamma(v)$.

Suppose that $M_0 = N_0$. The algebra $A$ is a finitely generated algebra, so all standard finite-dimensional filtrations of the algebra $A$ are equivalent. In particular, $F \sim G$ and so one can choose natural numbers $a, b, c, d$ such that

$$A_i \subseteq B_{ai+b} \quad \text{and} \quad B_i \subseteq A_{ci+d} \quad \text{for} \ i \gg 0.$$  

Then $N_i = B_iN_0 \subseteq A_{ci+d}M_0 = M_{ci+d}$ for all $i \geq 0$, hence $N_0 = M_0 \subseteq A_{v(ci+d)}N_{i,\text{gen}} \subseteq B_{av(ci+d)+b}N_{i,\text{gen}}$, therefore $\mu(i) \leq av(ci + d) + b$ for all $i \geq 0$, hence $\gamma(\mu) \leq \gamma(v)$. By
symmetry, we get the opposite inequality which implies \( \gamma(\mu) = \gamma(\nu) \). Now, \( \gamma(v_{F,M_0}) = \gamma(v_{F,N_0}) = \gamma(v_{G,N_0}) \).

(2) The algebra \( A \) is simple, equivalently, it is a simple (left) \( A \otimes A^0 \)-module where \( A^0 \) is the opposite algebra to \( A \). The opposite algebra has the standard filtration \( F^0 = \{ A^0_i \} \), opposite to the filtration \( F \). The tensor product of algebras \( A \otimes A^0 \), so-called, the enveloping algebra of \( A \otimes A^0 \), has the standard filtration \( F \otimes F^0 = \{ C_n \} \) which is the tensor product of the standard filtrations \( F \) and \( F^0 \), that is, \( C_n = \sum \{ A_i \otimes A_j^0, i + j \leq n \} \). Let \( v_{F \otimes F^0, K} \) be the return function of the \( A \otimes A^0 \)-module \( A \) associated with the filtration \( F \otimes F^0 \) and the generating subspace \( K \). Then

\[
v_F(i) \leq v_{F \otimes F^0, K}(i) \leq 2v_F(i) \quad \text{for all } i \geq 0,
\]

and so

\[
\gamma(v_F) = \gamma(v_{F \otimes F^0, K}), \quad (1)
\]

and, by the first statement, we have \( \gamma(v_F) = \gamma(v_{F \otimes F^0, K}) = \gamma(v_{G \otimes G^0, K}) = \gamma(v_G) \), as required. Using a similar argument as in the proof of the first statement one can prove that \( \gamma(\lambda_F) = \gamma(\lambda_G) \). We leave this as an exercise. \( \square \)

**Definition [4].** \( \text{fd}(M) = \gamma(v_{F,M_0}) \) is the filter dimension of the \( A \)-module \( M \), and \( \text{fd}(A) := \text{fd}(A \otimes A^0) \) is the filter dimension of the algebra \( A \). If, in addition, the algebra \( A \) is simple, then \( \text{fd}(A) = \gamma(v_F) \), and \( \text{lfd}(A) := \gamma(\lambda_F) \) is called the left filter dimension of the algebra \( A \).

By the previous lemma the definitions make sense (both filter dimensions do not depend on the choice of the standard filtration \( F \) for the algebra \( A \)).

By Lemma 2.1, \( \text{lfd}(A) \leq \text{fd}(A) \).

**Question.** What is the filter dimension of a polynomial algebra?

### 3. The first filter inequality

In this chapter, \( d(A) \) means either the filter dimension \( \text{fd}(A) \) or the left filter dimension \( \text{lfd}(A) \) of a simple finitely generated algebra \( A \) (i.e. \( d = \text{fd} \), \( \text{lfd} \)). Both filter dimensions appear naturally when one tries to find a lower bound for the holonomic number (Theorem 3.1) and an upper bound (Theorem 4.2) for the (left and right) Krull dimension (in the sense of Rentzschler and Gabriel, [22]) of simple finitely generated algebras.

The next theorem is a generalization of the Bernstein’s inequality (Theorem 1.1) to the class of simple finitely generated algebras.

**Theorem 3.1 (The first filter inequality, [4,6]).** Let \( A \) be a simple finitely generated algebra. Then

\[
\text{GK}(M) \geq \frac{\text{GK}(A)}{d(A) + \max\{d(A), 1\}}
\]

for all non-zero finitely generated \( A \)-modules \( M \) where \( d = \text{fd} \), \( \text{lfd} \).
Let \( \lambda = \lambda_F \) be the left return function associated with a standard filtration \( F \) of the algebra \( A \) and let \( 0 \neq a \in A_i \). It suffices to prove the inequality for \( \lambda \) (since \( \text{fd}(A) \geq \text{lfld}(A) \)). It follows from the inclusion

\[
Aa M_{\lambda(i)} = Aa A_{\lambda(i)} M_0 \supseteq 1 M_0 = M_0
\]

that the linear map

\[
A_i \to \text{Hom}(M_{\lambda(i)}, M_{\lambda(i)+i}), \quad a \mapsto (m \mapsto am),
\]

is injective, so \( \dim A_i \leq \dim M_{\lambda(i)} \dim M_{\lambda(i)+i} \). Using the above elementary properties of the degree (see also [19, 8.1.7]), we have

\[
\text{GK}(A) = \gamma(\dim A_i) \leq \gamma(\dim M_{\lambda(i)}) + \gamma(\dim M_{\lambda(i)+i}) \\
\leq \gamma(\dim M_i) \gamma(\lambda) + \gamma(\dim M_i) \max \{ \gamma(\lambda), 1 \} \\
= \text{GK}(M) (\text{lfld}A + \max \{ \text{lfld}A, 1 \}) \\
\leq \text{GK}(M) (\text{lfld}A + \max \{ \text{lfld}A, 1 \}).
\]

The result above gives a lower bound for the holonomic number of a simple finitely generated algebra

\[
h_A \geq \frac{\text{GK}(A)}{d(A) + \max \{ d(A), 1 \}}.
\]

**Theorem 3.2.** Let \( A \) be a finitely generated algebra. Then

\[
\text{GK}(M) \leq \text{GK}(A) \text{fd}(M)
\]

for any simple \( A \)-module \( M \).

**Proof.** Let \( v = v_{F,K,m} \) be the return function of the module \( M \) associated with a standard finite-dimensional filtration \( F = \{ A_i \} \) of the algebra \( A \) and a fixed non-zero element \( m \in M \). Let \( \pi : M \to K \) be a non-zero linear map satisfying \( \pi(m) = 1 \). Then, for any \( i \geq 0 \) and any \( 0 \neq u \in M_i : 1 = \pi(m) \in \pi(A_{v(i)}u) \), and so the linear map

\[
M_i \to \text{Hom}(A_{v(i)}, K), \quad u \mapsto (a \mapsto \pi(au)),
\]

is an injective map hence \( \dim M_i \leq \dim A_{v(i)} \) and finally \( \text{GK}(M) \leq \text{GK}(A) \text{fd}(M) \). \( \square \)

**Corollary 3.3.** Let \( A \) be a simple finitely generated infinite-dimensional algebra. Then

\[
\text{fd}(A) \geq \frac{1}{2}.
\]
PROOF. The algebra $A$ is a finitely generated infinite-dimensional algebra hence $\text{GK}(A) > 0$. Clearly, $\text{GK}(A \otimes A^0) \leq \text{GK}(A) + \text{GK}(A^0) = 2\text{GK}(A)$. Applying Theorem 3.2 to the simple $A \otimes A^0$-module $M = A$ we finish the proof:

$$\text{GK}(A) = \text{GK}(A \otimes A^0) \leq \text{GK}(A \otimes A^0) \leq 2\text{GK}(A) \text{fd}(A)$$

hence $\text{fd}(A) \geq \frac{1}{2}$.

□

QUESTION. Is $\text{fd}(A) \geq 1$ for all simple finitely generated infinite-dimensional algebras $A$?

QUESTION. For which numbers $d \geq \frac{1}{2}$ there exists a simple finitely generated infinite-dimensional algebra $A$ with $\text{fd}(A) = d$?

COROLLARY 3.4. Let $A$ be a simple finitely generated infinite-dimensional algebra. Then

$$\text{fd}(M) \geq \frac{1}{\text{fd}(A) + \max\{\text{fd}(A), 1\}}$$

for all simple $A$-modules $M$.

PROOF. Applying Theorem 3.1 and Theorem 3.2, we have the result

$$\text{fd}(M) \geq \frac{\text{GK}(M)}{\text{GK}(A)} \geq \frac{\text{GK}(A)}{\text{GK}(A)(\text{fd}(A) + \max\{\text{fd}(A), 1\})} = \frac{1}{\text{fd}(A) + \max\{\text{fd}(A), 1\}}.$$ □

4. Krull, Gelfand–Kirillov and filter dimensions of simple finitely generated algebras

In this section, we prove the second filter inequality (Theorem 4.2) and apply both filter inequalities for giving short proofs of some classical results about the rings of differential operators on a smooth irreducible affine algebraic varieties (Theorems 1.1, 4.4, 4.5, 4.7).

We say that an algebra $A$ is (left) finitely partitive, [19, 8.3.17], if, given any finitely generated $A$-module $M$, there is an integer $n = n(M) > 0$ such that for every strictly descending chain of $A$-submodules of $M$:

$$M = M_0 \supset M_1 \supset \cdots \supset M_m$$

with $\text{GK}(M_i/M_{i+1}) = \text{GK}(M)$, one has $m \leq n$. McConnell and Robson write in their book [19, 8.3.17], that “yet no examples are known which fail to have this property.”

Recall that $\text{K.dim}$ denotes the (left) Krull dimension in the sense of Rentschler and Gabriel, [22].
**Lemma 4.1.** Let $A$ be a finitely partitive algebra with $\text{GK}(A) < \infty$. Let $a \in \mathbb{N}$, $b \geq 0$ and suppose that $\text{GK}(M) \geq a + b$ for all finitely generated $A$-modules $M$ with $\text{K.dim}(M) = a$, and that $\text{GK}(N) \in \mathbb{N}$ for all finitely generated $A$-modules $N$ with $\text{K.dim}(N) \geq a$. Then $\text{GK}(M) \geq \text{K.dim}(M) + b$ for all finitely generated $A$-modules $M$ with $\text{K.dim}(M) \geq a$. In particular, $\text{GK}(A) \geq \text{K.dim}(A) + b$.

**Remark.** It is assumed that a module $M$ with $\text{K.dim}(M) = a$ exists.

**Proof.** We use induction on $n = \text{K.dim}(M)$. The base of induction, $n = a$, is true. Let $n > a$. There exists a descending chain of submodules $M = M_1 \supset M_2 \supset \cdots$ with $\text{K.dim}(M_i/M_{i+1}) = n - 1$ for $i \geq 1$. By induction, $\text{GK}(M_i/M_{i+1}) \geq n - 1 + b$ for $i \geq 1$.

The algebra $A$ is finitely partitive, so there exists $i$ such that $\text{GK}(M) > \text{GK}(M_i/M_{i+1})$, so $\text{GK}(M) - 1 \geq \text{GK}(M_i/M_{i+1}) \geq n - 1 + b$, since $\text{GK}(M) \in \mathbb{N}$, hence $\text{GK}(M) \geq \text{K.dim}(M) + b$. Since $\text{K.dim}(A) \geq \text{K.dim}(M)$ for all finitely generated $A$-modules $M$ we have $\text{GK}(A) \geq \text{K.dim}(A) + b$. □

**Theorem 4.2 [5].** Let $A$ be a simple finitely generated finitely partitive algebra with $\text{GK}(A) < \infty$. Suppose that the Gelfand–Kirillov dimension of every finitely generated $A$-module is a natural number. Then

$$
\text{K.dim}(M) \leq \text{GK}(M) - \frac{\text{GK}(A)}{\text{d}(A) + \max\{\text{d}(A), 1\}}
$$

for any non-zero finitely generated $A$-module $M$. In particular,

$$
\text{K.dim}(A) \leq \text{GK}(A) \left(1 - \frac{1}{\text{d}(A) + \max\{\text{d}(A), 1\}}\right).
$$

**Proof.** Applying the lemma above to the family of finitely generated $A$-modules of Krull dimension 0, by Theorem 3.1, we can put $a = 0$ and

$$
b = \frac{\text{GK}(A)}{\text{d}(A) + \max\{\text{d}(A), 1\}},
$$

and the result follows. □

Let $K$ be a field of characteristic zero and $B$ be a commutative $K$-algebra. The ring of ($K$-linear) differential operators $\mathcal{D}(B)$ on $B$ is defined as $\mathcal{D}(B) = \bigcup_{i=0}^{\infty} \mathcal{D}_i(B)$ where $\mathcal{D}_0(B) = \text{End}_K(B) \cong B ((x \mapsto bx) \leftrightarrow b)$,

$$
\mathcal{D}_i(B) = \{ u \in \text{End}_K(B): [u, r] \in \mathcal{D}_{i-1}(B) \text{ for each } r \in B \}.
$$

Note that the $\{\mathcal{D}_i(B)\}$ is, so-called, the order filtration for the algebra $\mathcal{D}(B)$:

$$
\mathcal{D}_0(B) \subseteq \mathcal{D}_1(B) \subseteq \cdots \subseteq \mathcal{D}_i(B) \subseteq \cdots \quad \text{and} \quad \mathcal{D}_i(B) \mathcal{D}_j(B) \subseteq \mathcal{D}_{i+j}(B),
$$

$i, j \geq 0$. 

The subalgebra $\Delta(B)$ of $\text{End}_K(B)$ generated by $B \equiv \text{End}_K(B)$ and by the set $\text{Der}_K(B)$ of all $K$-derivations of $B$ is called the derivation ring of $B$. The derivation ring $\Delta(B)$ is a subring of $\mathcal{D}(B)$.

Let the finitely generated algebra $B$ be a regular commutative domain of Krull dimension $n < \infty$. In geometric terms, $B$ is the coordinate ring $\mathcal{O}(X)$ of a smooth irreducible affine algebraic variety $X$ of dimension $n$. Then

- $\text{Der}_K(B)$ is a finitely generated projective $B$-module of rank $n$;
- $\mathcal{D}(B) = \Delta(B)$;
- $\mathcal{D}(B)$ is a simple (left and right) Noetherian domain with $\text{GK}\mathcal{D}(B) = 2n$ ($n = \text{GK}(B) = \text{Kdim}(B)$);
- $\mathcal{D}(B) = \Delta(B)$ is an almost centralizing extension of $B$;
- the associated graded ring $\text{gr}\mathcal{D}(B) = \bigoplus \mathcal{D}_i(B)/\mathcal{D}_{i-1}(B)$ is a commutative domain;
- the Gelfand–Kirillov dimension of every finitely generated $\mathcal{D}(B)$-module is a natural number.

For the proofs of the statements above the reader is referred to [19, chapter 15]. So, the domain $\mathcal{D}(B)$ is a simple finitely generated infinite-dimensional Noetherian algebra, [19, chapter 15].

**Example.** Let $P_n = K[X_1, \ldots, X_n]$ be a polynomial algebra. $\text{Der}_K(P_n) = \bigoplus_{i=1}^n P_n \frac{\partial}{\partial X_i}$,

$$\mathcal{D}(P_n) = \Delta(P_n) = K \left[ X_1, \ldots, X_n, \frac{\partial}{\partial X_1}, \ldots, \frac{\partial}{\partial X_n} \right]$$

is the ring of differential operators with polynomial coefficients, i.e. the $n$-th Weyl algebra $A_n$.

In Section 5, we prove the following result.

**Theorem 4.3** [5]. The filter dimension and the left filter dimension of the ring of differential operators $\mathcal{D}(B)$ are both equal to 1.

As an application we compute the Krull dimension of $\mathcal{D}(B)$.

**Theorem 4.4** [19, Chapter 15].

$$\text{Kdim}\mathcal{D}(B) = \frac{\text{GK}(\mathcal{D}(B))}{2} = \text{Kdim}(B).$$

**Proof.** The second equality is clear: $(\text{GK}(\mathcal{D}(B)) = 2\text{GK}(B) = 2\text{Kdim}(B))$. It follows from Theorems 4.2 and 4.3 that

$$\text{Kdim}\mathcal{D}(B) \leq \frac{\text{GK}(\mathcal{D}(B))}{2} = \text{Kdim}(B).$$

The map $I \to \mathcal{D}(B)I$ from the set of left ideals of $B$ to the set of left ideals of $\mathcal{D}(B)$ is injective, thus $\text{Kdim}(B) \leq \text{Kdim}(\mathcal{D}(B))$. \qed
This result shows that for the ring of differential operators on a smooth irreducible affine algebraic variety the inequality in Theorem 4.2 is an equality.

**THEOREM 4.5** [19, 15.4.3]. Let $M$ be a non-zero finitely generated $\mathcal{D}(B)$-module. Then

$$\text{GK}(M) \geq \frac{\text{GK}(\mathcal{D}(B))}{2} = \text{K.dim}(B).$$

**PROOF.** By Theorems 3.1 and 4.3,

$$\text{GK}(M) \geq \frac{\text{GK}(\mathcal{D}(B))}{2} = \frac{2\text{GK}(B)}{2} = \text{GK}(B) = \text{K.dim}(B).$$

So, for the ring of differential operators on a smooth affine algebraic variety the inequality in Theorem 3.1 is in fact an equality.

In general, it is difficult to find the exact value for the filter dimension but for the Weyl algebra $A_n$ it is easy and one can find it directly.

**THEOREM 4.6.** Both the filter dimension and the left filter dimension of the Weyl algebra $A_n$ over a field of characteristic zero are equal to 1.

**PROOF.** Denote by $a_1, \ldots, a_{2n}$ the canonical generators of the Weyl algebra $A_n$ and denote by $F = \{A_{n,i} \mid i \geq 0\}$ the standard filtration associated with the canonical generators. The associated graded algebra $\text{gr} A_n := \bigoplus_{i \geq 0} A_{n,i}/A_{n,i-1}$ ($A_{n,-1} = 0$) is a polynomial algebra in $2n$ variables, so

$$\text{GK}(A_n) = \text{GK}(\text{gr} A_n) = 2n.$$  

For every $i \geq 0$:

$$\text{ad} a_j : A_{n,i} \to A_{n,i-1}, \ x \mapsto \text{ad} a_j(x) := a_j x - x a_j.$$

The algebra $A_n$ is central ($Z(A_n) = K$), so

$$\text{ad} a_j(x) = 0 \text{ for all } j = 1, \ldots, 2n \iff x \in Z(A_n) = K = A_{n,0}.$$  

These two facts imply $v_F(i) \leq i$ for $i \geq 0$, and so $d(A_n) \leq 1$.

The $A_n$-module $P_n := K[X_1, \ldots, X_n] \simeq A_n/(A_n \partial_1 + \cdots + A_n \partial_n)$ has Gelfand–Kirillov dimension $n$. By Theorem 3.1 applied to the $A_n$-module $P_n$, we have

$$2n = \text{GK}(A_n) \leq n \left(d(A) + \max\{d(A), 1\}\right),$$

hence $d(A_n) \geq 1$, and so $d(A_n) = 1$.  

**PROOF OF THE BERNSTEIN’S INEQUALITY (THEOREM 1.1).** Since $\text{GK}(A_n) = 2n$ and $d(A_n) = 1$, Theorem 3.1 gives $\text{GK}(M) \geq \frac{2n}{2} = n$.  

One also gets a short proof of the following result of Rentschler and Gabriel.

**Theorem 4.7 [22].** If $\text{char } K = 0$ then the Krull dimension of the Weyl algebra $A_n$ is

$$K \text{.dim}(A_n) = n.$$

**Proof.** Putting $G(K(A_n)) = 2n$ and $d(A_n) = 1$ into the second formula of Theorem 4.2 we have $K \text{.dim}(A_n) \leq \frac{2d}{d} = n$. The polynomial algebra $P_n = K[X_1, \ldots, X_n]$ is the subalgebra of $A_n$ such that $A_n$ is a free right $P_n$-module. The map $I \mapsto A_nI$ from the set of left ideals of the polynomial algebra $P_n$ to the set of left ideals of the Weyl algebra $A_n$ is injective, thus $n = K \text{.dim}(P_n) \leq K \text{.dim}(A_n)$, and so $K \text{.dim}(A_n) = n$. □

### 5. Filter dimension of the ring of differential operators on a smooth irreducible affine algebraic variety (proof of Theorem 4.3)

Let $K$ be a field of characteristic 0 and let the algebra $B$ be as in the previous section, i.e. $B$ is a finitely generated regular commutative algebra which is a domain. We keep the notations of the previous section. Recall that the derivation ring $\Delta = \Delta(B)$ coincides with the ring of differential operators $D(B)$, [19, 15.5.6], and is a simple finitely generated finitely partitive $K$-algebra, [19, 15.3.8, 15.1.21]. We refer the reader to [19, Chapter 15] for basic definitions. We aim to prove Theorem 4.3.

Let $\{B_i\}$ and $\{\Delta_i\}$ be standard finite-dimensional filtrations on $B$ and $\Delta$ respectively such that $B_i \subseteq \Delta_i$ for all $i \geq 0$. Then the enveloping algebra $\Delta^e := \Delta \otimes \Delta^0$ can be equipped with the standard finite-dimensional filtration $\{\Delta^e_i\}$ which is the tensor product of the filtrations $\{\Delta_i\}$ and $\{\Delta^0_i\}$ of the algebras $\Delta$ and $\Delta^0$ respectively.

Then $B \simeq \Delta/\Delta \text{Der}_K B$ is a simple left $\Delta$-module [19, 15.3.8] with $G(K(\Delta)) = 2G(K(B))$, [19, 15.3.2]. By Theorem 3.1,

$$d(\Delta) + \max\{d(\Delta), 1\} \geq \frac{G(K(\Delta))}{G(K(B))} = \frac{2G(K(B))}{G(K(B))} = 2,$$

hence $d(\Delta) \geq 1$. It remains to prove the opposite inequality. For this, we recall some properties of $\Delta$ (see [19, Chapter 15], for details).

Given $0 \neq c \in B$, denote by $B_c$ the localization of the algebra $B$ at the powers of the element $c$, then $\Delta(B_c) \simeq \Delta(B)_c$ and the map $\Delta(B) \rightarrow \Delta(B)_c$, $d \rightarrow d/1$, is injective, [19, 5.1.25]. There is a finite subset $\{c_1, \ldots, c_t\}$ of $B$ such that the algebra $\bigcap_{i=1}^t \Delta(B_{c_i})$ is left and right faithfully flat over its subalgebra $\Delta$.

$$\sum_{i=1}^t B_{c_i} = B \quad (\text{see the proof of 15.2.13, [19]).}$$
For each $c = c_i$, $\text{Der}_K(B_c)$ is a free $B_c$-module with a basis $\partial_j = \frac{\partial}{\partial x_j}$, $j = 1, \ldots, n$, for some $x_1, \ldots, x_n \in B$, [19, 15.2.13]. Note that the choice of the $x_j$-th depends on the choice of the $c_i$. Then

$$\Delta(B)_c \simeq \Delta(B_c) = B_c(\partial_1, \ldots, \partial_n) \supseteq K(x_1, \ldots, x_n, \partial_1, \ldots, \partial_n).$$

Fix $c = c_i$. We aim to prove the following statement:

- There exist natural numbers $a$, $b$, $\alpha$, and $\beta$, such that for any $0 \neq d \in \Delta_k$ there exists $w \in \Delta^e_{\alpha k + b}$: $wd = c^\alpha k + \beta$. \((\ast)\)

Suppose that we are done. Then one can choose the numbers $a$, $b$, $\alpha$, and $\beta$ such that \((\ast)\) holds for all $i = 1, \ldots, t$. It follows from (2) that

$$\sum_{i=1}^t f_i c_i \in \Delta_v, \text{ and set } N(k) = \alpha k + \beta,$$  

where the $w_i$ are from \((\ast)\), i.e. $w_i \in \Delta^e_{\alpha k + b}$, $w_i d = c^\alpha_i N(k)$. So, $w = \sum_{i=1}^t g_i w_i \in \Delta^e_{\alpha k + \beta}$ and so $d(\Delta) \leq 1$, as required.

Fix $c = c_i$. By [19, 15.1.24], $\text{Der}_K(B_c) \simeq \text{Der}_K(B)_c$ and $\text{Der}_K B$ can be seen as a finitely generated $B$-submodule of $\text{Der}_K(B_c)$, [19, 15.1.7].

The algebra $B$ contains the polynomial subalgebra $P = K[x_1, \ldots, x_n]$. The polynomial algebra $P$ has the natural filtration $P = \bigcup_{i \geq 0} P_i$ by the total degree of the variables. Fix a natural number $l$ such that $P_i \subseteq B_l$, then $P_i \subseteq B_l$ for all $i \geq 0$. We denote by $Q = K(x_1, \ldots, x_n)$ the field of fractions of $P$. The field of fractions, say $L$, of the algebra $B$ has the same transcendence degree $n$ as the field of rational functions $Q$. The algebra $B$ is a finitely generated algebra, hence $L$ is a finite field extension of $Q$ of dimension, say $m$, over $Q$. Let $e_1, \ldots, e_m \in B$ be a $Q$-basis for the vector space $L$ over $Q$. Note that $L = QB$.

One can find a natural number $\beta \geq 1$ and a non-zero polynomial $p \in P_\beta$ such that

$$\{B_1, e_je_k \mid j, k = 1, \ldots, m\} \subseteq \sum_{\alpha=1}^m p^{-1} P_\beta e_\alpha.$$  

Then $B_k \subseteq \sum_{j=1}^m p^{-2k} P_{2\beta} e_j$ and $B_k e_i \subseteq \sum_{j=1}^m p^{-3k} P_{3\beta} e_j$ for all $k \geq 1$ and $i = 1, \ldots, m$. Let $0 \neq d \in B_k$. The $m \times m$ matrix of the bijective $Q$-linear map $L \to L$, $x \mapsto dx$, with respect to the basis $e_1, \ldots, e_m$ has entries from the set $p^{-3k} P_{3\beta}$. So, its characteristic polynomial

$$\chi_d(t) = t^m + \alpha_{m-1} t^{m-1} + \cdots + \alpha_0.$$
has coefficients in $p^{-3mk} P_{3m\beta k}$, and $\alpha_0 \neq 0$ as $x \mapsto dx$ is a bijection. Now,

$$
P_{6m\beta k} \ni p^{-3mk} \alpha_0 = \sum_{m=-1}^{m-2} \cdots \alpha_1 d \\
\in B_{4m\beta k} P_{3m\beta k} d \subseteq B_{m\beta k}(4+3l) d. \tag{3}
$$

Let $\delta_1, \ldots, \delta_t$ be a set of generators for the left $B$-module $\text{Der}_K(B)$. Then

$$
\partial_i \in \sum_{j=1}^t c_{-l_i} B_{l_1} \delta_j \quad \text{for } i = 1, \ldots, n,
$$

for some natural numbers $l_i' \leq l_1$. Fix a natural number $l_2$ such that $\delta_j(B_1) \subseteq B_{l_2}$ and $\delta_j(c) \in B_{l_2}$ for $j = 1, \ldots, t$. Then

$$
\partial^\alpha(B_k) \subseteq c^{-|\alpha| l_1'} B_{k+3|\alpha| (l_1'+l_2)} \quad \text{for all } \alpha \in \mathbb{N}^n, k \geq 1,
$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$, $\partial^\alpha = \partial_{\alpha_1} \cdots \partial_{\alpha_n}$. It follows from (3) that one can find $\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq 6m\beta k$ and

$$
1 \in K^* \partial^\alpha \left( p^{-3mk} \alpha_0 \right) \subseteq \partial^\alpha \left( B_{m\beta k}(4+3l) d \right) \\
\subseteq c^{-pk} \Delta^l_{qk+r} d,
$$

where $p, q, r \in \mathbb{N}$ and $K^* = K \setminus \{0\}$. Now $(*)$ follows.

In fact we have proved the following corollary.

**Corollary 5.1.** There exist natural numbers $a$ and $b$ such that for any $0 \neq d \in \Delta_k$ there exists an element $w \in \Delta^l_{ak+b}$ satisfying $wd = 1$.

6. **Multiplicity for the filter dimension, holonomic modules over simple finitely generated algebras**

In this section, we introduce a concept of multiplicity for the filter dimension and a concept of holonomic module for (some) finitely generated algebras. We will prove that a holonomic module has finite length (Theorem 6.8). The multiplicity for the filter dimension is a key ingredient in the proof.

First we recall the definition of multiplicity in the commutative situation and then for certain non-commutative algebras (somewhat commutative algebras).

### 6.1. Multiplicity in the commutative situation

Let $B$ be a commutative finitely generated $K$-algebra with a standard finite-dimensional filtration $F = \{B_i\}$, and let $M$ be a finitely generated $B$-module with a finite-dimensional
generating subspace, say $M_0$, and with the standard filtration $\{M_i = B_i M_0\}$ attached to it. Then there exists a polynomial $p(t) = tl^d + \cdots \in \mathbb{Q}[t]$ with rational coefficients of degree $d = \text{GK}(M)$ such that

$$\dim_K(M_i) = p(i) \quad \text{for all } i \gg 0.$$ 

The polynomial $p(t)$ is called the Hilbert polynomial of the $B$-module $M$. The Hilbert polynomial does depend on the filtration $\{M_i\}$ of the module $M$ but its leading coefficient $l$ does not. The number $e(M) = d!l$ is called the multiplicity of the $B$-module $M$. It is a natural number which does depend on the filtration $F$ of the algebra $B$.

In the case when $M = B$ is the homogeneous coordinate ring of a projective algebraic variety $X \subseteq \mathbb{P}^m$ equipped with the natural filtration that comes from the grading of the graded algebra $B$, the multiplicity is the degree of $X$, the number of points in which $X$ meets a generic plane of complementary degree in $\mathbb{P}^m$ ($K$ is an algebraically closed field).

**6.2. Somewhat commutative algebras**

A $K$-algebra $R$ is called a somewhat commutative algebra if it has a finite-dimensional filtration $R = \bigcup_{i \geq 0} R_i$ such that the associated graded algebra $\text{gr} R := \bigoplus_{i \geq 0} R_i / R_{i-1}$ is a commutative finitely generated $K$-algebra where $R_{-1} = 0$ and $R_0 = K$. Then the algebra $R$ is a Noetherian finitely generated algebra since $\text{gr} R$ is so. A finitely generated module over a somewhat commutative algebra has a Gelfand–Kirillov dimension which is a natural number. We refer the reader to the books [16,19] for the properties of somewhat commutative algebras.

**DEFINITION.** For a somewhat commutative algebra $R$ we define the holonomic number,

$$h_R := \min \{ \text{GK}(M) \mid M \neq 0 \text{ is a finitely generated } R\text{-module} \}.$$ 

**DEFINITION.** A finitely generated $R$-module $M$ is called a holonomic module if $\text{GK}(M) = h_R$. In other words, a non-zero finitely generated $R$-module is holonomic iff it has least Gelfand–Kirillov dimension. If $h_R = 0$ then every holonomic $R$-module is finite-dimensional and vice versa.

**EXAMPLES.** (1) The holonomic number of the Weyl algebra $A_n$ is $n$. The polynomial algebra $K[X_1, \ldots, X_n] \simeq A_n / \sum^n_{j=1} A_n \partial_j$ with the natural action of the ring of differential operators $A_n = K[X_1, \ldots, X_n, \partial X_1, \ldots, \partial X_n$] is a simple holonomic $A_n$-module.

(2) Let $X$ be a smooth irreducible affine algebraic variety of dimension $n$. The ring of differential operators $\mathcal{D}(X)$ is a simple somewhat commutative algebra of Gelfand–Kirillov dimension $2n$ with holonomic number $h_{\mathcal{D}(X)} = n$. The algebra $\mathcal{O}(X)$ of regular functions of the variety $X$ is a simple $\mathcal{D}(X)$-module with respect to the natural action of the algebra $\mathcal{D}(X)$. In more detail, $\mathcal{O}(X) \simeq \mathcal{D}(X)/\mathcal{D}(X) \text{Der}_K(\mathcal{O}(X))$ where $\text{Der}_K(\mathcal{O}(X))$ is the $\mathcal{O}(X)$-module of derivations of the algebra $\mathcal{O}(X)$. 
Let \( R = \bigcup_{i \geq 0} R_i \) be a somewhat commutative algebra. The associated graded algebra \( \text{gr} R \) is a commutative affine algebra. Let us choose homogeneous algebra generators of the algebra \( \text{gr} R \), say \( y_1, \ldots, y_s \), of graded degrees \( 1 \leq k_1, \ldots, k_s \) respectively (that is \( y_i \in R_{k_i}/R_{k_i-1} \)). A filtration \( \Gamma = \{ \Gamma_i \mid i \geq 0 \} \) of an \( R \)-module \( M = \bigcup_{i=0}^{\infty} \Gamma_i \) is called good if the associated graded \( \text{gr} R \)-module \( \text{gr} \Gamma M := \bigoplus_{i \geq 0} \Gamma_i/\Gamma_i-1 \) is finitely generated. An \( R \)-module \( M \) has a good filtration iff it is finitely generated, and if \( \{ \Gamma_i \} \) and \( \{ \Omega_i \} \) are two good filtrations of \( M \), then there exists a natural number \( t \) such that \( \Gamma_i \subseteq \Omega_i+t \) and \( \Omega_i \subseteq \Gamma_i+t \) for all \( i \).

Lemma 6.1. Let \( R = \bigcup_{i \geq 0} R_i \) be a somewhat commutative algebra, \( k = \text{lcm}(k_1, \ldots, k_s) \), and let \( M \) be a finitely generated \( R \)-module with good filtration \( \Gamma = \{ \Gamma_i \} \).

1. There exist \( k \) polynomials \( \gamma_0, \ldots, \gamma_{k-1} \in \mathbb{Q}[t] \) with coefficients from \( \lbrack k!^{\text{GK}(M)} \lbrack \text{GK}(M)! \rbrack^{-1} \mathbb{Z} \) such that \( \dim \Gamma_i = \gamma_j(i) \) for all \( i \gg 0 \) and \( j \equiv i \pmod{k} \).

2. The polynomials \( \gamma_j \) have the same degree \( \text{GK}(M) \) and the same leading coefficient \( e(M)/\text{GK}(M)! \) where \( e(M) \) is called the multiplicity of \( M \). The multiplicity \( e(M) \) does not depend on the choice of the good filtration \( \Gamma \).

Remark. A finitely generated \( R \)-module \( M \) has \( e(M) = 0 \) iff \( \text{dim}_K(M) < \infty \).

Lemma 6.2. Let \( 0 \to N \to M \to L \to 0 \) be an exact sequence of modules over a somewhat commutative algebra \( R \). Then \( \text{GK}(M) = \max\{\text{GK}(N), \text{GK}(L)\} \), and if \( \text{GK}(N) = \text{GK}(M) = \text{GK}(L) \) then \( e(M) = e(N) + e(L) \).

Corollary 6.3. Let the algebra \( R \) be as in lemma 6.1 with holonomic number \( h > 0 \).

1. Let \( M \) be a holonomic \( R \)-module with multiplicity \( e(M) \). The \( R \)-module \( M \) has finite length \( \leq e(M)k^h \).

2. Every non-zero submodule or factor module of a holonomic \( R \)-module is a holonomic module.

Proof. This follows directly from Lemma 6.2.

6.3. Multiplicity

Let \( f \) be a function from \( \mathbb{N} \) to \( \mathbb{R}_+ = \{ r \in \mathbb{R} : r \geq 0 \} \), the leading coefficient of \( f \) is a non-zero limit (if it exists)

\[
\text{lc}(f) = \lim_{i \to \infty} \frac{f(i)}{i^d} \neq 0,
\]

where

\[
\frac{f(i)}{i^d} \to \text{lc}(f) \text{ as } i \to \infty.
\]
where $d = \gamma(f)$. If $d \in \mathbb{N}$, we define the multiplicity $e(f)$ of $f$ by

$$e(f) = d! \, \text{lc}(f).$$

The factor $d!$ ensures that the multiplicity $e(f)$ is a positive integer in some important cases. If $f(t) = a_d t^d + a_{d-1} t^{d-1} + \cdots + a_0$ is a polynomial of degree $d$ with real coefficients then $\text{lc}(f) = a_d$ and $e(f) = d! a_d$.

**LEMMA 6.4.** Let $A$ be a finitely generated algebra equipped with a standard finite-dimensional filtration $F = \{A_i\}$ and $M$ be a finitely generated $A$-module with generating finite-dimensional subspaces $M_0$ and $N_0$.

1. If $\text{lc}(v_{F,M_0})$ exists then so does $\text{lc}(v_{F,N_0})$, and $\text{lc}(v_{F,M_0}) = \text{lc}(v_{F,N_0})$.
2. If $\text{lc}(\dim A_i M_0)$ exists then so does $\text{lc}(\dim A_i N_0)$, and $\text{lc}(\dim A_i M_0) = \text{lc}(\dim A_i N_0)$.

**PROOF.** (1) The module $M$ has two filtrations $\{M_i = A_i M_0\}$ and $\{N_i = A_i N_0\}$. Let $v = v_{F,M_0}$ and $\mu = v_{F,N_0}$. Choose a natural number $s$ such that $M_0 \subseteq N_s$ and $N_0 \subseteq M_s$, so $N_i \subseteq M_{i+s}$ and $M_i \subseteq N_{i+s}$ for $i \geq 0$. Since $M_0 \subseteq A_{v(i+s)} N_i, \text{gen}$ for each $i$ and $N_0 \subseteq A_{\mu(i+s)} M_0$, we have $N_0 \subseteq A_{v(i+s)} N_i, \text{gen}$, hence, $\mu(i) \leq v(i+s) + s$. By symmetry, $v(i) \leq \mu(i+s) + s$, so if $\text{lc}(\mu)$ exists then so does $\text{lc}(v)$ and $\text{lc}(\mu) = \text{lc}(v)$.

(2) Since $\dim N_i \leq \dim M_{i+s}$ and $\dim M_i \leq \dim N_{i+s}$ for $i \geq 0$, the statement is clear. □

Lemma 6.4 shows that the leading coefficients of the functions $\dim A_i M_0$ and $v_{F,M_0}$ (if they exist) do not depend on the choice of the generating subspace $M_0$. So, denote them by $l(M) = l_{F}(M)$ and $L(M) = L_{F}(M)$ respectively (if they exist). If GK($M$) (resp. d($A$)) is a natural number, then we denote by $e(M) = e_{F}(M)$ (resp. $E(M) = E_{F}(M)$) the multiplicity of the function $\dim A_i M_0$ (resp. $v_{F,M_0}$).

We denote by $L(A) = L_{F}(A)$ the leading coefficient $L_{F}(A \otimes_{A^0} A)$ of the return function $v_{F \otimes F^0, K}$ of the $A \otimes A^0$-module $A$.

### 6.4. Holonomic modules

**DEFINITION.** Let $A$ be a finitely generated $K$-algebra, and $h_A$ be its holonomic number. A non-zero finitely generated $A$-module $M$ is called a **holonomic** $A$-module if GK($M$) = $h_A$. We denote by $\text{hol}(A)$ the set of all the holonomic $A$-modules.

Since the holonomic number is an infimum it is not clear at the outset that there will be modules which achieve this dimension. Clearly, $\text{hol}(A) \neq \emptyset$ if the Gelfand–Kirillov dimension of every finitely generated $A$-module is a natural number.

A non-zero submodule or a factor module of a holonomic is a holonomic module (since the Gelfand–Kirillov dimension of a submodule or a factor module does not exceed the
Gelfand–Kirillov of the module). If, in addition, the finitely generated algebra \( A \) is left Noetherian and finitely partitive then each holonomic \( A \)-module \( M \) has finite length and each simple sub-factor of \( M \) is a holonomic module.

Let us consider algebras \( A \) having the following properties:

- (S) \( A \) is a simple finitely generated infinite-dimensional algebra.
- (N) There exists a standard finite-dimensional filtration \( F = \{ A_i \} \) of the algebra \( A \) such that the associated graded algebra \( \text{gr } A := \bigoplus_{i \geq 0} A_i / A_{i-1} \), \( A_{-1} = 0 \), is left Noetherian.
- (D) \( \text{GK}(A) < \infty \), \( \text{fd}(A) < \infty \), both \( l(A) = l_F(A) \) and \( L(A) = L_F(A) \) exist.
- (H) For every holonomic \( A \)-module \( M \) there exists \( l(M) = l_F(M) \).

In many cases we use a weaker form of the condition (D).

- (D') \( \text{GK}(A) < \infty \), \( d = \text{fd}(A) < \infty \), there exists \( l(A) = l_F(A) \) and a positive number \( c > 0 \) such that \( \nu(i) \leq ci^d \) for \( i \gg 0 \) where \( \nu \) is the return function \( \nu_{F \otimes F^0, K} \) of the left \( A \otimes A^0 \)-module \( A \).

It follows from (N) that \( A \) is a left Noetherian algebra.

**Lemma 6.5** [4].

1. The Weyl algebra \( A_n \) over a field of characteristic zero with the standard finite-dimensional filtration \( F = \{ A_{n,i} \} \) associated with the canonical generators satisfies the conditions (S), (N), (D), (H). The return function \( \nu_F(i) = i \) for \( i \geq 0 \), and so the leading coefficient of \( \nu_F \) is \( L_F(A_n) = 1 \).

2. \( \nu_{G,K}(i) = i \) for \( i \geq 0 \) and \( L_G(P_n) = 1 \) where \( \nu_{G,K} \) is the return function of the \( A_n \)-module \( P_n = K[X_1, \ldots, X_n] = A_n / \langle A_n \partial_1 + \cdots + A_n \partial_n \rangle \) with the usual filtration \( G = \{ P_{n,i} \} \) of the polynomial algebra.

**Proof.** (1) The only fact that we need to prove is that \( \nu_F(i) = i \) for \( i \geq 0 \). We keep the notation of Theorem 4.6. In the proof of Theorem 4.6 we have seen that \( \nu_F(i) \leq i \) for \( i \geq 0 \).

It remains to prove the reverse inequality.

Each element \( u \) in \( A_n \) can be written in a unique way as a finite sum \( u = \sum \lambda_{\alpha \beta} X^\alpha \partial^\beta \) where \( \lambda_{\alpha \beta} \in K \) and \( X^\alpha \) denotes the monomial \( X_1^{\alpha_1} \cdots X_n^{\alpha_n} \) and similarly \( \partial^\beta \) denotes the monomial \( \partial_1^{\beta_1} \cdots \partial_n^{\beta_n} \). The element \( u \) belongs to \( A_{n,m} \) iff \( |\alpha| + |\beta| \leq m \), where \( |\alpha| = \alpha_1 + \cdots + \alpha_n \). If \( \alpha \in K[X_1, \ldots, X_n] \), then

\[
\partial^m_i \alpha = \sum_{j=0}^{m} \binom{m}{j} \frac{\partial^j \alpha}{\partial X_i^j} \partial^{m-j}_i, \quad m \in \mathbb{N}.
\]

It follows that for any \( v \in \sum A_{n,i} \otimes A^0_{n,j} \), \( i + j < m \), the element \( v X^m = \sum \lambda_{\alpha \beta} X^\alpha \partial^\beta \) has the coefficient \( \lambda_{0,0} = 0 \), hence it could not be a non-zero scalar, and so \( \nu(i) \geq i \) for all \( i \geq 0 \). Hence \( \nu(i) = i \) all \( i \geq 0 \) and then \( L_F(A_n) = 1 \).

(2) The standard filtration of the \( A_n \)-module \( P_n \) associated with the generating subspace \( K \) coincides with the usual filtration of the polynomial algebra \( P_n \). Since \( \partial_j(P_{n,i}) \subseteq P_{n,i-1} \) for all \( i \geq 0 \) and \( j \), \( \nu_{G,K}(i) \leq i \) for \( i \geq 0 \). Using the same arguments as above we see that for any \( u \in \sum_{j=0}^{i-1} A_{n,j} \otimes A^0_{n,i-j-1} \) the element \( u X_1^i \) belongs to the ideal of \( P_n \) generated by \( X_1 \), hence, \( \nu_{G,K}(i) \geq i \), and so \( \nu_{G,K}(i) = i \) for all \( i \geq 0 \) and \( L_G(P_n) = 1 \). \( \square \)
THEOREM 6.6 [4]. Assume that an algebra $A$ satisfies the conditions (S), (H), (D), resp. $(D')$, for some standard finite-dimensional filtration $F = \{A_i\}$ of $A$. Then for every holonomic $A$-module $M$ its leading coefficient is bounded from below by a non-zero constant:

$$l(M) \geq \sqrt[4]{\frac{l(A)}{(L(A)L'(A))^{h_A}}}.$$ 

where

$$L'(A) = \begin{cases} 
L(A), & \text{if } d(A) > 1, \\
L(A) + 1, & \text{if } d(A) = 1, \\
1, & \text{if } d(A) < 1,
\end{cases}$$

resp.

$$l(M) \geq \sqrt[4]{\frac{l(A)}{(c(c+1))^{h_A}}}.$$ 

PROOF. Let $M_0$ be a generating finite-dimensional subspace of $M$ and $\{M_i = A_iM_0\}$ be the standard finite-dimensional filtration of $M$. In the proof of Theorem 3.1 we proved that $\dim A_i \leq \dim M_{\lambda(i)} \dim M_{\lambda(i)+i}$ for $i \geq 0$ where $\lambda$ is the left return function of the algebra $A$ associated with the filtration $F$. Since $\lambda(i) \leq v(i)$ for $i \geq 0$ we have $\dim A_i \leq \dim M_{v(i)} \dim M_{v(i)+i}$, hence, if (D) holds then

$$l(A)i^{\text{GK}(A)} + \cdots \leq l^2(M)\left(L(A)L'(A)\right)^{\text{GK}(M)}l^{\text{GK}(M)(\text{fd}(A)+\max\{\text{fd}(A),1\})} + \cdots,$$

where three dots denote smaller terms.

If $(D')$ holds then

$$l(A)i^{\text{GK}(A)} + \cdots \leq l^2(M)(c(c+1))^{\text{GK}(M)}l^{\text{GK}(M)(\text{fd}(A)+\max\{\text{fd}(A),1\})} + \cdots.$$ 

The module $M$ is holonomic, i.e. $\text{GK}(A) = \text{GK}(M)(\text{fd}(A)+\max\{\text{fd}(A),1\})$. Now, comparing the “leading” coefficients in the inequalities above we finish the proof. \qed

Let $A$ be as in Theorem 6.6. We attach to the algebra $A$ two positive numbers $c_A$ and $c_A'$ in the cases (D) and $(D')$ respectively:

$$c_A = \sqrt{\frac{l(A)}{(L(A)L'(A))^{h_A}}} \quad \text{and} \quad (c_A') = \sqrt[4]{\frac{l(A)}{(c(c+1))^{h_A}}}.$$ 

COROLLARY 6.7. Assume that an algebra $A$ satisfies the conditions (S), (N), (H), (D) or $(D')$. Let $0 \to N \to M \to L \to 0$ be an exact sequence of non-zero finitely generated $A$-modules. Then $M$ is holonomic if and only if $N$ and $L$ are holonomic, in that case $l(M) = l(N) + l(L)$. 


PROOF. The algebra $A$ is left Noetherian, so the module $M$ is finitely generated iff both $N$ and $L$ are so. The proof of Proposition 3.11, [19, p. 295], shows that we can choose finite-dimensional generating subspaces $N_0, M_0, L_0$ of the modules $N, M, L$ respectively such that the sequences

$$0 \to N_i = A_i N_0 \to M_i = A_i M_0 \to L_i = A_i L_0 \to 0$$

are exact for all $i$, hence, $\dim M_i = \dim N_i + \dim L_i$ and the results follow. □

**THEOREM 6.8** [4]. Suppose that the conditions (S), (N), (H), (D) (resp. (D’)) hold. Then each holonomic $A$-module $M$ has finite length which is less or equal to $l(M)/c_A$ (resp. $l(M)/c_A’$).

PROOF. If $M = M_1 \supset M_2 \supset \cdots \supset M_m \supset M_{m+1} = 0$ is a chain of distinct submodules, then by corollary 6.7 and theorem 6.6

$$l(M) = \sum_{i=1}^m l(M_i/M_{i+1}) \geq mc_A \quad (\text{resp. } l(M) \geq mc_A’),$$

thus $m \leq l(M)/c_A$ (resp. $m \leq l(M)/c_A’$). □

**7. Filter dimension and commutative subalgebras of simple finitely generated algebras and their division algebras**

In this section, using the first and the second filter inequalities, we obtain (i) an upper bound for the Gelfand–Kirillov dimension of commutative subalgebras of simple finitely generated infinite-dimensional algebras (Theorem 7.2), and (ii) an upper bound for the transcendence degree of subfields of quotient division rings of (certain) simple finitely generated infinite-dimensional algebras (Theorems 7.4 and 7.5).

For certain classes of algebras and their division algebras the maximum Gelfand–Kirillov dimension/transcendence degree over the commutative subalgebras/subfields were found in [1,12,17,13–15,2], and [23].

Recall that

the Gelfand–Kirillov dimension $\text{GK}(C) = \text{the Krull dimension } \text{K.dim}(C)$

$= \text{the transcendence degree } \text{tr.deg}_K(C)$

for every commutative finitely generated algebra $C$ which is a domain.

**7.1. An upper bound for the Gelfand–Kirillov dimensions of commutative subalgebras of simple finitely generated algebras**

**PROPOSITION 7.1.** Let $A$ and $C$ be finitely generated algebras such that $C$ is a commutative domain with field of fractions $Q$. $B := C \otimes A$, and $B := Q \otimes A$. Let $M$ be a finitely generated $B$-module such that $\mathcal{M} := B \otimes_B M \neq 0$. Then $\text{GK}(B M) \geq \text{GK}_Q(B M) + \text{GK}(C)$. 

**Remark.** $\text{GK}_Q$ stands for the Gelfand–Kirillov dimension over the field $Q$.

**Proof.** Let us fix standard filtrations $\{A_i\}$ and $\{C_i\}$ for the algebras $A$ and $C$ respectively. Let $h(t) \in \mathbb{Q}[t]$ be the *Hilbert polynomial* for the algebra $C$, i.e. $\dim_K(C_i) = h(i)$ for $i \gg 0$. Recall that $\text{GK}(C) = \deg(h(t))$. The algebra $B$ has a standard filtration $\{B_i\}$ which is the tensor product of the standard filtrations $\{C_i\}$ and $\{A_i\}$ of the algebras $C$ and $A$, i.e. $B_i := \sum_{j=0}^{i} C_j \otimes A_{i-j}$. By the assumption, the $B$-module $M$ is finitely generated, so $M = BM_0$ where $M_0$ is a finite-dimensional generating subspace for $M$. Then the $B$-module $M$ has a standard filtration $\{M_i := B_iM_0\}$. The $Q$-algebra $B$ has a standard (finite-dimensional over $Q$) filtration $\{B_i := Q \otimes A_i\}$, and the $B$-module $M$ has a standard (finite-dimensional over $Q$) filtration $\{M_i := B_iM'_0 = QA_iM'_0\}$ where $M'_0$ is the image of the vector space $M_0$ under the $B$-module homomorphism $M \to M, m \mapsto m' := 1 \otimes_B m$.

For each $i \gg 0$, one can fix a $K$-subspace, say $L_i$, of $A_iM_0'$ such that $\dim_K(QA_iM'_0) = \dim_K(L_i)$. Now, $B_{2i} \supseteq C_i \otimes A_i$ implies $\dim_K(B_{2i}M_0) \geq \dim_K((C_i \otimes A_i)M_0)$, and $((C_i \otimes A_i)M_0)^\gamma \supseteq C_iL_i$ implies $\dim_K(((C_i \otimes A_i)M_0)^\gamma) \geq \dim_K(C_iL_i) = \dim_K((C_i) \dim_K(L_i) = \dim_K(C_i) \dim_Q(M_i))$. It follows that

$$\begin{align*}
\text{GK}(B,M) &= \gamma \left( \dim_K(M_i) \right) \geq \gamma \left( \dim_K((C_i \otimes A_i)M_0) \right) \\
&\geq \gamma \left( \dim_K ((C_i \otimes A_i)M_0)^\gamma \right) \geq \gamma \left( \dim_K(C_i) \dim_Q(M_i) \right) \\
&= \gamma \left( \dim_K(C_i) \right) + \gamma \left( \dim_Q(M_i) \right) \\
&= \text{GK}(C) + \text{GK}_Q(B,M).
\end{align*}$$

Recall that $d = \dim A_{\text{fd}}$. A $K$-algebra $A$ is called *central* if its centre $Z(A) = K$.

**Theorem 7.2** [7]. Let $A$ be a central simple finitely generated $K$-algebra of Gelfand–Kirillov dimension $0 < n < \infty$ (over $K$). Let $C$ be a commutative subalgebra of $A$. Then

$$\text{GK}(C) \leq \text{GK}(A) \left( 1 - \frac{1}{f_A + \max\{f_A, 1\}} \right),$$

where $f_A := \max\{d_{Q_m}(Q_m \otimes A) \mid 0 \leq m \leq n\}$, $Q_0 := K$, and $Q_m := K(x_1, \ldots, x_m)$ is a rational function field in indeterminates $x_1, \ldots, x_m$.

**Proof.** Let $P_m = K[x_1, \ldots, x_m]$ be a polynomial algebra over the field $K$. Then $Q_m$ is its field of fractions and $\text{GK}(P_m) = m$. Suppose that $P_m$ is a subalgebra of $A$. Then $m = \text{GK}(P_m) \leq \text{GK}(A) = n$. For each $m \geq 0$, $Q_m \otimes A$ is a central simple $Q_m$-algebra, [19, 9.6.9], of Gelfand–Kirillov dimension (over $Q_m$) $\text{GK}_{Q_m}(Q_m \otimes A) = \text{GK}(A) > 0$, hence

$$\text{GK}(A) = \text{GK}(AA) \geq \text{GK}(A_{P_m}) = \text{GK}(P_m \otimes A) \quad (P_m \text{ is commutative}).$$
\[
\geq \text{GK}_{Q_m}(Q_m \otimes A(Q_m \otimes P_m A)) + \text{GK}(P_m) \quad \text{(Lemma 7.1)}
\]
\[
\geq \frac{\text{GK}(A)}{d_{Q_m}(Q_m \otimes A) + \max\{d_{Q_m}(Q_m \otimes A), 1\}} + m \quad \text{(Theorem 3.1)}.
\]

Hence,
\[
m \leq \text{GK}(A) \left(1 - \frac{1}{d_{Q_m}(Q_m \otimes A) + \max\{d_{Q_m}(Q_m \otimes A), 1\}}\right) \leq \text{GK}(A),
\]
and so
\[
\text{GK}(C) \leq \text{GK}(A) \left(1 - \frac{1}{f_A + \max\{f_A, 1\}}\right).
\]

As a consequence we have a short proof of the following well-known result.

**Corollary 7.3.** Let \( K \) be an algebraically closed field of characteristic zero, \( X \) be a smooth irreducible affine algebraic variety of dimension \( n := \dim(X) > 0 \), and \( C \) be a commutative subalgebra of the ring of differential operators \( D(X) \). Then \( \text{GK}(C) \leq n \).

**Proof.** The algebra \( D(X) \) is central since \( K \) is an algebraically closed field of characteristic zero [19, chapter 15]. By Theorem 4.3, \( f_{D(X)} = 1 \), and then, by Theorem 7.2,
\[
\text{GK}(C) \leq 2n \left(1 - \frac{1}{1 + 1}\right) = n.
\]

**Remark.** For the ring of differential operators \( D(X) \) the upper bound in Theorem 7.2 for the Gelfand–Kirillov dimension of commutative subalgebras of \( D(X) \) is an exact upper bound since as we mentioned above the algebra \( \mathcal{O}(X) \) of regular functions on \( X \) is a commutative subalgebra of \( D(X) \) of Gelfand–Kirillov dimension \( n \).

### 7.2. An upper bound for the transcendence degree of subfields of quotient division algebras of simple finitely generated algebras

Recall that the transcendence degree \( \text{tr.deg}_K(L) \) of a field extension \( L \) of a field \( K \) coincides with the Gelfand–Kirillov dimension \( \text{GK}_K(L) \), and, by the Goldie’s theorem, a left Noetherian semiprime algebra \( A \) has a quotient algebra \( D = D_A \) (i.e. \( D = S^{-1}A \) where \( S \) is the set of regular elements = the set of non-zerodivisors of \( A \)). As a rule, the quotient algebra \( D \) has infinite Gelfand–Kirillov dimension and is not a finitely generated algebra (e.g., the quotient division algebra \( D(X) \) of the ring of differential operators \( D(X) \) on each smooth irreducible affine algebraic variety \( X \) of dimension \( n > 0 \) over a field \( K \) of characteristic zero contains a non-commutative free subalgebra since \( D(X) \supseteq D(\mathbb{A}^1) \) and the first Weyl division algebra \( D(\mathbb{A}^1) \) has this property, [18]). So, if we want to find an upper bound
for the transcendence degree of subfields in the quotient algebra $D$ we can not apply Theorem 7.2. Nevertheless, imposing some natural (mild) restrictions on the algebra $A$ one can obtain exactly the same upper bound for the transcendence degree of subfields in the quotient algebra $D_A$ as the upper bound for the Gelfand–Kirillov dimension of commutative subalgebras in $A$.

**Theorem 7.4** [7]. Let $A$ be a simple finitely generated $K$-algebra such that $0 < n := \text{GK}(A) < \infty$, all the algebras $Q_m \otimes A$, $m \geq 0$, are simple finitely partitive algebras where $Q_0 := K$, $Q_m := K(x_1, \ldots, x_m)$ is a rational function field and, for each $m \geq 0$, the Gelfand–Kirillov dimension (over $Q_m$) of every finitely generated $Q_m \otimes A$-module is a natural number. Let $B = S^{-1}A$ be the localization of the algebra $A$ at a left Ore subset $S$ of $A$. Let $L$ be a (commutative) subfield of the algebra $B$ that contains $K$. Then

$$\text{tr.deg}_K(L) \leq \text{GK}(A) \left(1 - \frac{1}{f_A + \max\{f_A, 1\}}\right),$$

where $f_A := \max\{d_{Q_m}(Q_m \otimes A) \mid 0 \leq m \leq n\}$.

**Proof.** It follows immediately from a definition of the Gelfand–Kirillov dimension that $\text{GK}_{K'}(K' \otimes C) = \text{GK}(C)$ for any $K$-algebra $C$ and any field extension $K'$ of $K$. In particular, $\text{GK}_{Q_m}(Q_m \otimes A) = \text{GK}(A)$ for all $m \geq 0$. By theorem 4.2,

$$\text{K.dim}(Q_m \otimes A) \leq \text{GK}(A) \left(1 - \frac{1}{d_{Q_m}(Q_m \otimes A) + \max\{d_{Q_m}(Q_m \otimes A), 1\}}\right).$$

Let $L$ be a subfield of the algebra $B$ that contains $K$. Suppose that $L$ contains a rational function field (isomorphic to) $Q_m$ for some $m \geq 0$.

$$m = \text{tr.deg}_K(Q_m) \leq \text{K.dim}(Q_m \otimes Q_m)$$

$$\leq \text{K.dim}(Q_m \otimes B)$$

(by [19, 6.5.3] since $Q_m \otimes B$ is a free $Q_m \otimes Q_m$-module)

$$= \text{K.dim}(Q_m \otimes S^{-1}A) = \text{K.dim}(S^{-1}(Q_m \otimes A))$$

$$\leq \text{K.dim}(Q_m \otimes A)$$

(by [19, 6.5.3(ii)(b)])

$$\leq \text{GK}(A) \left(1 - \frac{1}{d_{Q_m}(Q_m \otimes A) + \max\{d_{Q_m}(Q_m \otimes A), 1\}}\right) \leq \text{GK}(A).$$

Hence

$$\text{tr.deg}_K(L) \leq \text{GK}(A) \left(1 - \frac{1}{f_A + \max\{f_A, 1\}}\right).$$

Recall that every somewhat commutative algebra $A$ is a Noetherian finitely generated finitely partitive algebra of finite Gelfand–Kirillov dimension, the Gelfand–Kirillov di-
mension of every finitely generated $A$-modules is an integer, and (Quillen’s lemma): the ring $\text{End}_A(M)$ is algebraic over $K$ (see [19, Chapter 8] or [16] for details).

**Theorem 7.5** [7]. Let $A$ be a central simple somewhat commutative infinite-dimensional $K$-algebra and let $D = D_A$ be its quotient algebra. Let $L$ be a subfield of $D$ that contains $K$. Then the transcendence degree of the field $L$ (over $K$)

$$\text{tr.deg}_K(L) \leq \text{GK}(A) \left(1 - \frac{1}{f_A + \max\{f_A, 1\}}\right),$$

where $f_A := \max\{d_{Q_m}(Q_m \otimes A) \mid 0 \leq m \leq \text{GK}(A)\}$.

**Proof.** The algebra $A$ is a somewhat commutative algebra, so it has a finite-dimensional filtration $A = \bigcup_{i \geq 0} A_i$ such that the associated graded algebra is a commutative finitely generated algebra. For each integer $m \geq 0$, the $Q_m$-algebra $Q_m \otimes A = \bigcup_{i \geq 0} Q_m \otimes A_i$ has the finite-dimensional filtration (over $Q_m$) such that the associated graded algebra $\text{gr}(Q_m \otimes A) = \bigoplus_{i \geq 0} Q_m \otimes A_i / Q_m \otimes A_{i-1} \simeq Q_m \otimes \text{gr}(A)$ is a commutative finitely generated $Q_m$-algebra. So, $Q_m \otimes A$ is a somewhat commutative $Q_m$-algebra.

By the assumption $\dim_K(A) = \infty$, hence $\text{dim}_K(\text{gr}(A)) = \infty$ which implies $\text{GK}(\text{gr}(A)) > 0$, and so $\text{GK}(A) > 0$ (since $\text{GK}(A) = \text{GK}(\text{gr}(A))$). The algebra $A$ is a central simple $K$-algebra, so $Q_m \otimes A$ is a central simple $Q_m$-algebra, [19, 9.6.9]. Now, Theorem 7.5 follows from Theorem 7.4 applied to $B = D$. \hfill \Box

**Theorem 7.6.** Let $K$ be an algebraically closed field of characteristic zero, $\mathcal{D}(X)$ be the ring of differential operators on a smooth irreducible affine algebraic variety $X$ of dimension $n > 0$, and $D(X)$ be the quotient division ring for $\mathcal{D}(X)$. Let $L$ be a (commutative) subfield of $D(X)$ that contains $K$. Then $\text{tr.deg}_K(L) \leq n$.

**Remark.** This inequality is, in fact, an exact upper bound for the transcendence degree of subfields in $D(X)$ since the field of fractions $Q(X)$ for the algebra $\mathcal{O}(X)$ is a commutative subfield of the division ring $D(X)$ with $\text{tr.deg}_K(Q(X)) = n$.

**Proof.** Since $Q_m \otimes \mathcal{D}_K(\mathcal{O}(X)) \simeq \mathcal{D}_{Q_m}(Q_m \otimes \mathcal{O}(X))$ and $d(\mathcal{D}(Q_m \otimes \mathcal{O}(X))) = 1$ for all $m \geq 0$ we have $f_{\mathcal{D}(X)} = 1$. Now, Theorem 7.6 follows from Theorem 7.5,

$$\text{tr.deg}_K(L) \leq 2n \left(1 - \frac{1}{1 + 1}\right) = n. \quad \Box$$

Following [15] for a $K$-algebra $A$ define the commutative dimension

$$\text{Cdim}(A) := \max\{\text{GK}(C) \mid C \text{ is a commutative subalgebra of } A\}.$$

The commutative dimension $\text{Cdim}(A)$ (if finite) is the largest non-negative integer $m$ such that the algebra $A$ contains a polynomial algebra in $m$ variables ([15, 1.1], or [19, 8.2.14]). So, $\text{Cdim}(A) = \mathbb{N} \cup \{\infty\}$. If $A$ is a subalgebra of $B$ then $\text{Cdim}(A) \leq \text{Cdim}(B)$. 

---

**Filter dimension**

101
COROLLARY 7.7. Let $X$ and $Y$ be smooth irreducible affine algebraic varieties of dimensions $n$ and $m$ respectively, let $D(X)$ and $D(Y)$ be quotient division rings for the rings of differential operators $\mathcal{D}(X)$ and $\mathcal{D}(Y)$. Then there is no $K$-algebra embedding $D(X) \rightarrow D(Y)$ if $n > m$.

PROOF. By Theorem 7.6, $\operatorname{Cdim}(D(X)) = n$ and $\operatorname{Cdim}(D(Y)) = m$. Suppose that there is a $K$-algebra embedding $D(X) \rightarrow D(Y)$. Then $n = \operatorname{Cdim}(D(X)) \leq \operatorname{Cdim}(D(Y)) = m$. □

For the Weyl algebras $A_n = \mathcal{D}(\mathbb{A}^n)$ and $A_m = \mathcal{D}(\mathbb{A}^m)$ the result above was proved by Gelfand and Kirillov in [12]. They introduced a new invariant of an algebra $A$, the so-called (Gelfand–Kirillov) transcendence degree $\operatorname{GKtr.deg}(A)$, and proved that $\operatorname{GKtr.deg}(D_n) = 2n$. Recall that

$$\operatorname{GKtr.deg}(A) := \sup_{V} \inf_{b} \operatorname{GK}(K[bV]),$$

where $V$ ranges over the finite-dimensional subspaces of $A$ and $b$ ranges over the regular elements of $A$. Another proofs of the corollary based on different ideas were given by A. Joseph, [14], and R. Resco, [23], see also [19, 6.6.19]. Joseph’s proof is based on the fact that the centralizer of any isomorphic copy of the Weyl algebra $A_n$ in its division algebra $D_n := \mathcal{D}(\mathbb{A}^n)$ reduces to scalars ([15, 4.2]), Resco proved that $\operatorname{Cdim}(D_n) = n$ ([23, 4.2]) using the result of Rentschler and Gabriel, [22], that $\operatorname{K.dim}(A_n) = n$ (over an arbitrary field of characteristic zero).

8. Filter dimension and isotropic subalgebras of Poisson algebras

In this section, we apply Theorem 7.2 to obtain an upper bound for the Gelfand–Kirillov dimension of isotropic subalgebras of certain Poisson algebras (Theorem 8.1).

Let $(P, \{\cdot, \cdot\})$ be a Poisson algebra over the field $K$. Recall that $P$ is an associative commutative $K$-algebra which is a Lie algebra with respect to the bracket $\{\cdot, \cdot\}$ for which Leibniz’s rule holds:

$$\{a, xy\} = \{a, x\}y + x\{a, y\} \quad \text{for all } a, x, y \in P,$$

which means that the inner derivation $\text{ad}(a): P \rightarrow P, x \mapsto \{a, x\}$, of the Lie algebra $P$ is also a derivation of the associative algebra $P$. Therefore, to each Poisson algebra $P$ one can attach an associative subalgebra $A(P)$ of the ring of differential operators $\mathcal{D}(P)$ with coefficients from the algebra $P$ which is generated by $P$ and $\text{ad}(P) := \{\text{ad}(a) | a \in P\}$. If $P$ is a finitely generated algebra then so is the algebra $A(P)$ with $\operatorname{GK}(A(P)) \leq \operatorname{GK}(\mathcal{D}(P)) < \infty$.

EXAMPLE. Let $P_{2n} = K[x_1, \ldots, x_{2n}]$ be the Poisson polynomial algebra over a field $K$ of characteristic zero equipped with the Poisson bracket

$$\{f, g\} = \sum_{i=1}^{n} \left( \frac{\partial f}{\partial x_i} \frac{\partial g}{\partial x_{n+i}} - \frac{\partial f}{\partial x_{n+i}} \frac{\partial g}{\partial x_i} \right).$$
The algebra $A(P_{2n})$ is generated by the elements
\[ x_1, \ldots, x_{2n}, \quad \text{ad}(x_i) = \frac{\partial}{\partial x_{n+i}}, \quad \text{ad}(x_{n+i}) = -\frac{\partial}{\partial x_i}, \quad i = 1, \ldots, n. \]
So, the algebra $A(P_{2n})$ is canonically isomorphic to the Weyl algebra $A_{2n}$.

**Definition.** We say that a Poisson algebra $P$ is a strongly simple Poisson algebra if
1. $P$ is a finitely generated (associative) algebra which is a domain,
2. the algebra $A(P)$ is central simple, and
3. for each set of algebraically independent elements $a_1, \ldots, a_m$ of the algebra $P$ such that $\{a_i, a_j\} = 0$ for all $i, j = 1, \ldots, m$, the (commuting) elements $a_1, \ldots, a_m$, ad$(a_1), \ldots, \text{ad}(a_m)$ of the algebra $A(P)$ are algebraically independent.

**Theorem 8.1** [7]. Let $P$ be a strongly simple Poisson algebra, and $C$ be an isotropic subalgebra of $P$, i.e., $\{C, C\} = 0$. Then
\[ \text{GK}(C) \leq \frac{\text{GK}(A(P))}{2} \left( 1 - \frac{1}{f_{A(P)} + \text{max}\{f_{A(P)}, 1\}} \right), \]
where $f_{A(P)} := \text{max}\{d_{Q_m}(Q_m \otimes A(P)) \mid 0 \leq m \leq \text{GK}(A(P))\}$.

**Proof.** By assumption the finitely generated algebra $P$ is a domain, hence the finitely generated algebra $A(P)$ is a domain (as a subalgebra of the domain $\mathcal{D}(Q(P))$, the ring of differential operators with coefficients from the field of fractions $Q(P)$ for the algebra $P$). It suffices to prove the inequality for isotropic subalgebras of the Poisson algebra $P$ that are polynomial algebras. So, let $C$ be an isotropic polynomial subalgebra of $P$ in $m$ variables, say $a_1, \ldots, a_m$. By assumption, the commuting elements $a_1, \ldots, a_m$, ad$(a_1), \ldots, \text{ad}(a_m)$ of the algebra $A(P)$ are algebraically independent. So, the Gelfand–Kirillov dimension of the subalgebra $C'$ of $A(P)$ generated by these elements is equal to $2m$. By Theorem 7.2,
\[ 2\text{GK}(C) = 2m = \text{GK}(C') \leq \text{GK}(A(P)) \left( 1 - \frac{1}{f_{A(P)} + \text{max}\{f_{A(P)}, 1\}} \right), \]
and this proves the inequality. □

**Corollary 8.2.**
1. The Poisson polynomial algebra $P_{2n} = K[x_1, \ldots, x_{2n}]$ (with the Poisson bracket) over a field $K$ of characteristic zero is a strongly simple Poisson algebra, the algebra $A(P_{2n})$ is canonically isomorphic to the Weyl algebra $A_{2n}$.
2. The Gelfand–Kirillov dimension of every isotropic subalgebra of the polynomial Poisson algebra $P_{2n}$ is $\leq n$.

**Proof.** (1) The third condition in the definition of strongly simple Poisson algebra is the only statement we have to prove. So, let $a_1, \ldots, a_m$ be algebraically independent elements
of the algebra \( P_{2n} \) such that \( \{a_i, a_j\} = 0 \) for all \( i, j = 1, \ldots, m \). One can find polynomials, say \( a_{m+1}, \ldots, a_{2n} \), in \( P_{2n} \) such that the elements \( a_1, \ldots, a_{2n} \) are algebraically independent, hence the determinant \( d \) of the Jacobian matrix \( J := (\frac{\partial a_i}{\partial x_j}) \) is a non-zero polynomial. Let \( X = \langle \{x_i, a_j\} \rangle \) and \( Y = \langle \{a_i, a_j\} \rangle \) be, so-called, the Poisson matrices associated with the elements \( \{x_i\} \) and \( \{a_i\} \). It follows from \( Y = J^T X J \) that \( \det(Y) = d^2 \det(X) \neq 0 \) since \( \det(X) \neq 0 \). The derivations

\[ \delta_i := d^{-1} \det \begin{pmatrix} \{a_1, a_1\} & \ldots & \{a_1, a_{i-1}\} & \{a_1, a_{i+1}\} & \ldots & \{a_1, a_{2n}\} \\ \{a_2, a_1\} & \ldots & \{a_2, a_{i-1}\} & \{a_2, a_{i+1}\} & \ldots & \{a_2, a_{2n}\} \\ \vdots & & & & & \\ \{a_{2n}, a_1\} & \ldots & \{a_{2n}, a_{i-1}\} & \{a_{2n}, a_{i+1}\} & \ldots & \{a_{2n}, a_{2n}\} \end{pmatrix}, \]

\( i = 1, \ldots, 2n \), of the rational function field \( Q_{2n} = K(x_1, \ldots, x_{2n}) \) satisfy the following properties: \( \delta_i(a_j) = \delta_j, \) the Kronecker delta. For each \( i \) and \( j \), the kernel of the derivation \( \Delta_{ij} := \delta_i \delta_j - \delta_j \delta_i \in \text{Der}_K(Q_{2n}) \) contains \( 2n \) algebraically independent elements \( a_1, \ldots, a_{2n} \). Hence \( \Delta_{ij} = 0 \) since the field \( Q_{2n} \) is algebraic over its subfield \( K(a_1, \ldots, a_{2n}) \) and \( \text{char}(K) = 0 \). So, the subalgebra, say \( W \), of the ring of differential operators \( D(Q_{2n}) \) generated by the elements \( a_1, \ldots, a_{2n}, \delta_1, \ldots, \delta_{2n} \) is isomorphic to the Weyl algebra \( A_{2n} \), and so \( \text{GK}(W) = \text{GK}(A_{2n}) = 4n \).

Let \( U \) be the \( K \)-subalgebra of \( D(Q_{2n}) \) generated by the elements \( x_1, \ldots, x_{2n}, \delta_1, \ldots, \delta_{2n}, \) and \( d^{-1} \). Let \( P' \) be the localization of the polynomial algebra \( P_{2n} \) at the powers of the element \( d \). Then \( \delta_1, \ldots, \delta_{2n} \in \sum_{i=1}^{2n} P' \text{ad}(a_i) \) and \( \text{ad}(a_1), \ldots, \text{ad}(a_{2n}) \in \sum_{i=1}^{2n} P' \delta_i \), hence the algebra \( U \) is generated (over \( K \)) by \( P' \) and \( \text{ad}(a_1), \ldots, \text{ad}(a_{2n}) \). The algebra \( U \) can be viewed as a subalgebra of the ring of differential operators \( D(P') \). Now, the inclusions, \( W \subseteq U \subseteq D(P') \) imply \( 4n = \text{GK}(W) \leq \text{GK}(U) \leq \text{GK}(D(P')) = 2\text{GK}(P') = 4n \), therefore \( \text{GK}(U) = 4n \). The algebra \( U \) is a factor algebra of an iterated Ore extension \( V = P'[t_1; \text{ad}(a_1)\cdots [t_{2n}; \text{ad}(a_{2n})] \). Since \( P' \) is a domain, so is the algebra \( V \). The algebra \( P' \) is a finitely generated algebra of Gelfand–Kirillov dimension \( 2n \), hence \( \text{GK}(V) = \text{GK}(P') + 2n = 4n \) (by [19, 8.2.11]). Since \( \text{GK}(V) = \text{GK}(U) \) and any proper factor algebra of \( V \) has Gelfand–Kirillov dimension strictly less than \( \text{GK}(V) \) (by [19, 8.3.5], since \( V \) is a domain), the algebras \( V \) and \( U \) must be isomorphic. Therefore, the (commuting) elements \( a_1, \ldots, a_m, \text{ad}(a_1), \ldots, \text{ad}(a_m) \) of the algebra \( U \) (and \( A(P) \)) must be algebraically independent.

(2) Let \( C \) be an isotropic subalgebra of the Poisson algebra \( P_{2n} \). Note that \( f_{A(P_{2n})} = f_{A_{2n}} = 1 \) and \( \text{GK}(A_{2n}) = 4n \). By Theorem 8.1,

\[ \text{GK}(C) \leq 4n \leq \frac{2}{2} \left( \frac{1}{1} \cdot \frac{1}{1+1} \right) = n. \]

\( \Box \)

REMARK. This result means that for the Poisson polynomial algebra \( P_{2n} \) the right-hand side of the inequality of Theorem 8.1 is the exact upper bound for the Gelfand–Kirillov dimension of isotropic subalgebras in \( P_{2n} \) since the polynomial subalgebra \( K[x_1, \ldots, x_n] \) of \( P_{2n} \) is isotropic.
References

Section 4E
Lie Algebras
This page intentionally left blank
# Gelfand–Tsetlin Bases for Classical Lie Algebras

**A.I. Molev**  
*School of Mathematics and Statistics, University of Sydney, NSW 2006, Australia*  
*E-mail: alexm@maths.usyd.edu.au*

## Contents

1. Introduction ........................................ 111  
2. Gelfand–Tsetlin bases for representations of $\mathfrak{gl}_n$  
   2.1. Construction of the basis: lowering and raising operators ........................................ 119  
   2.2. Mickelsson algebra theory ........................................ 121  
   2.3. Mickelsson–Zhelobenko algebra $Z(\mathfrak{gl}_n, \mathfrak{gl}_{n-1})$ ........................................ 124  
   2.4. Characteristic identities ........................................ 129  
   2.5. Quantum minors ........................................ 133  
Bibliographical notes ........................................ 135  
3. Weight bases for representations of $\mathfrak{o}_N$ and $\mathfrak{sp}_{2N}$  
   3.1. Raising and lowering operators ........................................ 138  
   3.2. Branching rules, patterns and basis vectors ........................................ 142  
   3.3. Yangians and their representations ........................................ 146  
   3.4. Yangian action on the multiplicity space ........................................ 151  
   3.5. Calculation of the matrix elements ........................................ 155  
Bibliographical notes ........................................ 158  
4. Gelfand–Tsetlin bases for representations of $\mathfrak{o}_N$ ........................................ 158  
   4.1. Lowering operators for the reduction $\mathfrak{o}_{2n+1} \downarrow \mathfrak{o}_{2n}$ ........................................ 159  
   4.2. Lowering operators for the reduction $\mathfrak{o}_{2n} \downarrow \mathfrak{o}_{2n-1}$ ........................................ 160  
   4.3. Basis vectors ........................................ 162  
Bibliographical notes ........................................ 164  
Acknowledgements ........................................ 164  
References ........................................ 164
1. Introduction

The theory of semisimple Lie algebras and their representations lies in the heart of modern mathematics. It has numerous connections with other areas of mathematics and physics. The simple Lie algebras over the field of complex numbers were classified in the work of Cartan and Killing in the 1930s. There are four infinite series $A_n$, $B_n$, $C_n$, $D_n$ which are called the classical Lie algebras, and five exceptional Lie algebras $E_6$, $E_7$, $E_8$, $F_4$, $G_2$. The structure of these Lie algebras is uniformly described in terms of certain finite sets of vectors in a Euclidean space called root systems. Due to the Weyl complete reducibility theorem, the theory of finite-dimensional representations of the semisimple Lie algebras is largely reduced to the study of irreducible representations. The irreducibles are parametrized by their highest weights. The characters and dimensions are explicitly known by the Weyl formula. The reader is referred to, e.g., the books of Bourbaki, [11], Dixmier, [19], Humphreys, [61], or Goodman and Wallach, [45], for a detailed exposition of the theory.

However, the Weyl formula for the dimension does not use any explicit construction of the representations. Such constructions remained unknown until 1950 when Gelfand and Tsetlin published two short papers, [41] and [42] (in Russian), where they solved the problem for the general linear Lie algebras (type $A_n$) and the orthogonal Lie algebras (types $B_n$ and $D_n$), respectively. Later, Baird and Biedenharn, [4] (1963), commented on [41] as follows:

This paper is extremely brief (three pages) and does not appear to have been translated in either the usual journal translations or the translations on group-theoretical subjects of the American Mathematical Society, or even referred to in the review articles on group theory by Gelfand himself. Moreover, the results are presented without the slightest hint as to the methods employed and contain not a single reference or citation of other work. In an effort to understand the meaning of this very impressive work, we were led to develop the proofs ... 

Baird and Biedenharn employed the calculus of Young patterns to derive the Gelfand–Tsetlin formulas. Their interest to the formulas was also motivated by the connection with the fundamental Wigner coefficients; see Section 2.4 below.

A year earlier (1962) Zhelobenko published an independent paper, [167], where he derived the branching rules for all classical Lie algebras. In his approach the representations are realized in a space of polynomials satisfying the “indicator system” of differential equations. He outlined a method to construct the lowering operators and to derive the matrix element formulas for the case of the general linear Lie algebra $g_{n}$. An explicit “infinitesimal” form for the lowering operators as elements of the enveloping algebra was found by Nagel and Moshinsky, [106] (1964), and independently by Hou Pei-yu, [59] (1966). The latter work relies on Zhelobenko’s results, [167], and also contains a derivation of the Gelfand–Tsetlin formulas alternative to that of Baird and Biedenharn. This approach was further developed in the book by Zhelobenko, [168], which contains its detailed account.

The work of Nagel and Moshinsky was extended to the orthogonal Lie algebras $o_N$ by Pang and Hecht, [133], and Wong, [164], who produced explicit infinitesimal expres-
sions for the lowering operators and gave a derivation of the formulas of Gelfand and Tsetlin, [42].

During the half a century passed since the work of Gelfand and Tsetlin, many different approaches were developed to construct bases of the representations of the classical Lie algebras. New interpretations of the lowering operators and new proofs of the Gelfand–Tsetlin formulas were discovered by several authors. In particular, Gould, [46–48,50], employed the characteristic identities of Bracken and Green, [12,54], to calculate the Wigner coefficients and matrix elements of generators of $\mathfrak{gl}_n$ and $\mathfrak{o}_N$. The extremal projector discovered by Asherova, Smirnov and Tolstoy, [1–3], turned out to be a powerful instrument in the representation theory of simple Lie algebras. It plays an essential role in the theory of Mickelsson algebras developed by Zhelobenko which has a wide spectrum of applications from the branching rules and reduction problems to the classification of Harish-Chandra modules; see Zhelobenko’s expository paper, [173], and his book, [174]. Two different quantum minor interpretations of the lowering and raising operators were given by Nazarov and Tarasov, [109], and the author, [96]. These techniques are based on the theory of quantum algebras called the Yangians and allow an independent derivation of the matrix element formulas. We shall discuss the above approaches in more detail in Sections 2.3, 2.4 and 2.5 below.

A quite different method to construct modules over the classical Lie algebras is developed in the papers by King and El-Sharkaway, [69], Berele, [6], King and Welsh, [70], Koike and Terada, [73], Proctor, [138], Nazarov, [107]. In particular, bases in the representations of the orthogonal and symplectic Lie algebras parametrized by $\mathfrak{o}_N$-standard or $\mathfrak{sp}_{2n}$-standard Young tableaux are constructed. This method provides an algorithm for the calculation of the representation matrices. It is based on the Weyl realization of the representations of the classical groups in tensor spaces; see Weyl, [159]. A detailed exposition of the theory of the classical groups together with many recent developments are presented in the book by Goodman and Wallach, [45].

Bases with special properties in the universal enveloping algebra for a simple Lie algebra $\mathfrak{g}$ and in some modules over $\mathfrak{g}$ were constructed by Lakshmibai, Musili and Seshadri, [75], Littelmann, [81,82], Chari and Xi, [15] (monomial bases); De Concini and Kazhdan, [18], Xi, [166] (special bases and their $q$-analogues); Gelfand and Zelevinsky, [44], Retakh and Zelevinsky, [140], Mathieu, [85] (good bases); Lusztig, [83], Kashiwara, [66], Du, [35,36] (canonical or crystal bases); see also Mathieu, [86], for a review and more references. Algorithms for computing the global crystal bases of irreducible modules for the classical Lie algebras were recently given by Leclerc and Toffin, [76], and Lecouvey, [77,78]. In general, no explicit formulas are known, however, for the matrix elements of the generators in such bases other than those of Gelfand and Tsetlin type. It is known, although, that for the canonical bases the matrix elements of the standard generators are nonnegative integers. Some classes of representations of the symplectic, odd orthogonal and the Lie algebras of type $G_2$ were explicitly constructed by Donnelly, [23,24,26], and Donnelly, Lewis and Pervine, [27]. The constructions were applied to establish combinatorial properties of the supporting graphs of the representations and were inspired by the earlier results of Proctor, [134,135,137]. Another graph-theoretic approach is developed by Wildberger, [160–163], to construct simple Lie algebras and their minuscule representations; see also Stembridge, [146].
We now discuss the main idea which leads to the construction of the Gelfand–Tsetlin bases. The first point is to regard a given classical Lie algebra not as a single object but as a part of a chain of subalgebras with natural embeddings. We illustrate this idea using representations of the symmetric groups $\mathfrak{S}_n$ as an example. Consider the chain of subgroups

$$\mathfrak{S}_1 \subset \mathfrak{S}_2 \subset \cdots \subset \mathfrak{S}_n,$$

(1.1)

where the subgroup $\mathfrak{S}_k$ of $\mathfrak{S}_{k+1}$ consists of the permutations which fix the index $k + 1$ of the set $\{1, 2, \ldots, k + 1\}$. The irreducible representations of the group $\mathfrak{S}_n$ are indexed by partitions $\lambda$ of $n$. A partition $\lambda = (\lambda_1, \ldots, \lambda_l)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l$ is depicted graphically as a Young diagram which consists of $l$ left-justified rows of boxes so that the top row contains $\lambda_1$ boxes, the second row $\lambda_2$ boxes, etc. Denote by $V(\lambda)$ the irreducible representation of $\mathfrak{S}_n$ corresponding to the partition $\lambda$. One of the central results of the representation theory of the symmetric groups is the following branching rule which describes the restriction of $V(\lambda)$ to the subgroup $\mathfrak{S}_{n-1}$:

$$V(\lambda)|_{\mathfrak{S}_{n-1}} = \bigoplus_{\mu} V'(\mu),$$

summed over all partitions $\mu$ whose Young diagram is obtained from that of $\lambda$ by removing one box. Here $V'(\mu)$ denotes the irreducible representation of $\mathfrak{S}_{n-1}$ corresponding to a partition $\mu$. Thus, the restriction of $V(\lambda)$ to $\mathfrak{S}_{n-1}$ is multiplicity-free, i.e., it contains each irreducible representation of $\mathfrak{S}_{n-1}$ at most once. This makes it possible to obtain a natural parameterization of the basis vectors in $V(\lambda)$ by taking its further restrictions to the subsequent subgroups of the chain (1.1). Namely, the basis vectors will be parametrized by sequences of partitions

$$\lambda^{(1)} \to \lambda^{(2)} \to \cdots \to \lambda^{(n)} = \lambda,$$

where $\lambda^{(k)}$ is obtained from $\lambda^{(k+1)}$ by removing one box. Equivalently, each sequence of this type can be regarded as a standard tableau of shape $\lambda$ which is obtained by writing the numbers $1, \ldots, n$ into the boxes of $\lambda$ in such a way that the numbers increase along the rows and down the columns. In particular, the dimension of $V(\lambda)$ equals the number of standard tableaux of shape $\lambda$. There is only one irreducible representation of the trivial group $\mathfrak{S}_1$ therefore the procedure defines basis vectors up to a scalar factor. The corresponding basis is called the Young basis. The symmetric group $\mathfrak{S}_n$ is generated by the adjacent transpositions $s_i = (i, i + 1)$. The construction of the representation $V(\lambda)$ can be completed by deriving explicit formulas for the action of the elements $s_i$ in the basis which are also due to A. Young. This realization of $V(\lambda)$ is usually called Young’s orthogonal (or seminormal) form. The details can be found, e.g., in James and Kerber, [63], and Sagan, [141]; see also Okounkov and Vershik, [112], where an alternative construction of the Young basis is produced. Branching rules and the corresponding Bratteli diagrams were employed by Halverson and Ram, [56], Leduc and Ram, [79], Ram, [139], to compute irreducible representations of the Iwahori–Hecke algebras and some families of centralizer algebras.
Quite a similar method can be applied to representations of the classical Lie algebras. Consider the general linear Lie algebra $\mathfrak{gl}_n$ which consists of complex $n \times n$-matrices with the usual matrix commutator. The chain (1.1) is now replaced by

$$\mathfrak{gl}_1 \subset \mathfrak{gl}_2 \subset \cdots \subset \mathfrak{gl}_n,$$

with natural embeddings $\mathfrak{gl}_k \subset \mathfrak{gl}_{k+1}$. The orthogonal Lie algebra $\mathfrak{o}_N$ can be regarded as a subalgebra of $\mathfrak{gl}_N$ which consists of skew-symmetric matrices. Again, we have a natural chain

$$\mathfrak{o}_2 \subset \mathfrak{o}_3 \subset \cdots \subset \mathfrak{o}_N.$$  \hspace{1cm} (1.2)

Both restrictions $\mathfrak{gl}_n \downarrow \mathfrak{gl}_{n-1}$ and $\mathfrak{o}_N \downarrow \mathfrak{o}_{N-1}$ are multiplicity-free so that the application of the argument which we used for the chain (1.1) produces basis vectors in an irreducible representation of $\mathfrak{gl}_n$ or $\mathfrak{o}_N$. With an appropriate normalization, these bases are precisely those of Gelfand and Tsetlin given in [41] and [42]. Instead of the standard tableaux, the basis vectors here are parametrized by combinatorial objects called the Gelfand–Tsetlin patterns.

However, this approach does not work for the symplectic Lie algebras $\mathfrak{sp}_{2n}$ since the restriction $\mathfrak{sp}_{2n} \downarrow \mathfrak{sp}_{2n-2}$ is not multiplicity-free. The multiplicities are given by Zhelobenko’s branching rule, [167], which was re-discovered later by Hegerfeldt, [58]. Various attempts to fix this problem were made by several authors. A natural idea is to introduce an intermediate Lie algebra “$\mathfrak{sp}_{2n-1}$” and try to restrict an irreducible representation of $\mathfrak{sp}_{2n}$ first to this subalgebra and then to $\mathfrak{sp}_{2n-2}$ in the hope to get simple spectra in the two restrictions. Such intermediate subalgebras and their representations were studied by Gelfand and Zelevinsky, [43], Proctor, [136], Shtepin, [142]. The drawback of this approach is the fact that the Lie algebra $\mathfrak{sp}_{2n-1}$ is not reductive so that the restriction of an irreducible representation of $\mathfrak{sp}_{2n}$ to $\mathfrak{sp}_{2n-1}$ is not completely reducible. In some sense, the separation of multiplicities can be achieved by constructing a filtration of $\mathfrak{sp}_{2n-1}$-modules; cf. Shtepin, [142].

Another idea is to use the restriction $\mathfrak{gl}_{2n} \downarrow \mathfrak{sp}_{2n}$. Gould and Kalnins, [51,53], constructed a basis for the representations of the symplectic Lie algebras parametrized by a subset of the Gelfand–Tsetlin $\mathfrak{gl}_{2n}$-patterns. Some matrix element formulas are also derived by using the $\mathfrak{gl}_{2n}$-action. A similar observation is made independently by Kirillov, [71], and Proctor, [136]. A description of the Gelfand–Tsetlin patterns for $\mathfrak{sp}_{2n}$ and $\mathfrak{o}_N$ can be obtained by regarding them as fixed points of involutions of the Gelfand–Tsetlin patterns for the corresponding Lie algebra $\mathfrak{gl}_N$.

The lowering operators in the symplectic case were given by Mickelsson, [91]; see also Bincer, [9,10]. The application of ordered monomials in the lowering operators to the highest vector yields a basis of the representation. However, the action of the Lie algebra generators in such a basis does not seem to be computable. The reason is the fact that, unlike the cases of $\mathfrak{gl}_n$ and $\mathfrak{o}_N$, the lowering operators do not commute so that the basis depends on the chosen ordering. A “hidden symmetry” has been needed (cf. Cherednik, [17]) to make

---

3 Some western authors referred to Hegerfeldt’s result as the original derivation of the rule.
a natural choice of an appropriate combination of the lowering operators. New ideas which led to a construction of a Gelfand–Tsetlin type basis for any irreducible finite-dimensional representation of $\mathfrak{sp}_{2n}$ came from the theory of quantized enveloping algebras. This is a part of the theory of quantum groups originating from the work of Drinfeld, [28,30], and Jimbo, [64]. A particular class of quantized enveloping algebras called twisted Yangians introduced by Olshanski, [118], plays the role of the hidden symmetries for the construction of the basis. We refer the reader to the book by Chari and Pressley, [14], and the review papers [103] and [104] for detailed expositions of the properties of these algebras and their origins. For each classical Lie algebra we attach the Yangian $Y(N) = Y(\mathfrak{gl}_N)$, or the twisted Yangian $Y^{\pm}(N)$ as follows

$$
\begin{array}{cccc}
type A_n & type B_n & type C_n & type D_n \\
Y(n+1) & Y^+(2n+1) & Y^-(2n) & Y^+(2n).
\end{array}
$$

The algebra $Y(N)$ was first introduced in the work of Faddeev and the St.-Petersburg school in relation with the inverse scattering method; see for instance Takhtajan and Faddeev, [147]. Kulish and Sklyanin, [74]. Olshanski, [118], introduced the twisted Yangians in relation with his centralizer construction; see also [105]. In particular, he established the following key fact which plays an important role in the basis construction. Given irreducible representations $V(\lambda)$ and $V'(\mu)$ of $\mathfrak{sp}_{2n}$ and $\mathfrak{sp}_{2n-2}$, respectively, there exists a natural irreducible action of the algebra $Y^-(2)$ on the space $\text{Hom}_{\mathfrak{sp}_{2n-2}}(V'(\mu), V(\lambda))$. The homomorphism space is isomorphic to the subspace $V(\lambda)^{\mu}$ of $V(\lambda)$ which is spanned by the highest vectors of weight $\mu$ for the subalgebra $\mathfrak{sp}_{2n-2}$. Finite-dimensional irreducible representations of the twisted Yangians were classified later in [97]. In particular, it turned out that the representation $V(\lambda)^{\mu}$ of $Y^-(2)$ can be extended to the Yangian $Y(2)$. Another proof of this fact was given recently by Nazarov, [107]. The algebra $Y(2)$ and its representations are well-studied; see Tarasov, [149], Chari and Pressley, [13]. A large class of representations of $Y(2)$ admits Gelfand–Tsetlin-type bases associated with the inclusion $Y(1) \subset Y(2)$; see [96]. This allows one to get a natural basis in the space $V(\lambda)^{\mu}$ and then by induction to get a basis in the entire space $V(\lambda)$. Moreover, it turns out to be possible to write down explicit formulas for the action of the generators of the symplectic Lie algebra in this basis; see the author’s paper [98] for more details. This construction together with the work of Gelfand and Tsetlin thus provides explicit realizations of all finite-dimensional irreducible representations of the classical Lie algebras.

The same method can be applied to the pairs of the orthogonal Lie algebras $\mathfrak{o}_{N-2} \subset \mathfrak{o}_N$. Here the corresponding space $V(\lambda)^{\mu}$ is a natural $Y^+(2)$-module which can also be extended to a $Y(2)$-module. This leads to a construction of a natural basis in the representation $V(\lambda)$ and allows one to explicitly calculate the representation matrices; see [99,100]. This realization of $V(\lambda)$ is alternative to that of Gelfand and Tsetlin, [42]. To compare the two constructions, note that the basis of [42] in the orthogonal case lacks the weight property, i.e., the basis vectors are not eigenvectors for the Cartan subalgebra. The reason for that is the fact that the chain (1.2) involves Lie algebras of different types ($B$ and $D$) and the embeddings are not compatible with the root systems. In the new approach we use instead the chains

$$
\mathfrak{o}_2 \subset \mathfrak{o}_4 \subset \cdots \subset \mathfrak{o}_{2n} \quad \text{and} \quad \mathfrak{o}_3 \subset \mathfrak{o}_5 \subset \cdots \subset \mathfrak{o}_{2n+1}.
$$
The embeddings here “respect” the root systems so that the basis of \( V(\lambda) \) possesses the weight property in both the symplectic and orthogonal cases. However, the new weight bases, in turn, lack the orthogonality property of the Gelfand–Tsetlin bases: the latter are orthogonal with respect to the standard inner product in the representation space \( V(\lambda) \). It is an open problem to construct a natural basis of \( V(\lambda) \) in the \( B, C \) and \( D \) cases which would simultaneously accommodate the two properties.

This chapter is structured as follows. In Section 2 we review the construction of the Gelfand–Tsetlin basis for the general linear Lie algebra and discuss its various versions. We start by applying the most elementary approach which consists of using explicit formulas for the lowering operators in a way similar to the pioneering works of the sixties. Remarkably, these operators admit several other presentations which reflect different approaches to the problem developed in the literature. First, we outline the general theory of extremal projectors and Mickelsson algebras as a natural way to work with lowering operators. Next, we describe the \( \mathfrak{gl}_n \)-type Mickelsson–Zhelobenko algebra which is then used to prove the branching rule and derive the matrix element formulas. Further, we outline the Gould construction based upon the characteristic identities. Finally, we produce quantum minor formulas for the lowering operators inspired by the Yangian approach and describe the action of the Drinfeld generators in the Gelfand–Tsetlin basis.

In Section 3 we produce weight bases for representations of the orthogonal and symplectic Lie algebras. Here we describe the relevant Mickelsson–Zhelobenko algebra, formulate the branching rules and construct the basis vectors. Then we outline the properties of the (twisted) Yangians and their representations and explain their relationship with the lowering and raising operators. Finally, we sketch the main ideas in the calculation of the matrix element formulas.

Section 4 is devoted to the Gelfand–Tsetlin bases for the orthogonal Lie algebras. We outline the basis construction along the lines of the general method of Mickelsson algebras.

At the end of each section we give brief bibliographical comments pointing towards the original articles and to the references where the proofs or further details can be found.

2. Gelfand–Tsetlin bases for representations of \( \mathfrak{gl}_n \)

Let \( E_{ij}, i, j = 1, \ldots, n \), denote the standard basis of the general linear Lie algebra \( \mathfrak{gl}_n \) over the field of complex numbers. The subalgebra \( \mathfrak{gl}_{n-1} \) is spanned by the basis elements \( E_{ij} \) with \( i, j = 1, \ldots, n - 1 \). Denote by \( \mathfrak{h} = \mathfrak{h}_n \) the diagonal Cartan subalgebra in \( \mathfrak{gl}_n \). The elements \( E_{11}, \ldots, E_{nn} \) form a basis of \( \mathfrak{h} \).

Finite-dimensional irreducible representations of \( \mathfrak{gl}_n \) are in a one-to-one correspondence with \( n \)-tuples of complex numbers \( \lambda = (\lambda_1, \ldots, \lambda_n) \) such that

\[
\lambda_i - \lambda_{i+1} \in \mathbb{Z}_+ \quad \text{for} \quad i = 1, \ldots, n - 1.
\]  

(2.1)

Here \( \mathbb{Z}_+ = \{ i \in \mathbb{Z}: i \geq 0 \} \).
Such an $n$-tuple $\lambda$ is called the highest weight of the corresponding representation which we shall denote by $L(\lambda)$. It contains a unique, up to a multiple, nonzero vector $\xi$ (the highest weight vector (highest vector)) such that $E_{ii} \xi = \lambda_i \xi$ for $i = 1, \ldots, n$ and $E_{ij} \xi = 0$ for $1 \leq i < j \leq n$.

The following theorem is the branching rule for the reduction $\mathfrak{gl}_n \downarrow \mathfrak{gl}_{n-1}$.

**Theorem 2.1.** The restriction of $L(\lambda)$ to the subalgebra $\mathfrak{gl}_{n-1}$ is isomorphic to the direct sum of pairwise inequivalent irreducible representations

$$L(\lambda)|_{\mathfrak{gl}_{n-1}} \cong \bigoplus_{\mu} L'(\mu),$$

summed over the highest weights $\mu$ satisfying the betweenness conditions

$$\lambda_i - \mu_i \in \mathbb{Z}_+ \text{ and } \mu_i - \lambda_{i+1} \in \mathbb{Z}_+ \text{ for } i = 1, \ldots, n-1.$$  

(2.2)

The rule can presumably be attributed to I. Schur who was the first to discover the representation-theoretic significance of a particular class of symmetric polynomials which now bear his name. Without loss of generality we may regard $\lambda$ as a partition: we can take the composition of $L(\lambda)$ with an appropriate automorphism of $U(\mathfrak{gl}_n)$ which sends $E_{ij}$ to $E_{ij} + \delta_{ija}$ for some $a \in \mathbb{C}$. The character of $L(\lambda)$ regarded as a $\text{GL}_n$-module is the Schur polynomial $s_\lambda(x)$, $x = (x_1, \ldots, x_n)$, defined by

$$s_\lambda(x) = \text{tr}(g, L(\lambda)),$$

where $x_1, \ldots, x_n$ are the eigenvalues of $g \in \text{GL}_n$. The Schur polynomial is symmetric in the $x_i$ and can be given by the explicit combinatorial formula

$$s_\lambda(x) = \sum_T x^T,$$

(2.3)

summed over the semistandard tableaux $T$ of shape $\lambda$ (cf. Remark 2.2 below), where $x^T$ is the monomial containing $x_i$ with the power equal to the number of occurrences of $i$ in $T$; see, e.g., Macdonald, [84, Chapter 1], or Sagan, [141, Chapter 4], for more details. To find out what happens when $L(\lambda)$ is restricted to $\text{GL}_{n-1}$ we just need to put $x_n = 1$ into formula (2.3). The right-hand side will then be written as the sum of the Schur polynomials $s_\mu(x_1, \ldots, x_{n-1})$ with $\mu$ satisfying (2.2).

On the other hand, the multiplicity-freeness of the reduction $\mathfrak{gl}_n \downarrow \mathfrak{gl}_{n-1}$ can be explained by the fact that the vector space $\text{Hom}_{\mathfrak{gl}_{n-1}}(L'(\mu), L(\lambda))$ bears a natural irreducible representation of the centralizer $U(\mathfrak{gl}_n)$; see, e.g., Dixmier, [19, Section 9.1]. However, the centralizer is a commutative algebra and therefore if the homomorphism space is nonzero then it must be one-dimensional.

The branching rule is implicit in the formulas of Gelfand and Tsetlin, [41]. A proof based upon an explicit realization of the representations of $\text{GL}_n$ was given by Zhelobenko, [167]. We outline a proof of Theorem 2.1 below in Section 2.3 which employs the modern theory
of Mickelson algebras following Zhelobenko, [174]. Two other proofs can be found in Goodman and Wallach, [45, Chapters 8 and 12].

Successive applications of the branching rule to the subalgebras of the chain
\[
\mathfrak{gl}_1 \subset \mathfrak{gl}_2 \subset \cdots \subset \mathfrak{gl}_{n-1} \subset \mathfrak{gl}_n
\]
yield a parameterization of basis vectors in \(L(\lambda)\) by the combinatorial objects called the Gelfand–Tsetlin patterns. Such a pattern \(\Lambda\) (associated with \(\lambda\)) is an array of row vectors
\[
\begin{array}{cccc}
\lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \\
\lambda_{n-1,1} & \cdots & \lambda_{n-1,n-1} \\
\cdots & \cdots & \cdots \\
\lambda_{21} & \lambda_{22} & \cdots & \cdots \\
\lambda_{11} & & & \\
\end{array}
\]
where the upper row coincides with \(\lambda\) and the following conditions hold
\[
\lambda_{ki} - \lambda_{k-1,i} \in \mathbb{Z}_+, \quad \lambda_{k-1,i} - \lambda_{k,i+1} \in \mathbb{Z}_+, \quad i = 1, \ldots, k - 1, \quad (2.4)
\]
for each \(k = 2, \ldots, n\).

**Remark 2.2.** If the highest weight \(\lambda\) is a partition then there is a natural bijection between the patterns associated with \(\lambda\) and the semistandard \(\lambda\)-tableaux with entries in \(\{1, \ldots, n\}\). Namely, the pattern \(\Lambda\) can be viewed as the sequence of partitions
\[
\lambda^{(1)} \subset \lambda^{(2)} \subset \cdots \subset \lambda^{(n)} = \lambda,
\]
with \(\lambda^{(k)} = (\lambda_{k1}, \ldots, \lambda_{kk})\). Conditions (2.4) mean that the skew diagram \(\lambda^{(k)}/\lambda^{(k-1)}\) is a horizontal strip; see, e.g., Macdonald, [84, Chapter 1]. The corresponding semistandard tableau is obtained by placing the entry \(k\) into each box of \(\lambda^{(k)}/\lambda^{(k-1)}\).

The Gelfand–Tsetlin basis of \(L(\lambda)\) is provided by the following theorem. Let us set \(l_{ki} = \lambda_{ki} - i + 1\).

**Theorem 2.3.** There exists a basis \(\{\xi_{\Lambda}\}\) in \(L(\lambda)\) parametrized by all patterns \(\Lambda\) such that the action of generators of \(\mathfrak{gl}_n\) is given by the formulas
\[
E_{kk}\xi_{\Lambda} = \left(\sum_{i=1}^{k} \lambda_{ki} - \sum_{i=1}^{k-1} \lambda_{k-1,i}\right)\xi_{\Lambda}, \quad (2.5)
\]
\[
E_{k,k+1}\xi_{\Lambda} = -\sum_{i=1}^{k} \frac{(l_{ki} - l_{k+1,1}) \cdots (l_{ki} - l_{k,k+1})}{(l_{ki} - l_{k1}) \cdots (l_{ki} - l_{kk})} \xi_{\Lambda + \delta_{ki}}, \quad (2.6)
\]
\[
E_{k+1,k}\xi_{\Lambda} = \sum_{i=1}^{k} \frac{(l_{ki} - l_{k-1,1}) \cdots (l_{ki} - l_{k-1,k-1})}{(l_{ki} - l_{k1}) \cdots (l_{ki} - l_{kk})} \xi_{\Lambda - \delta_{ki}}. \quad (2.7)
\]
The arrays $\Lambda \pm \delta_{ki}$ are obtained from $\Lambda$ by replacing $\lambda_{ki}$ by $\lambda_{ki} \pm 1$. It is supposed that $\xi_\Lambda = 0$ if the array $\Lambda$ is not a pattern; the symbol $\wedge$ indicates that the zero factor in the denominator is skipped.

A construction of the basis vectors is given in Theorem 2.7 below. A derivation of the matrix element formulas (2.5)–(2.7) is outlined in Section 2.3.

The vector space $L(\lambda)$ is equipped with a contravariant inner product $\langle \cdot, \cdot \rangle$. It is uniquely determined by the conditions

$$\langle \xi, \xi \rangle = 1 \quad \text{and} \quad \langle E_{ij} \eta, \zeta \rangle = \langle \eta, E_{ji} \zeta \rangle$$

for any vectors $\eta, \zeta \in L(\lambda)$ and any indices $i, j$. In other words, for the adjoint operator for $E_{ij}$ with respect to the inner product we have $(E_{ij})^* = E_{ji}$.

**Proposition 2.4.** The basis $\{\xi_\Lambda\}$ is orthogonal with respect to the inner product $\langle \cdot, \cdot \rangle$. Moreover, we have

$$\langle \xi_\Lambda, \xi_\Lambda \rangle = \prod_{k=2}^{n} \prod_{1 \leq i < j < k} \frac{(l_{ki} - l_{k-1,j})!}{(l_{k-1,i} - l_{k-1,j})!} \prod_{1 \leq l < j \leq k} \frac{(l_{ki} - l_{kj} - 1)!}{(l_{k-1,i} - l_{kj} - 1)!}.$$

The formulas of Theorem 2.3 can therefore be rewritten in the orthonormal basis

$$\xi_\Lambda = \xi_\Lambda / \|\xi_\Lambda\|, \quad \|\xi_\Lambda\|^2 = \langle \xi_\Lambda, \xi_\Lambda \rangle. \quad (2.8)$$

They were presented in this form in the original work by Gelfand and Tsetlin, [41]. A proof of Proposition 2.4 will be outlined in Section 2.3.

### 2.1. Construction of the basis: lowering and raising operators

For each $i = 1, \ldots, n - 1$ introduce the following elements of the universal enveloping algebra $U(\mathfrak{gl}_n)$

$$z_{in} = \sum_{i > i_1 > \cdots > i_s \geq 1} E_{ii_1} E_{i_1 i_2} \cdots E_{i_s-1 i_s} E_{i_s n} (h_i - h_{j_1}) \cdots (h_i - h_{j_s}), \quad (2.9)$$

$$z_{ni} = \sum_{i < i_1 < \cdots < i_s < n} E_{ii_1} E_{i_1 i_2} \cdots E_{i_s i_{s+1}} E_{ni_s} (h_i - h_{j_1}) \cdots (h_i - h_{j_s}), \quad (2.10)$$

where $s$ runs over nonnegative integers, $h_i = E_{ii} - i + 1$ and $\{j_1, \ldots, j_s\}$ is the complementary subset to $\{i_1, \ldots, i_s\}$ in the set $\{1, \ldots, i-1\}$ or $\{i+1, \ldots, n-1\}$, respectively. For instance,

$$z_{13} = E_{13}, \quad z_{23} = E_{23} (h_2 - h_1) + E_{21} E_{13},$$

$$z_{32} = E_{32}, \quad z_{31} = E_{31} (h_1 - h_2) + E_{21} E_{32}.$$
Consider now the irreducible finite-dimensional representation $L(\lambda)$ of $\mathfrak{gl}_n$ with the highest weight $\lambda = (\lambda_1, \ldots, \lambda_n)$ and the highest vector $\xi$. Denote by $L(\lambda)^+$ the subspace of $\mathfrak{gl}_{n-1}$-highest vectors in $L(\lambda)$:

$$L(\lambda)^+ = \{ \eta \in L(\lambda) \mid E_{ij} \eta = 0, \ 1 \leq i < j < n \}.$$ 

Given a $\mathfrak{gl}_{n-1}$-weight $\mu = (\mu_1, \ldots, \mu_{n-1})$ we denote by $L(\lambda)_{\mu}^+$ the corresponding weight subspace in $L(\lambda)^+$:

$$L(\lambda)_{\mu}^+ = \{ \eta \in L(\lambda)^+ \mid E_{ii} \eta = \mu_i \eta, \ i = 1, \ldots, n-1 \}.$$ 

The main property of the elements $z_{ni}$ and $z_{ni}$ is described by the following lemma.

**Lemma 2.5.** Let $\eta \in L(\lambda)_{\mu}^+$. Then for any $i = 1, \ldots, n-1$ we have

$$z_{in} \eta \in L(\lambda)_{\mu+\delta_i}^+ \quad \text{and} \quad z_{ni} \eta \in L(\lambda)_{\mu-\delta_i}^+,$$

where the weight $\mu \pm \delta_i$ is obtained from $\mu$ by replacing $\mu_i$ with $\mu_i \pm 1$.

This result allows us to regard the elements $z_{in}$ and $z_{ni}$ as operators in the space $L(\lambda)^+$. They are called the raising and lowering operators, respectively. By the branching rule (Theorem 2.1) the space $L(\lambda)_{\mu}^+$ is one-dimensional if the conditions (2.2) hold and it is zero otherwise. The following lemma will be proved in Section 2.3.

**Lemma 2.6.** Suppose that $\mu$ satisfies the betweenness conditions (2.2). Then the vector

$$\xi_{\mu} = z_{n1}^{\lambda_1-\mu_1} \cdots z_{n, n-1}^{\lambda_{n-1}-\mu_{n-1}} \xi,$$

is nonzero. Moreover, the space $L(\lambda)_{\mu}^+$ is spanned by $\xi_{\mu}$.

The $\mathfrak{u}(\mathfrak{gl}_{n-1})$-span of each nonzero vector $\xi_{\mu}$ is a $\mathfrak{gl}_{n-1}$-module isomorphic to $L'_{\mu}$. Iterating the construction of the vectors $\xi_{\mu}$ for each pair of Lie algebras $\mathfrak{gl}_{k-1} \subset \mathfrak{gl}_k$ we shall be able to get a basis in the entire space $L(\lambda)$.

**Theorem 2.7.** The basis vectors $\xi_{\lambda}$ of Theorem 2.3 can be given by the formula

$$\xi_{\lambda} = \prod_{k=2, \ldots, n} \left( z_{k1}^{\lambda_{k1}-\lambda_{k-1,1}} \cdots z_{k, k-1}^{\lambda_{k, k-1}-\lambda_{k-1, k-1}} \right) \xi,$$

where the factors in the product are ordered according to increasing indices.
2.2. Mickelsson algebra theory

The lowering and raising operators $z_{ni}$ and $z_{in}$ in the space $L(\lambda)^+$ (see Lemma 2.5) satisfy some quadratic relations with rational coefficients in the parameters of the highest weights. These relations can be regarded in a representation independent form with a suitable interpretation of the coefficients as rational functions in the elements of the Cartan subalgebra $\mathfrak{h}$. In this abstract form the algebras of lowering and raising operators were introduced by Mickelsson, [92], who, however, did not use any rational extensions of the algebra $U(\mathfrak{h})$. The importance of this extension was realized by Zhelobenko, [169,170], who developed a general structure theory of these algebras which he called Mickelsson algebras. Another important ingredient is the theory of extremal projectors which originated in the work of Asherova, Smirnov and Tolstoy, [1–3], and was further developed by Zhelobenko, [173,174].

Let $\mathfrak{g}$ be a Lie algebra over $\mathbb{C}$ and let $\mathfrak{k}$ be a subalgebra reductive in $\mathfrak{g}$. This means that the adjoint $\mathfrak{k}$-module $\mathfrak{g}$ is completely reducible. In particular, $\mathfrak{k}$ is a reductive Lie algebra. Fix a Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{k}$ and a triangular decomposition

$$\mathfrak{k} = \mathfrak{k}^- \oplus \mathfrak{h} \oplus \mathfrak{k}^+.$$ 

The subalgebras $\mathfrak{k}^-$ and $\mathfrak{k}^+$ are respectively spanned by the negative and positive root vectors $e_{-\alpha}$ and $e_{\alpha}$ with $\alpha$ running over the set of positive roots $\Delta^+$ of $\mathfrak{k}$ with respect to $\mathfrak{h}$. The root vectors will be assumed to be normalized in such a way that

$$[e_\alpha, e_{-\alpha}] = h_\alpha, \quad \alpha(h_\alpha) = 2$$  \hspace{1cm} (2.12)

for all $\alpha \in \Delta^+$.

Let $J = U(\mathfrak{g})\mathfrak{k}^+$ be the left ideal of $U(\mathfrak{g})$ generated by $\mathfrak{k}^+$. Its normalizer $\text{Norm} J$ is the subalgebra of $U(\mathfrak{g})$ defined by

$$\text{Norm} J = \{ u \in U(\mathfrak{g}) \mid Ju \subseteq J \}.$$ 

Then $J$ is a two-sided ideal of $\text{Norm} J$ and the Mickelsson algebra $S(\mathfrak{g}, \mathfrak{k})$ is defined as the quotient

$$S(\mathfrak{g}, \mathfrak{k}) = \text{Norm} J / J.$$ 

Let $R(\mathfrak{h})$ denote the field of fractions of the commutative algebra $U(\mathfrak{h})$. In what follows it is convenient to consider the extension $U'(\mathfrak{g})$ of the universal enveloping algebra $U(\mathfrak{g})$ defined by

$$U'(\mathfrak{g}) = U(\mathfrak{g}) \otimes_{U(\mathfrak{h})} R(\mathfrak{h}).$$
Let $J' = U'(g)t^+$ be the left ideal of $U'(g)$ generated by $t^+$. Exactly as with the ideal $J$ above, $J'$ is a two-sided ideal of the normalizer $\text{Norm} J'$ and the Mickelsson–Zhelobenko algebra $Z(g, t)$ is defined as the quotient

$$Z(g, t) = \text{Norm} J'/J'.$$

Clearly, $Z(g, t)$ is an extension of the Mickelsson algebra $S(g, k)$,

$$Z(g, t) = S(g, t) \otimes U(h) R(h).$$

An equivalent definition of the algebra $Z(g, t)$ can be given by using the quotient space

$$M(g, t) = U'(g)/J'.$$

The Mickelsson–Zhelobenko algebra $Z(g, t)$ coincides with the subspace of $t$-highest vectors in $M(g, t)$

$$Z(g, t) = M(g, t)^+, $$

where

$$M(g, t)^+ = \{ v \in M(g, t) \mid t^+ v = 0 \}.$$ 

The algebraic structure of the algebra $Z(g, t)$ can be described with the use of the extremal projector for the Lie algebra $t$. In order to define it, suppose that the positive roots are $\Delta^+ = \{ \alpha_1, \ldots, \alpha_m \}$. Consider the vector space $F_\mu(t)$ of formal series of weight $\mu$ monomials

$$e_{-\alpha_1}^{k_1} \cdots e_{-\alpha_m}^{k_m} e_{\alpha_m}^{r_m} \cdots e_{\alpha_1}^{r_1}$$

with coefficients in $R(h)$, where

$$(r_1 - k_1)\alpha_1 + \cdots + (r_m - k_m)\alpha_m = \mu.$$

Introduce the space $F(t)$ as the direct sum

$$F(t) = \bigoplus_\mu F_\mu(t).$$

That is, the elements of $F(t)$ are finite sums $\sum x_\mu$ with $x_\mu \in F_\mu(t)$. It can be shown that $F(t)$ is an algebra with respect to the natural multiplication of formal series. The algebra

$^{4}$Zhelobenko sometimes used the names $Z$-algebra or extended Mickelsson algebra. The author believes the new name is more appropriate and justified from the scientific point of view.
F(\mathfrak{t}) is equipped with a Hermitian anti-involution (antilinear involutive anti-automorphism) defined by
\[ e_\alpha^* = e_{-\alpha}, \quad \alpha \in \Delta^+. \]

Further, call an ordering of the positive roots *normal* if any composite root lies between its components. For instance, there are precisely two normal orderings for the root system of type $B_2$,
\[ \Delta^+ = \{ \alpha, \alpha + \beta, \alpha + 2\beta, \beta \} \quad \text{and} \quad \Delta^+ = \{ \beta, \alpha + 2\beta, \alpha + \beta, \alpha \}, \]
where $\alpha$ and $\beta$ are the simple roots. In general, the number of normal orderings coincides with the number of reduced decompositions of the longest element of the corresponding Weyl group.

For any $\alpha \in \Delta^+$ introduce an element of $F(\mathfrak{t})$ by
\[ p_\alpha = 1 + \sum_{k=1}^{\infty} e_{-\alpha}^k e_\alpha^k (-1)^k k!(h_\alpha + \rho(h_\alpha) + 1) \cdots (h_\alpha + \rho(h_\alpha) + k), \tag{2.13} \]
where $h_\alpha$ is defined in (2.12) and $\rho$ is the half sum of the positive roots. Finally, define the extremal projector $p = p_\mathfrak{t}$ by
\[ p = p_{\alpha_1} \cdots p_{\alpha_n} \]
with the product taken in a normal ordering of the positive roots $\alpha_i$.

**Theorem 2.8.** The element $p \in F(\mathfrak{t})$ does not depend on the normal ordering on $\Delta^+$ and satisfies the conditions
\[ e_\alpha p = pe_{-\alpha} = 0 \quad \text{for all } \alpha \in \Delta^+. \tag{2.14} \]
Moreover, $p^* = p$ and $p^2 = p$.

In fact, the relations (2.14) uniquely determine the element $p$, up to a factor from $R(\mathfrak{t})$. The extremal projector naturally acts on the vector space $M(\mathfrak{g}, \mathfrak{t})$. The following corollary states that the Mickelsson–Zhelobenko algebra coincides with its image.

**Corollary 2.9.** We have
\[ Z(\mathfrak{g}, \mathfrak{t}) = pM(\mathfrak{g}, \mathfrak{t}). \]

To get a more precise description of the algebra $Z(\mathfrak{g}, \mathfrak{t})$ consider a $\mathfrak{t}$-module decomposition
\[ \mathfrak{g} = \mathfrak{t} \oplus \mathfrak{p}. \]
Choose a weight basis $e_1, \ldots, e_n$ (with respect to the adjoint action of $\mathfrak{h}$) of the complementary module $\mathfrak{p}$.

**Theorem 2.10.** The elements 

$$a_i = p e_i, \quad i = 1, \ldots, n,$$

are generators of the Mickelsson–Zhelobenko algebra $Z(\mathfrak{g}, \mathfrak{t})$. Moreover, the monomials 

$$a_1^{k_1} \cdots a_n^{k_n}, \quad k_i \in \mathbb{Z}_+,$$

form a basis of $Z(\mathfrak{g}, \mathfrak{t})$.

It can be proved that the generators $a_i$ of $Z(\mathfrak{g}, \mathfrak{t})$ satisfy quadratic defining relations; see [173]. For the pairs $(\mathfrak{g}, \mathfrak{t})$ relevant to the constructions of bases of Gelfand–Tsetlin type, the relations can be explicitly written down; cf. Sections 2.3 and 3.1 below.

Regarding $Z(\mathfrak{g}, \mathfrak{t})$ as a right $R(\mathfrak{h})$-module, it is possible to introduce the normalized elements 

$$z_i = a_i \pi_i, \quad \pi_i \in U(\mathfrak{h}),$$

by multiplying $a_i$ by its right denominator $\pi_i$. Therefore the $z_i$ can be viewed as elements of the Mickelsson algebra $S(\mathfrak{g}, \mathfrak{t})$.

To formulate the final theorem of this section, for any $\mathfrak{g}$-module $V$ set 

$$V^+ = \{ v \in V \mid \mathfrak{t}^+ v = 0 \}.$$

**Theorem 2.11.** Let $V = U(\mathfrak{g}) v$ be the cyclic $U(\mathfrak{g})$-module generated by an element $v \in V^+$. Then the subspace $V^+$ is linearly spanned by the elements 

$$z_1^{k_1} \cdots z_n^{k_n} v, \quad k_i \in \mathbb{Z}_+.$$

**2.3. Mickelsson–Zhelobenko algebra $Z(\mathfrak{gl}_n, \mathfrak{gl}_{n-1})$**

For any positive integer $m$ consider the general linear Lie algebra $\mathfrak{gl}_m$. The positive roots of $\mathfrak{gl}_m$ with respect to the diagonal Cartan subalgebra $\mathfrak{h}$ (with the standard choice of the positive root system) are naturally enumerated by the pairs $(i, j)$ with $1 \leq i < j \leq m$. In accordance with the general theory outlined in the previous section, for each pair introduce a formal series $p_{ij} \in F(\mathfrak{gl}_m)$ by 

$$p_{ij} = 1 + \sum_{k=1}^{\infty} (E_{ji})^k (E_{ij})^k \frac{(-1)^k}{k!(h_i - h_j + 1) \cdots (h_i - h_j + k)},$$

where $E_{ij}$ are the elementary root vectors.
where, as before, $h_i = E_{ii} - i + 1$. Then define the element $p = p_m$ by

$$p = \prod_{i<j} p_{ij},$$

where the product is taken in a normal ordering on the pairs $(i, j)$. By Theorem 2.8,

$$E_{ij} p = p E_{ji} = 0 \quad \text{for } 1 \leq i < j \leq m. \quad (2.15)$$

Now set $m = n - 1$. By Theorem 2.10, ordered monomials in the elements $E_{nn}$, $p E_{in}$ and $p E_{ni}$ with $i = 1, \ldots, n - 1$ form a basis of $Z(\mathfrak{gl}_n, \mathfrak{gl}_{n-1})$ as a left or right $R(\mathfrak{h})$-module. These elements can explicitly be given by

$$p E_{in} = \sum_{i > i_1 > \cdots > i_s \geq 1} E_{i_1 i_2} E_{i_1 i_3} \cdots E_{i_s i_{s-1}} E_{i_{s+1} n} \frac{1}{(h_i - h_{i_1}) \cdots (h_i - h_{i_s})},$$

$$p E_{ni} = \sum_{i < i_1 < \cdots < i_s < n} E_{i_1 i} E_{i_2 i} \cdots E_{i_{s+1} i_{s+2}} E_{i_{s+2} n} \frac{1}{(h_i - h_{i_1}) \cdots (h_i - h_{i_s})}, \quad (2.16)$$

where $s = 0, 1, \ldots$. Indeed, by choosing appropriate normal orderings on the positive roots, we can write

$$p E_{in} = p_{i_1} \cdots p_{i_{s-1}} E_{in} \quad \text{and} \quad p E_{ni} = p_{i, i+1} \cdots p_{i, n-1} E_{ni}. $$

The lowering and raising operators introduced in Section 2.1 coincide with the normalized generators:

$$z_{in} = p E_{in} (h_i - h_{i-1}) \cdots (h_i - h_1),$$

$$z_{ni} = p E_{ni} (h_i - h_{i+1}) \cdots (h_i - h_{n-1}), \quad (2.17)$$

which belong to the Mickelsson algebra $S(\mathfrak{gl}_n, \mathfrak{gl}_{n-1})$. Thus, Lemma 2.5 is an immediate corollary of (2.15).

**PROPOSITION 2.12.** The lowering and raising operators satisfy the following relations

$$z_{ni} z_{nj} = z_{nj} z_{ni} \quad \text{for all } i, j, \quad (2.18)$$

$$z_{in} z_{nj} = z_{nj} z_{in} \quad \text{for } i \neq j, \quad (2.19)$$

and

$$z_{in} z_{ni} = \prod_{j=1, j \neq i}^{n} (h_j - h_j - 1) + \sum_{j=1}^{n-1} z_{nj} z_{jn} \prod_{k=1, k \neq j}^{n-1} \frac{h_i - h_k - 1}{h_j - h_k}. \quad (2.20)$$
PROOF. We use the properties of $p$. Assume that $i < j$. Then (2.15) and (2.16) imply that in $Z(\mathfrak{gl}_n, \mathfrak{gl}_{n-1})$

$$pE_{ni}pE_{nj} = pE_{ni}E_{nj}, \quad pE_{nj}pE_{ni} = pE_{ni}E_{nj} \frac{h_i - h_j + 1}{h_i - h_j}.$$ 

Now (2.18) follows from (2.17). The proof of (2.19) is similar. The “long” relation (2.20) can be verified by analogous but more complicated direct calculations. We give a different proof based upon the properties of the Capelli determinant $C(u)$. Consider the $n \times n$-matrix

$$E$$

whose $ij$-th entry is $E_{ij}$ and let $u$ be a formal variable. Then $C(u)$ is the polynomial with coefficients in the universal enveloping algebra $U(\mathfrak{gl}_n)$ defined by

$$C(u) = \sum_{\sigma \in S_n} \text{sgn} \sigma \cdot (u + E)_{\sigma(1),1} \cdots (u + E - n + 1)_{\sigma(n),n}. \quad (2.21)$$

It is well known that all its coefficients belong to the center of $U(\mathfrak{gl}_n)$ and generate the center; see, e.g., Howe and Umeda, [60]. This also easily follows from the properties of the quantum determinant of the Yangian for the Lie algebra $\mathfrak{gl}_n$; see, e.g., [104]. Therefore, these coefficients act in $L(\lambda)$ as scalars which can be easily found by applying $C(u)$ to the highest vector $\xi$:

$$C(u)|_{L(\lambda)} = (u + l_1) \cdots (u + l_n), \quad l_i = \lambda_i - i + 1. \quad (2.22)$$

On the other hand, the center of $U(\mathfrak{gl}_n)$ is a subalgebra in the normalizer $\text{Norm} J$. We shall keep the same notation for the image of $C(u)$ in the Mickelsson–Zhelobenko algebra $Z(\mathfrak{gl}_n, \mathfrak{gl}_{n-1})$. To get explicit expressions of the coefficients of $C(u)$ in terms of the lowering and raising operators we consider $C(u)$ modulo the ideal $J'$ and apply the projection $p$. A straightforward calculation yields two alternative formulas

$$C(u) = (u + E_{nn}) \prod_{i=1}^{n-1} (u + h_i - 1) - \sum_{i=1}^{n-1} z_{in}z_{ni} \prod_{j=1, j \neq i}^{n-1} \frac{u + h_j - 1}{h_i - h_j} \quad (2.23)$$

and

$$C(u) = \prod_{i=1}^{n} (u + h_i) - \sum_{i=1}^{n-1} z_{ni}z_{in} \prod_{j=1, j \neq i}^{n-1} \frac{u + h_j}{h_i - h_j}. \quad (2.24)$$

The formulas show that $C(u)$ can be regarded as an interpolation polynomial for the products $z_{in}z_{ni}$ and $z_{ni}z_{in}$. Namely, for $i = 1, \ldots, n - 1$, we have

$$C(-h_i + 1) = (-1)^{n-1} z_{in}z_{ni} \quad \text{and} \quad C(-h_i) = (-1)^{n-1} z_{ni}z_{in} \quad (2.25)$$

with the agreement that when we evaluate $u$ in $U(h)$ we write the coefficients of the polynomial to the left from powers of $u$. Comparing the values of (2.23) and (2.24) at $u = -h_i + 1$ we get (2.20). □
Note that the relation inverse to (2.20) can be obtained by comparing the values of (2.23) and (2.24) at \( u = -h_i \).

Next we outline the proofs of the branching rule (Theorem 2.1) and the formulas for the basis elements of \( L(\lambda^+) \) (Lemma 2.6). The module \( L(\lambda) \) is generated by the highest vector \( \xi \) and we have

\[ z_{in} \xi = 0, \quad i = 1, \ldots, n - 1. \]

So, by Theorem 2.11, the vector space \( L(\lambda)^+ \) is spanned by the elements

\[ z_{n1}^{k_1} \cdots z_{n,n-1}^{k_{n-1}} \xi, \quad k_i \in \mathbb{Z}_+. \tag{2.26} \]

Let us set \( \mu_i = \lambda_i - k_i \) for \( 1 \leq i \leq n - 1 \) and denote the vector (2.26) by \( \xi_{\mu} \). That is,

\[ \xi_{\mu} = z_{n1}^{\lambda_1 - \mu_1} \cdots z_{n,n-1}^{\lambda_{n-1} - \mu_{n-1}} \xi. \tag{2.27} \]

It is now sufficient to show that the vector \( \xi_{\mu} \) is nonzero if and only if the betweenness conditions (2.2) hold. The linear independence of the vectors \( \xi_{\mu} \) will follow from the fact that their weights are distinct. If \( \xi_{\mu} \neq 0 \) then using the relations (2.18) we conclude that each vector \( z_{ni}^{\lambda_i - \mu_i} \xi \) is nonzero. On the other hand, \( z_{ni}^{k_i} \xi \) is a \( \mathfrak{gl}_{n-1} \)-highest vector of the weight obtained from \( (\lambda_1, \ldots, \lambda_{n-1}) \) by replacing \( \lambda_i \) with \( \lambda_i - k_i \). Therefore, if \( k_i \geq \lambda_i - \lambda_i + 1 + 1 \) then the conditions (2.1) are violated for this weight which implies \( z_{ni}^{k_i} \xi = 0 \). Hence, \( \lambda_i - \mu_i \leq \lambda_i - \lambda_i + 1 \) for each \( i \), and \( \mu \) satisfies (2.2).

For the proof of the converse statement we shall employ the following key lemma which will also be used for the proof of Theorem 2.3.

**LEMMA 2.13.** We have for each \( i = 1, \ldots, n - 1 \)

\[ z_{in} \xi_{\mu} = -(m_i - l_i) \cdots (m_i - l_n) \xi_{\mu + \delta_i}, \tag{2.28} \]

where

\[ m_i = \mu_i - i + 1, \quad l_i = \lambda_i - i + 1. \]

Here \( \xi_{\mu + \delta_i} = 0 \) if \( \lambda_i = \mu_i \).

**PROOF.** The relation (2.19) implies that if \( \lambda_i = \mu_i \) then \( z_{in} \xi_{\mu} = 0 \) which agrees with (2.28). Now let \( \lambda_i - \mu_i \geq 1 \). Using (2.18) and (2.25) we obtain

\[ z_{in} \xi_{\mu} = z_{in} z_{ni} \xi_{\mu + \delta_i} = (-1)^{n-1} c(-h_i + 1) \xi_{\mu + \delta_i} = (-1)^{n-1} c(-m_i) \xi_{\mu + \delta_i}. \]

The relation (2.28) now follows from (2.22) and the centrality of \( c(u) \). \( \square \)

If the betweenness conditions (2.2) hold then by Lemma 2.13, applying appropriate operators \( z_{in} \) repeatedly to the vector \( \xi_{\mu} \) we can obtain the highest vector \( \xi \) with a nonzero coefficient. This gives \( \xi_{\mu} \neq 0 \).
Thus, we have proved that the vectors $\xi_A$ defined in (2.11) form a basis of the representation $L(\lambda)$. The orthogonality of the basis vectors (Proposition 2.4) is implied by the fact that the operators $pE_{ni}$ and $pE_{in}$ are adjoint to each other with respect to the restriction of the inner product $\langle \cdot, \cdot \rangle$ to the subspace $L(\lambda)^+$. Therefore, for the adjoint operator to $z_{ni}$ we have

$$z_{ni}^* = z_{in} \frac{(h_i - h_{i+1} - 1) \cdots (h_i - h_{n-1} - 1)}{(h_i - h_1) \cdots (h_i - h_{i-1})}$$

and Proposition 2.4 follows from Lemma 2.13 by induction.

Now we outline a derivation of formulas (2.5)–(2.7). First, since $E_{nn}z_{ni} = z_{ni}(E_{nn} + 1)$ for any $i$, we have

$$E_{nn}\xi_\mu = \left( \sum_{i=1}^n \lambda_i - \sum_{i=1}^{n-1} \mu_i \right) \xi_\mu$$

which implies (2.5). To prove (2.6) is suffices to calculate $E_{n-1,n}\xi_{\mu\nu}$, where

$$\xi_{\mu\nu} = z_{n-1,1}^{\mu_1-v_1} \cdots z_{n-1,n-2}^{\mu_n-v_n} \xi_\mu$$

and the $v_i$ satisfy the betweenness conditions

$$\mu_i - v_i \in \mathbb{Z}_+ \quad \text{and} \quad v_i - \mu_{i+1} \in \mathbb{Z}_+ \quad \text{for} \quad i = 1, \ldots, n - 2.$$

Since $E_{n-1,n}$ commutes with the $z_{n-1,i}$,

$$E_{n-1,n}\xi_{\mu\nu} = z_{n-1,1}^{\mu_1-v_1} \cdots z_{n-1,n-2}^{\mu_n-v_n} E_{n-1,n}\xi_\mu.$$

The following lemma is implied by the explicit formulas for the lowering and raising operators (2.9) and (2.10).

**LEMMA 2.14.** We have the relation in $U'(\mathfrak{gl}_n)$ modulo the ideal $J'$,

$$E_{n-1,n} = \sum_{i=1}^{n-1} z_{n-1,i} z_{in} \frac{1}{(h_i - h_1) \cdots \wedge_i \cdots (h_i - h_{n-1})},$$

where $z_{n-1,n-1} = 1$.

By Lemmas 2.13 and 2.14,

$$E_{n-1,n}\xi_{\mu\nu} = - \sum_{i=1}^{n-1} \frac{(m_i - l_1) \cdots (m_i - l_n)}{(m_i - m_1) \cdots \wedge_i \cdots (m_i - m_{n-1})} \xi_{\mu + \delta_i, \nu}$$

(2.29)
which proves (2.6). To prove (2.7) we use Proposition 2.4. Relation (2.29) implies that

\[ E_{n,n-1} \xi_{\mu \nu} = \sum_{i=1}^{n-1} c_i(\mu, \nu) \xi_{\mu \delta_i \nu} \]

for some coefficients \( c_i(\mu, \nu) \). Apply the operator \( z_{j,n-1} \) to both sides of this relation. Since \( z_{j,n-1} \) commutes with \( E_{n,n-1} \) we obtain from Lemma 2.13 a recurrence relation for the \( c_i(\mu, \nu) \): if \( \mu_j - \nu_j \geq 1 \) then

\[ c_i(\mu, \nu + \delta_j) = c_i(\mu, \nu) \frac{m_i - \gamma_j - 1}{m_i - \gamma_j} \]

where \( \gamma_j = \nu_j - j + 1 \). The proof is completed by induction. The initial values of \( c_i(\mu, \nu) \) are found by applying the relation

\[ E_{n,n-1} z_{n-1,i} = z_{n,i} \frac{1}{h_i - h_{n-1}} + z_{n-1,i} E_{n,n-1} \frac{h_i - h_{n-1} - 1}{h_i - h_{n-1}} \]

to the vector \( \xi_{\mu} \) and taking into account that \( E_{n,n-1} = z_{n,n-1} \). Performing the calculation we get

\[ E_{n,n-1} \xi_{\mu \nu} = \sum_{i=1}^{n-1} \frac{(m_i - \gamma_1) \cdots (m_i - \gamma_{n-2})}{(m_i - m_1) \cdots \wedge_i \cdots (m_i - m_{n-1})} \xi_{\mu \delta_i \nu} \]

thus proving (2.7).

2.4. Characteristic identities

Denote by \( L \) the vector representation of \( \mathfrak{gl}_n \) and consider its contragredient \( L^* \). Note that \( L^* \) is isomorphic to \( L(0, \ldots, 0, -1) \). Let \( \{ e_1, \ldots, e_n \} \) denote the basis of \( L^* \) dual to the canonical basis \( \{ e_1, \ldots, e_n \} \) of \( L \). Introduce the \( n \times n \)-matrix \( E \) whose \( ij \)-th entry is the generator \( E_{ij} \). We shall interpret \( E \) as the element

\[ E = \sum_{i,j=1}^{n} e_{ij} \otimes E_{ij} \in \text{End} L^* \otimes \mathfrak{u}(\mathfrak{gl}_n), \]

where the \( e_{ij} \) are the standard matrix units acting on \( L^* \) by \( e_{ij} e_k = \delta_{jk} e_i \). The basis element \( E_{ij} \) of \( \mathfrak{gl}_n \) acts on \( L^* \) as \(-e_{ji}\) and hence \( E \) may also be thought of as the image of the element

\[ e = -\sum_{i,j=1}^{n} E_{ji} \otimes E_{ij} \in U(\mathfrak{gl}_n) \otimes U(\mathfrak{gl}_n). \]
On the other hand, using the standard coproduct $\Delta$ on $U(\mathfrak{g}l_n)$ defined by
\[ \Delta(E_{ij}) = E_{ij} \otimes 1 + 1 \otimes E_{ij}, \]
we can write $e$ in the form
\[ e = \frac{1}{2} \left( z \otimes 1 + 1 \otimes z - \Delta(z) \right), \] (2.30)
where $z$ is the second-order Casimir element
\[ z = \sum_{i,j=1}^{n} E_{ij}E_{ji} \in U(\mathfrak{g}l_n). \]

We have the tensor product decomposition
\[ L^* \otimes L(\lambda) \simeq L(\lambda - \delta_1) \oplus \cdots \oplus L(\lambda - \delta_n), \] (2.31)
where $L(\lambda - \delta_i)$ is considered to be zero if $\lambda_i = \lambda_{i+1}$. On the level of characters this is a particular case of the Pieri rule for the expansion of the product of a Schur polynomial by an elementary symmetric polynomial; see, e.g., Macdonald, [84, Chapter 1]. The Casimir element $z$ acts as a scalar operator in any highest weight representation $L(\lambda)$. The corresponding eigenvalue is given by
\[ z|_{L(\lambda)} = \sum_{i=1}^{n} \lambda_i(\lambda_i + n - 2i + 1). \]

Regarding now $E$ as an operator on $L^* \otimes L(\lambda)$ and using (2.30) we derive that the restriction of $E$ to the summand $L(\lambda - \delta_r)$ in (2.31) is the scalar operator with the eigenvalue $\lambda_r + n - r$ which we shall denote by $\alpha_r$. This implies the characteristic identity for the matrix $E$,
\[ \prod_{r=1}^{n} (E - \alpha_r) = 0, \] (2.32)
as an operator in $L^* \otimes L(\lambda)$. Moreover, the projection $P[r]$ of $L^* \otimes L(\lambda)$ to the summand $L(\lambda - \delta_r)$ can be written explicitly as
\[ P[r] = \frac{(E - \alpha_1) \cdots \wedge_r \cdots (E - \alpha_n)}{(\alpha_r - \alpha_1) \cdots \wedge_r \cdots (\alpha_r - \alpha_n)}, \]
with $\wedge_r$ indicating that the $r$-th factor is omitted. Together with (2.32) this yields the spectral decomposition of $E$,
\[ E = \sum_{r=1}^{n} \alpha_r P[r]. \] (2.33)
Consider the orthonormal Gelfand–Tsetlin bases \( \{ \zeta_A \} \) of \( L(\lambda) \) and \( \{ \zeta_{A(r)} \} \) of \( L(\lambda - \delta_r) \) for \( r = 1, \ldots, n \); see (2.8). Regarding the matrix element \( P[r]_{ij} \) as an operator in \( L(\lambda) \) we obtain

\[
\langle \zeta_A', P[r]_{ij} \zeta_A \rangle = \langle \epsilon_i \otimes \zeta_A', P[r](\epsilon_j \otimes \zeta_A) \rangle,
\]

where we have extended the inner products on \( L^* \) and \( L(\lambda) \) to \( L^* \otimes L(\lambda) \) by setting

\[
\langle \eta \otimes \zeta, \eta' \otimes \zeta' \rangle = \langle \eta, \eta' \rangle \langle \zeta, \zeta' \rangle
\]

with \( \eta, \eta' \in L^* \) and \( \zeta, \zeta' \in L(\lambda) \). Furthermore, using the expansions

\[
\epsilon_j \otimes \zeta_A = \sum_{s=1}^{n} \sum_{A(s)} \langle \epsilon_j \otimes \zeta_A, \zeta_{A(s)} \rangle \zeta_{A(s)},
\]

brings (2.34) to the form

\[
\sum_{A(r)} \langle \epsilon_i \otimes \zeta_{A'}, \zeta_{A(r)} \rangle \langle \epsilon_j \otimes \zeta_A, \zeta_{A(r)} \rangle,
\]

where we have used the fact that \( P[r] \) is the identity map on \( L(\lambda - \delta_r) \), and zero on \( L(\lambda - \delta_s) \) with \( s \neq r \). The numbers \( \langle \epsilon_i \otimes \zeta_{A'}, \zeta_{A(r)} \rangle \) are the Wigner coefficients (a particular case of the Clebsch–Gordan coefficients). They can be used to express the matrix elements of the generators \( E_{ij} \) in the Gelfand–Tsetlin basis as follows. Using the spectral decomposition (2.33) we get

\[
E_{ij} = \sum_{r=1}^{n} \alpha_r P[r]_{ij}.
\]

Therefore, we derive the following result from (2.34).

**THEOREM 2.15.** We have

\[
\langle \zeta_{A'}, E_{ij} \zeta_A \rangle = \sum_{r=1}^{n} \alpha_r \sum_{A(r)} \langle \epsilon_i \otimes \zeta_{A'}, \zeta_{A(r)} \rangle \langle \epsilon_j \otimes \zeta_A, \zeta_{A(r)} \rangle.
\]

Employing the characteristic identities for both the Lie algebras \( \mathfrak{gl}_{n+1} \) and \( \mathfrak{gl}_n \) it is possible to determine the values of the Wigner coefficients and thus to get an independent derivation of the formulas of Theorem 2.3. In fact, explicit formulas for the matrix elements of \( E_{ij} \) with \( |i - j| > 1 \) can also be given; see Gould, [48], for details.

The approach based upon the characteristic identities also leads to an alternative presentation of the lowering and raising operators. Taking \( \zeta_A \) to be the highest vector \( \xi \) in (2.34)
we conclude that $P[r]_{ij}\xi = 0$ for $j > r$. Consider now $\mathfrak{gl}_n$ as a subalgebra of $\mathfrak{gl}_{n+1}$. Suppose that $\xi$ is a highest vector of weight $\lambda$ in a representation $L(\lambda')$ of $\mathfrak{gl}_{n+1}$. The previous observation implies that the vector

$$\sum_{i=r}^{n} E_{n+1,i} P[r]_{ir} \xi$$

is again a $\mathfrak{gl}_n$-highest vector of weight $\lambda - \delta_r$.

**Proposition 2.16.** We have the following identity of operators on the space $L(\lambda')$:

$$pE_{n+1,r} = \sum_{i=r}^{n} E_{n+1,i} P[r]_{ir},$$

where $p$ is the extremal projector for $\mathfrak{gl}_n$.

**Outline of the proof.** Since both sides represent lowering operators they must be proportional. It is therefore sufficient to apply both sides to a vector $\xi \in L(\lambda')$ and compare the coefficients at $E_{n+1,r} \xi$. For the calculation we use the explicit formula (2.16) for $pE_{n+1,r}$ and the relation

$$P[r]_{ir} \xi = \prod_{s=r+1}^{n} \frac{h_r - h_s - 1}{h_r - h_s} \xi$$

which can be derived from the characteristic identities.

An analogous argument leads to a similar formula for the raising operators. Here one starts with the dual characteristic identity

$$\prod_{r=1}^{n} (\bar{E} - \bar{\alpha}_r) = 0,$$

where the $ij$-th matrix element of $\bar{E}$ is $-E_{ij}$, $\bar{\alpha}_r = -\lambda_r + r$ and the powers of $\bar{E}$ are defined recursively by

$$(\bar{E}^p)_{ij} = \sum_{k=1}^{n} (\bar{E}^{p-1})_{kj} \bar{E}_{ik}.$$ 

For any $r = 1, \ldots, n$ the dual projection operator is given by

$$\bar{P}[r] = \frac{(\bar{E} - \bar{\alpha}_1) \cdots \wedge_r \cdots (\bar{E} - \bar{\alpha}_n)}{(\bar{\alpha}_r - \bar{\alpha}_1) \cdots \wedge_r \cdots (\bar{\alpha}_r - \bar{\alpha}_n)}.$$
PROPOSITION 2.17. We have the following identity of operators on the space $L(\lambda')^+_\lambda$:

$$pE_{r,n+1} = \sum_{i=1}^r E_{i,n+1} \bar{P}[r]_{ri}.$$  

2.5. Quantum minors

For a complex parameter $u$ introduce the $n \times n$-matrix $E(u) = u1 + E$. Given sequences $a_1, \ldots, a_s$ and $b_1, \ldots, b_s$ of elements of $\{1, \ldots, n\}$ the corresponding quantum minor of the matrix $E(u)$ is defined by the following equivalent formulas:

$$E(u)_{a_1 \cdots a_s}^{b_1 \cdots b_s} = \sum_{\sigma \in S_s} \text{sgn} \sigma \cdot E(u)^{a_{\sigma(1)}b_1} \cdots E(u-s+1)^{a_{\sigma(s)}b_s},$$

(2.35)

$$= \sum_{\sigma \in S_s} \text{sgn} \sigma \cdot E(u-s+1)^{a_{\sigma(1)}b_1} \cdots E(u)^{a_{\sigma(s)}b_s}.$$  

(2.36)

This is a polynomial in $u$ with coefficients in $U(gl_n)$. It is skew symmetric under permutations of the indices $a_i$, or $b_i$.

For any index $1 \leq i < n$ introduce the polynomials

$$\tau_{ni}(u) = E(u)^{i+1 \cdots n}_{1 \cdots n-1} \quad \text{and} \quad \tau_{in}(u) = (-1)^{i-1} E(u)^{1 \cdots i}_{1 \cdots i-1 \cdots n}.$$  

For instance,

$$\tau_{13}(u) = E_{13}, \quad \tau_{23}(u) = -E_{23}(u + E_{11}) + E_{21}E_{13},$$

$$\tau_{32}(u) = E_{32}, \quad \tau_{31}(u) = E_{21}E_{32} - E_{31}(u + E_{22} - 1).$$

PROPOSITION 2.18. If $\eta \in L(\lambda)^+_{\mu}$ then

$$\tau_{ni}(-\mu_i)\eta \in L(\lambda)^+_{\mu - \delta_i} \quad \text{and} \quad \tau_{in}(-\mu_i + i - 1)\eta \in L(\lambda)^+_{\mu + \delta_i}.$$  

OUTLINE OF THE PROOF. The proof is based upon the following relations

$$[E_{ij}, E(u)_{b_1 \cdots b_j}^{a_1 \cdots a_s}] = \sum_{r=1}^s (\delta_{j a_r} E(u)_{b_1 \cdots b_r}^{a_1 \cdots a_s} - \delta_{i b_r} E(u)_{b_1 \cdots j \cdots b_s}^{a_1 \cdots a_s}),$$  

(2.37)

where $i$ and $j$ on the right-hand side are in the $r$-th slot.  

The relations (2.37) imply an important property of the quantum minors: for any indices $i$, $j$ we have

$$[E_{a_ib_j}, E(u)_{b_1 \cdots b_i}^{a_1 \cdots a_s}] = 0.$$
In particular, this implies the centrality of the Capelli determinant $C(u) = E(u)^{1\ldots n}$; see (2.21).

The lowering and raising operators of Proposition 2.18 can be shown to essentially coincide with those defined in Section 2.1. To write down the formulas we shall need to evaluate the variable $u$ in $U(h)$. To make this operation well-defined we use the agreement used in the evaluation of the Capelli determinant. See just below (2.25).

**Proposition 2.19.** We have the following identities for any $i = 1, \ldots, n - 1$

$$\tau_{ni}(-h_i - i + 1) = z_{ni} \quad \text{and} \quad \tau_{in}(-h_i) = z_{in}. \quad (2.38)$$

Using this interpretation of the lowering operators one can express the Gelfand–Tsetlin basis vector (2.11) in terms of the quantum minors $\tau_{ki}(u)$. The action of certain other quantum minors on these vectors can be explicitly found. This will provide one more independent proof of Theorem 2.3. For $m \geq 1$ introduce the polynomials $A_m(u)$, $B_m(u)$ and $C_m(u)$ by

$$A_m(u) = E(u)^{1\ldots m}, \quad B_m(u) = E(u)^{1\ldots m-1,m+1},$$

$$C_m(u) = E(u)^{1\ldots m-1,m+1}.$$

We use the notation $l_{mi} = \lambda_{mi} - i + 1$ and $l_i = \lambda_i - i + 1$.

**Theorem 2.20.** Let $\xi_A$ be the Gelfand–Tsetlin basis of $L(\lambda)$. Then

$$A_m(u)\xi_A = (u + l_m)\cdots (u + l_{mm})\xi_A, \quad (2.39)$$

$$B_m(-l_{mj})\xi_A = -\prod_{i=1}^{m+1}(l_{m+1,i} - l_{mj})\xi_{A+\delta_{mj}} \quad \text{for } j = 1, \ldots, m,$$

$$C_m(-l_{mj})\xi_A = \prod_{i=1}^{m-1}(l_{m-1,i} - l_{mj})\xi_{A-\delta_{mj}} \quad \text{for } j = 1, \ldots, m, \quad (2.40)$$

where $A \pm \delta_{mj}$ is obtained from $A$ by replacing the entry $\lambda_{mj}$ with $\lambda_{mj} \pm 1$.

Applying the Lagrange interpolation formula we can find the action of $B_m(u)$ and $C_m(u)$ for any $u$. Note that these polynomials have degree $m - 1$ with leading coefficients $E_{m,m+1}$ and $E_{m+1,m}$, respectively. Theorem 2.3 is therefore an immediate corollary of Theorem 2.20.

Formula (2.40) prompts a quite different construction of the basis vectors of $L(\lambda)$ which uses the polynomials $C_m(u)$ instead of the traditional lowering operators $z_{ni}$. Indeed, for a particular value of $u$, $C_m(u)$ takes a basis vector into another one, up to a factor. Given a pattern $\Lambda$ associated with $\lambda$, define the vector $\kappa_A \in L(\lambda)$ by

$$\kappa_A = \prod_{k=1,\ldots,n-1} (C_{n-1}(-l_{n-1,k} - 1) \cdots C_{n-1}(-l_k + 1))C_{n-1}(-l_k)$$
$C_{n-2}(-l_{n-2,k}-1) \cdots C_{n-2}(-l_k+1)C_{n-2}(-l_k)$
$\times \cdots \times C_k(-l_{k-1}-1) \cdots C_k(-l_k+1)C_k(-l_k) \xi.$

**THEOREM 2.21.** The vectors $\kappa_\Lambda$ with $\Lambda$ running over all patterns associated with $\lambda$ form a basis of $L(\lambda)$ and one has $\kappa_\Lambda = N_\Lambda \xi_\Lambda$ for a nonzero constant $N_\Lambda$.

The value of the constant $N_\Lambda$ can be found from (2.40). Using the relations between the elements $A_m(u)$, $B_m(u)$ and $C_m(u)$ one can derive Theorem 2.20 from Theorem 2.21 with the use of Proposition 2.22 below; see Nazarov and Tarasov, [109], for details.

Observe that $A_m(u)$ is the Capelli determinant (2.21) for the Lie algebra $gl_m$. Therefore, its coefficients $a_{mi}$ defined by

$$A_m(u) = u^m + a_{m1}u^{m-1} + \cdots + a_{mm}$$

are generators of the center of the enveloping algebra $U(gl_m)$. All together the elements $a_{mi}$ with $1 \leq i \leq m \leq n$ generate a commutative subalgebra $A_n$ of $U(gl_n)$ which is called the Gelfand–Tsetlin subalgebra. By (2.39), the basis vectors $\xi_\Lambda$ are simultaneous eigenvectors for the elements of the subalgebra $A_n$. Introduce the corresponding eigenvalues of the generators $a_{mi}$ by

$$a_{mi} \xi_\Lambda = \alpha_{mi}(\Lambda) \xi_\Lambda. \quad (2.41)$$

Thus, $\alpha_{mi}(\Lambda)$ is the $i$-th elementary symmetric polynomial in $l_{m1}, \ldots, l_{mm}$.

**PROPOSITION 2.22.** For any pattern $\Lambda$ associated with $\lambda$, the one-dimensional subspace of $L(\lambda)$ spanned by the basis vector $\xi_\Lambda$ is uniquely determined by the set of eigenvalues $\{\alpha_{mi}(\Lambda)\}$.

**Bibliographical notes**

The explicit formulas for the lowering and raising operators (2.9) and (2.10) first appeared in Nagel and Moshinsky, [106]; see also Hou Pei-yu, [59], and Zhelobenko, [168]. The derivation of the Gelfand–Tsetlin formulas outlined in Section 2.1 follows Zhelobenko, [168], and Asherova, Smirnov and Tolstoy, [2]. The extremal projectors were originally introduced by Asherova, Smirnov and Tolstoy, [1] (see also [3]). In a subsequent paper [2] the projectors were used to construct the lowering operators and derive the relations between them. A systematic study of the extremal projectors and the corresponding Mickelsson algebras was undertaken by Zhelobenko: a detailed exposition is given in his paper [173] and book [174]. The application to the Gelfand–Tsetlin formulas is contained in his paper [171]. Section 2.2 is a brief outline of the general results which are used in the basis constructions.

The first proof of Theorem 2.11 was given by van den Hombergh, [158], as an answer to the question posed by Mickelsson, [92]. A derivation of the relations in the Mickelsson–Zhelobenko algebra $Z(gl_n, gl_m)$ with the use of the Capelli-type determinants is contained
in the author’s paper [101]. A proof of the formulas (2.23) and (2.24) is also given there. The results of Section 2.4 are due to Gould, [46–49]. The characteristic identity (2.32) was proved by Green, [54]. The significance of the Wigner coefficients in mathematical physics is discussed in the book by Biedenharn and Louck, [8]. The definition (2.35) of the quantum minors is inspired by the theory of “quantum” algebras called the Yangians; see [103, 104] for a review of the theory. The polynomials $A_m(u)$, $B_m(u)$ and $C_m(u)$ are essentially the images of the Drinfeld generators of the Yangian $Y(n)$ under the evaluation homomorphism to the universal enveloping algebra $U(gl_n)$. The quantum minor presentation of the lowering operators (2.38) is due to the author, [96]; see also [101]. The construction of the Gelfand–Tsetlin basis vectors $\kappa_A$ with the use of the Drinfeld generators (Theorem 2.21) was devised by Nazarov and Tarasov, [109].

Analogs of the extremal projector were given by Tolstoy, [150–154], for a wide class of Lie (super)algebras and their quantized enveloping algebras. The corresponding super and quantum versions of the Mickelsson–Zhelobenko algebras are studied in [152–154]. An alternative “tensor formula” for the extremal projector was provided by Tolstoy and Draayer, [155]. The techniques of extremal projectors were applied by Khosrhshin and Tolstoy, [67], for calculation of the universal $R$-matrices for quantized enveloping algebras. A basis of Gelfand–Tsetlin type for representations of the exceptional Lie algebra $G_2$ was constructed by Sviridov, Smirnov and Tolstoy, [144,145].

Bases of Gelfand–Tsetlin type have been constructed for representations of various types of algebras. For the quantized enveloping algebra $U_q(gl_n)$ such bases were constructed by Jimbo, [65], Ueno, Takebayashi and Shibukawa, [157], Nazarov and Tarasov, [109], Tolstoy, [153]. The results of [109] include $q$-analogs of Theorems 2.20 and 2.21, while [153] contains matrix element formulas for the generators corresponding to arbitrary roots. Gould and Biedenharn, [52], developed pattern calculus for representations of the quantum group $U_q(u(n))$. Polynomial realizations of the Gelfand–Tsetlin basis for representations of $U_q(sl_3)$ were given by Dobrev and Truini, [20,21], and Dobrev, Mitov and Truini, [22].

Gelfand–Tsetlin bases for ‘generic’ representations of the Yangian $Y(n)$ were constructed in [96]. Theorem 2.20 was proved there in the more general context of representations of the Yangian of level $p$ for $gl_n$, which was previously introduced by Cherednik, [17]. In particular, the enveloping algebra $U(gl_n)$ coincides with the Yangian of level 1. A more general class of the tame Yangian modules was introduced and explicitly constructed by Nazarov and Tarasov, [110], via the trapezium or skew analogs of the Gelfand–Tsetlin patterns. Their approach was motivated by the so-called centralizer construction devised by Olshanski, [114,116,117], and also employed by Cherednik, [16,17]. Basis vectors in the tame Yangian modules are characterized in a way similar to Proposition 2.22. The skew Yangian modules were also studied in [101] with the use of the quantum Sylvester theorem and the Mickelsson algebras.

The center of $U(gl_n)$ possesses several natural families of generators and so does the Gelfand–Tsetlin subalgebra $A_n$. The corresponding eigenvalues in $L(\lambda)$ are known explicitly; see, e.g., [103] for a review. An alternative description of $A_n$ was given by Gelfand, Krob, Lascoux, Leclerc, Retakh and Thibon, [40, Section 7.3], as an application of their theory of noncommutative symmetric functions and quasi-determinants.

The combinatorics of the skew Gelfand–Tsetlin patterns is employed by Berenstein and Zelevinsky, [7], to obtain multiplicity formulas for the skew representations of $gl_n$. Ap-
plications to continuous piecewise linear actions of the symmetric group were found by Kirillov and Berenstein, [72].

The explicit realization of irreducible finite-dimensional representations of \( gl_n \) via the Gelfand–Tsetlin bases has important applications in the representation theory of quantum affine algebras and Yangians. In particular, Theorem 2.20 and its Yangian analog, [96], are crucial in the proof of the irreducibility criterion of the tensor products of the Yangian evaluation modules (a generalization to \( gl_n \) of Theorem 3.8 below); see [102].

Analogs of the Gelfand–Tsetlin bases for representations of some Lie superalgebras and their quantum analogs were given by Ottoson, [119, 120], Palev, [122–125], Palev, Stoilova and van der Jeugt, [131], Palev and Tolstoy, [132], Tolstoy, Istomina and Smirnov, [156]. Highest weight irreducible representations for the Lie (super)algebras of infinite matrices and their quantum analogs were constructed by Palev, [126, 127], and Palev and Stoilova, [128–130], via bases of Gelfand–Tsetlin-type.

The explicit formulas of Theorem 2.3 make it possible to define a class of infinite-dimensional representations of \( gl_n \) by altering the inequalities (2.4). Families of such representations were introduced by Gelfand and Graev, [39]. However, as was later observed by Lemire and Patera, [80], some necessary conditions were missing in [39], so that only a part of those families actually provides representations. A more general theory of the so-called Gelfand–Tsetlin modules is developed by Drozd, Futorny and Ovsienko, [31–34], Ovsienko, [121], and Mazorchuk, [87, 88]. The starting point of the theory is to axiomatize the property of the basis vectors (2.41) and to consider the module generated by an eigenvector for the Gelfand–Tsetlin subalgebra with a given arbitrary set of eigenvalues \( \{\alpha_m\} \).

Some \( q \)-analogs of such modules were constructed by Mazorchuk and Turowska, [90].

The formulas of Theorem 2.3 were applied by Olshanski, [113, 115], to study unitary representations of the pseudo-unitary groups \( U(p,q) \). In particular, he classified all irreducible unitarizable highest weight representations of the Lie algebra \( u(p,q) \), [113]. This work was extended by the author to a family of the Enright–Varadarajan modules over \( u(p,q) \), [95]. Analogs of the Gelfand–Tsetlin bases for the unitary highest weight modules were constructed in [94].

Applications of the Gelfand–Tsetlin bases in mathematical physics are reviewed in the books by Barut and Racznka, [5], and Biedenharn and Louck, [8].

3. Weight bases for representations of \( \mathfrak{o}_N \) and \( \mathfrak{sp}_{2n} \)

Let \( g_n \) denote the rank \( n \) simple complex Lie algebra of type \( B, C, \) or \( D \). That is,

\[
g_n = \mathfrak{o}_{2n+1}, \; \mathfrak{sp}_{2n}, \; \text{or} \; \mathfrak{o}_{2n},
\]

respectively. Let \( V(\lambda) \) denote the finite-dimensional irreducible representation of \( g_n \) with the highest weight \( \lambda \). The restriction of \( V(\lambda) \) to the subalgebra \( g_{n-1} \) is not multiplicity-free in general. This means that if \( V'(\mu) \) is the finite-dimensional irreducible representation of \( g_{n-1} \) with the highest weight \( \mu \), then the space

\[
\text{Hom}_{g_{n-1}}(V'(\mu), V(\lambda))
\]

(3.2)
need not be one-dimensional. In order to construct a basis of \( V(\lambda) \) associated with the chain of subalgebras

\[
\mathfrak{g}_1 \subset \mathfrak{g}_2 \subset \cdots \subset \mathfrak{g}_n
\]

we need to construct a basis of the space (3.2) which is isomorphic to the subspace \( V(\lambda)_\mu^+ \) of \( \mathfrak{g}_{n-1} \)-highest vectors of weight \( \mu \) in \( V(\lambda) \). The subspace \( V(\lambda)_\mu^+ \) possesses a natural structure of a representation of the centralizer \( C_n = U(\mathfrak{g}_n)_{\mathfrak{g}_n-1} \) in the universal enveloping algebra \( U(\mathfrak{g}_n) \). It was shown by Olshanski, [118], that there exist natural homomorphisms

\[
C_1 \leftarrow C_2 \leftarrow \cdots \leftarrow C_n \leftarrow C_{n+1} \leftarrow \cdots.
\]

The projective limit of this chain turns out to be an extension of the twisted Yangian \( Y^+(2) \) or \( Y^-(2) \), in the orthogonal and symplectic case, respectively; see [118,104] and [105] for the definition and properties of the twisted Yangians. In particular, there is an algebra homomorphism \( Y^\pm(2) \to C_n \) which allows one to equip the space \( V(\lambda)_\mu^+ \) with a \( Y^\pm(2) \)-module structure. By the results of [97], the representation \( V(\lambda)_\mu^+ \) can be extended to a larger algebra, the Yangian \( Y(2) \). This is a key fact which allows us to construct a natural basis in each space \( V(\lambda)_\mu^+ \). In the \( C \) and \( D \) cases the \( Y(2) \)-module \( V(\lambda)_\mu^+ \) is irreducible while in the \( B \) case it is a direct sum of two irreducible submodules. This does not lead, however, to major differences in the constructions, and the final formulas are similar in all the three cases.

The calculations of the matrix elements of the generators of \( \mathfrak{g}_n \) are based on the relationship between the twisted Yangian \( Y^\pm(2) \) and the Mickelsson–Zhelobenko algebra \( Z(\mathfrak{g}_n, \mathfrak{g}_{n-1}) \); see Section 2.2. We construct an algebra homomorphism \( Y^\pm(2) \to Z(\mathfrak{g}_n, \mathfrak{g}_{n-1}) \) which allows us to express the generators of the twisted Yangian, as operators in \( V(\lambda)_\mu^+ \), in terms of the lowering and raising operators.

### 3.1. Raising and lowering operators

Whenever possible we consider the three cases (3.1) simultaneously, unless otherwise stated. The rows and columns of \( 2n \times 2n \)-matrices will be enumerated by the indices \(-n, \ldots, -1, 1, \ldots, n\), while the rows and columns of \( (2n + 1) \times (2n + 1) \)-matrices will be enumerated by the indices \(-n, \ldots, -1, 0, 1, \ldots, n\). Accordingly, the index 0 will usually be skipped in the former case. For \(-n \leq i, j \leq n\) set

\[
F_{ij} = E_{ij} - \theta_{ij} E_{-j, -i},
\]

where the \( E_{ij} \) are the standard matrix units, and

\[
\theta_{ij} = \begin{cases} 
1 & \text{in the orthogonal case,} \\
\text{sgn} i \cdot \text{sgn} j & \text{in the symplectic case.}
\end{cases}
\]
The matrices $F_{ij}$ span the Lie algebra $\mathfrak{g}_n$. The subalgebra $\mathfrak{g}_{n-1}$ is spanned by the elements (3.3) with the indices $i, j$ running over the set $\{-n + 1, \ldots, n - 1\}$. Denote by $\mathfrak{h} = \mathfrak{h}_n$ the diagonal Cartan subalgebra in $\mathfrak{g}_n$. The elements $F_{11}, \ldots, F_{nn}$ form a basis of $\mathfrak{h}$.

The finite-dimensional irreducible representations of $\mathfrak{g}_n$ are in a one-to-one correspondence with $n$-tuples $\lambda = (\lambda_1, \ldots, \lambda_n)$ where the numbers $\lambda_i$ satisfy the conditions

$$\lambda_i - \lambda_{i+1} \in \mathbb{Z}_+ \quad \text{for } i = 1, \ldots, n - 1,$$

and

$$-2\lambda_1 \in \mathbb{Z}_+ \quad \text{for } \mathfrak{g}_n = \mathfrak{o}_{2n+1},$$
$$-\lambda_1 \in \mathbb{Z}_+ \quad \text{for } \mathfrak{g}_n = \mathfrak{sp}_{2n},$$
$$-\lambda_1 - \lambda_2 \in \mathbb{Z}_+ \quad \text{for } \mathfrak{g}_n = \mathfrak{o}_{2n}. $$

Such an $n$-tuple $\lambda$ is called the highest weight of the corresponding representation which we shall denote by $V(\lambda)$. It contains a unique, up to a constant factor, nonzero vector $\xi$ (the highest vector) such that

$$F_{ii} \xi = \lambda_i \xi \quad \text{for } i = 1, \ldots, n,$$
$$F_{ij} \xi = 0 \quad \text{for } -n \leqslant i < j \leqslant n.$$

Denote by $V(\lambda)^+$ the subspace of $\mathfrak{g}_{n-1}$-highest vectors in $V(\lambda)$:

$$V(\lambda)^+ = \{ \eta \in V(\lambda) \mid F_{ij} \eta = 0, \ -n < i < j < n \}. $$

Given a $\mathfrak{g}_{n-1}$-weight $\mu = (\mu_1, \ldots, \mu_{n-1})$ we denote by $V(\lambda)^+_\mu$ the corresponding weight subspace in $V(\lambda)^+$:

$$V(\lambda)^+_\mu = \{ \eta \in V(\lambda)^+ \mid F_{ii} \eta = \mu_i \eta, \ i = 1, \ldots, n - 1 \}. $$

Consider the Mickelsson–Zhelobenko algebra $Z(\mathfrak{g}_n, \mathfrak{g}_{n-1})$ introduced in Section 2.2. Let $p = p_{n-1}$ be the extremal projector for the Lie algebra $\mathfrak{g}_{n-1}$. It satisfies the conditions

$$F_{ij} p = p F_{ji} = 0 \quad \text{for } -n < i < j < n.$$

By Theorem 2.10, the elements

$$F_{nn}, \quad p F_{ia}, \quad a = -n, n, \ i = -n + 1, \ldots, n - 1, $$

are generators of $Z(\mathfrak{g}_n, \mathfrak{g}_{n-1})$ in the orthogonal case. In the symplectic case, the algebra $Z(\mathfrak{g}_n, \mathfrak{g}_{n-1})$ is generated by the elements (3.7) together with $F_{n,-n}$ and $F_{-n,n}$. To write down explicit formulas for the generators, introduce numbers $\rho_i$, where $i = 1, \ldots, n$, by

$$\rho_i = \begin{cases} 
-i + 1/2 & \text{for } \mathfrak{g}_n = \mathfrak{o}_{2n+1}, \\
-i & \text{for } \mathfrak{g}_n = \mathfrak{sp}_{2n}, \\
-i + 1 & \text{for } \mathfrak{g}_n = \mathfrak{o}_{2n}. 
\end{cases}$$
The numbers $-\rho_i$ are coordinates of the half-sum of positive roots with respect to the upper triangular Borel subalgebra. Now set

$$f_i = F_{ii} + \rho_i, \quad f_{-i} = -f_i$$

for $i = 1, \ldots, n$. Moreover, in the case of $\sigma_{2n+1}$ also set $f_0 = -1/2$. The generators $pF_{ia}$ can be given by a uniform expression in all the three cases. Let $a \in \{-n, n\}$ and $i \in \{-n + 1, \ldots, n - 1\}$. Then we have the modulo the ideal $J'$

$$pF_{ia} = F_{ia} + \sum_{i > i_1 > \cdots > i_s > -n} F_{ii_1} F_{ii_2} \cdots F_{i_{s-1}i_s} F_{is} \frac{1}{(f_i - f_{i_1}) \cdots (f_i - f_{i_s})},$$

(3.8)

summed over $s \geq 1$. It will be convenient to work with normalized generators of $Z(g_n, g_{n-1})$. Set

$$z_{ia} = pF_{ia}(f_i - f_{i-1}) \cdots (f_i - f_{-n+1})$$

in the $B, C$ cases, and

$$z_{ia} = pF_{ia}(f_i - f_{i-1}) \cdots (\hat{f}_{i-j}) \cdots (f_i - f_{-n+1})$$

in the $D$ case, where the hat indicates the factor to be omitted if it occurs. We shall also use the elements $z_{ai}$ defined by

$$z_{ai} = (-1)^{n-i} z_{i,-a} \quad \text{and} \quad z_{ai} = (-1)^{n-i} \text{sgn } a \cdot z_{-i,-a},$$

in the orthogonal and symplectic case, respectively. The elements $z_{ia}$ satisfy some quadratic relations which can be shown to be the defining relations of the algebra $Z(g_n, g_{n-1})$. In particular, we have for all $a, b \in \{-n, n\}$ and $i + j \neq 0$,

$$z_{ia} z_{ib} + z_{ja} z_{ib}(f_i - f_j - 1) = z_{ib} z_{ja}(f_i - f_j).$$

(3.9)

Thus, $z_{ia}$ and $z_{ja}$ commute for $i + j \neq 0$. Also, $z_{ia}$ and $z_{ib}$ commute for $i \neq 0$ and all values of $a$ and $b$. Analogs of the relation (2.20) in the algebra $Z(g_n, g_{n-1})$ can be explicitly written down as well. However, we shall avoid using them in a way similar to the proof of Lemma 2.13.

The elements $z_{ia}$ naturally act in the space $V(\lambda)^+$ by raising or lowering the weights. We have for $i = 1, \ldots, n - 1$:

$$z_{ia} : V(\lambda)_\mu^+ \to V(\lambda)_{\mu + \delta_i}^+, \quad z_{ai} : V(\lambda)_\mu^+ \to V(\lambda)_{\mu - \delta_i}^+,$$

where $\mu \pm \delta_i$ is obtained from $\mu$ by replacing $\mu_i$ with $\mu_i \pm 1$. In the $B$ case the operators $z_{0a}$ preserve each subspace $V(\lambda)_\mu^+$. 
We shall need the following element which can be checked to belong to the normalizer \( \text{Norm} J' \), and so it can be regarded as an element of the algebra \( Z(g_n, g_{n-1}) \):

\[
zn, -n = \sum_{n > i_1 > \cdots > i_s > -n} F_{ni_1} F_{i_1i_2} \cdots F_{i_s, -n} (f_n - f_{j_1}) \cdots (f_n - f_{j_k})
\]

in the \( B, C \) cases, and

\[
zn, -n = \sum_{n > i_1 > \cdots > i_s > -n} F_{ni_1} F_{i_1i_2} \cdots F_{i_s, -n} \frac{(f_n - f_{j_1}) \cdots (f_n - f_{j_k})}{2f_n}
\]

in the \( D \) case, where \( s = 0, 1, \ldots \) and \( \{j_1, \ldots, j_k\} \) is the complement to the subset \( \{i_1, \ldots, i_s\} \) in \( \{-n + 1, \ldots, n - 1\} \). The following is a counterpart of Lemma 2.14 and is crucial in the calculation of the matrix elements of the generators in the bases.

**LEMMA 3.1.** For \( a \in \{-n, n\} \) we have

\[
F_{n-1,a} = \sum_{i=-n+1}^{n-1} zn-i,i-a (f_i - f_{-n+1}) \cdots \wedge_i \cdots (f_i - f_{n-1})
\]

in the \( B, C \) cases, and

\[
F_{n-1,a} = \sum_{i=-n+1}^{n-1} zn-i,i-a (f_i - f_{-n+1}) \cdots \wedge_{-i,j} \cdots (f_i - f_{n-1})
\]

in the \( D \) case, where \( zn-1,n-1 = 1 \) and the equalities are considered in \( U'(g_n) \) modulo the ideal \( J' \). The wedge indicates the indices to be skipped.

In order to write down the basis vectors, introduce the interpolation polynomials \( Z_{n,-n}(u) \) with coefficients in the Mickelsson–Zhelobenko algebra \( Z(g_n, g_{n-1}) \) by

\[
Z_{n,-n}(u) = \sum_{i=1}^{n} z_{ni} z_{i,-n} \prod_{j=1}^{n} \frac{u^2 - g_j^2}{g_i^2 - g_j^2}
\]

(3.10)

in the \( B, C \) cases, and

\[
Z_{n,-n}(u) = \sum_{i=1}^{n} z_{ni} z_{i,-n} \prod_{j=1}^{n-1} \frac{u^2 - g_j^2}{g_i^2 - g_j^2}
\]

(3.11)

in the \( D \) case, where \( g_i = f_i + 1/2 \). Accordingly, we have

\[
Z_{n,-n}(g_i) = z_{ni} z_{i,-n}
\]

(3.12)
with the agreement that when \( u \) is evaluated in \( U(h) \), the coefficients of the polynomial \( Z_{n-u}(u) \) are written to the left of the powers of \( u \), as is the case in the formulas (3.10) and (3.11).

### 3.2. Branching rules, patterns and basis vectors

The restriction of \( V(\lambda) \) to the subalgebra \( g_{n-1} \) is given by

\[
V(\lambda) |_{g_{n-1}} \simeq \bigoplus_{\mu} c(\mu) V'(\mu),
\]

where \( V'(\mu) \) is the irreducible finite-dimensional representation of \( g_{n-1} \) with highest weight \( \mu \). The multiplicity \( c(\mu) \) coincides with the dimension of the space \( V(\lambda)_\mu^+ \), and its exact value is found from the Zhelobenko branching rules, [167]. We formulate them separately for each case recalling the conditions (3.5) and (3.6) on the highest weight \( \lambda \). In the formulas below we use the notation

\[
l_i = \lambda_i + \rho_i + 1/2, \quad \gamma_i = \nu_i + \rho_i + 1/2,
\]

where the \( \nu_i \) are the parameters defined in the branching rules.

A parameterization of basis vectors in \( V(\lambda) \) is obtained by applying the branching rules to its successive restrictions to the subalgebras of the chain

\[
g_1 \subset g_2 \subset \cdots \subset g_{n-1} \subset g_n.
\]

This leads to the definition of the Gelfand–Tsetlin patterns for the \( B, C \) and \( D \) types. Then we give formulas for the basis vectors of the representation \( V(\lambda) \). We use the notation

\[
l_{ki} = \lambda_{ki} + \rho_i + 1/2, \quad l'_{ki} = \lambda'_{ki} + \rho_i + 1/2,
\]

where the \( \lambda_{ki} \) and \( \lambda'_{ki} \) are the entries of the patterns defined below.

#### B type case

The multiplicity \( c(\mu) \) equals the number of \( n \)-tuples \( (\nu_1', \nu_2, \ldots, \nu_n) \) satisfying the inequalities

\[
-\lambda_1 \geq \nu_1' \geq \lambda_1 \geq \nu_2 \geq \lambda_2 \geq \cdots \geq \nu_{n-1} \geq \lambda_{n-1} \geq \nu_n \geq \lambda_n,
-\mu_1 \geq \nu_1' \geq \mu_1 \geq \nu_2 \geq \mu_2 \geq \cdots \geq \nu_{n-1} \geq \mu_{n-1} \geq \nu_n
\]

with \( \nu_1' \) and all the \( \nu_i \) being simultaneously integers or half-integers together with the \( \lambda_i \). Equivalently, \( c(\mu) \) equals the number of \( (n+1) \)-tuples \( \nu = (\sigma, \nu_1, \ldots, \nu_n) \), with the entries given by

\[
(\sigma, \nu_1) = \begin{cases} 
(0, \nu_1') & \text{if } \nu_1' \leq 0, \\
(1, -\nu_1') & \text{if } \nu_1' > 0.
\end{cases}
\]
Lemma 3.2. The vectors

\[ \xi_v = z_{n0}^{n-1} \prod_{i=1}^{n-1} z_{ni}^{\nu_i - \mu_i} \cdot \prod_{k=\lambda}^{\nu_n - 1} Z_{n,-n}(k) \xi \]

form a basis of the space \( V(\lambda) \).

Define the B type pattern \( \Lambda \) associated with \( \lambda \) as an array of the form

\[
\begin{array}{cccc}
\sigma_n & \lambda_n & \lambda_{n-1} & \cdots & \lambda_1 \\
\lambda'_n & \lambda'_{n-1} & \cdots & \cdots & \cdots \\
\sigma_{n-1} & \lambda_{n-1} & \lambda_{n-2} & \cdots & \lambda_2 \\
\lambda'_{n-1} & \lambda'_{n-2} & \cdots & \cdots & \cdots \\
\sigma_1 & \lambda_1 & \lambda_2 & \cdots & \lambda_1 \\
\lambda'_{1} & \lambda'_{2} & \cdots & \cdots & \cdots \\
\end{array}
\]

such that \( \lambda = (\lambda_n, \ldots, \lambda_{nn}) \), each \( \sigma_k \) is 0 or 1, the remaining entries are all nonpositive integers or nonpositive half-integers together with the \( \lambda_i \), and the following inequalities hold

\[
\lambda'_{k1} \geq \lambda_{k1} \geq \lambda'_{k2} \geq \cdots \geq \lambda'_{k,k-1} \geq \lambda_{k,k-1} \geq \lambda'_{kk} \geq \lambda_{kk}
\]

for \( k = 1, \ldots, n \), and

\[
\lambda'_{k1} \geq \lambda_{k-1,1} \geq \lambda'_{k2} \geq \cdots \geq \lambda'_{k,k-1} \geq \lambda_{k-1,k-1} \geq \lambda'_{kk}
\]

for \( k = 2, \ldots, n \). In addition, in the case of the integer \( \lambda_i \) the condition

\[
\lambda'_{k1} \leq -1 \quad \text{if } \sigma_k = 1
\]

should hold for all \( k = 1, \ldots, n \).

Theorem 3.3. The vectors

\[ \xi_{\Lambda} = \prod_{k=1,\ldots,n} \left( z_{k0}^{\sigma_0} \cdot \prod_{i=1}^{k-1} z_{ki}^{\lambda_i - \lambda_{k-1,i}} \cdot \prod_{i=1}^{\nu_n - \lambda_{k,i}} z_{i,-k}^{\lambda_{k,i} - \lambda_{k,i}} \cdot \prod_{j=l_{kk}}^{l_{kk} - 1} Z_{k,-k}(j) \right) \xi \]

parametrized by the patterns \( \Lambda \) form a basis of the representation \( V(\lambda) \).
C type case. The multiplicity $c(\mu)$ equals the number of $n$-tuples of integers $\nu = (\nu_1, \ldots, \nu_n)$ satisfying the inequalities

$$\begin{align*}
0 \geq & \nu_1 \geq \lambda_1 \geq \nu_2 \geq \lambda_2 \geq \cdots \geq \nu_{n-1} \geq \lambda_{n-1} \geq \nu_n \geq \lambda_n, \\
0 \geq & \mu_1 \geq \nu_1 \geq \mu_2 \geq \nu_2 \geq \cdots \geq \nu_{n-1} \geq \mu_{n-1} \geq \nu_n.
\end{align*}$$

(3.13)

**Lemma 3.4.** The vectors

$$\xi_\nu = \prod_{i=1}^{n-1} z_{\nu_i - \mu_i}^{-\nu_i - \lambda_i} \cdot \prod_{k=\nu_n}^\gamma \cdot Z_{\nu_n - \nu}(k) \xi$$

form a basis of the space $V(\lambda)^\perp$.

Define the $C$ type pattern $\Lambda$ associated with $\lambda$ as an array of the form

$$\begin{array}{cccccc}
\lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \\
\lambda'_{n1} & \lambda'_{n2} & \cdots & \lambda'_{nn} \\
\lambda_{n-1,1} & \cdots & \lambda_{n-1,n-1} \\
\lambda'_{n-1,1} & \cdots & \lambda'_{n-1,n-1} \\
\cdots & \cdots & \\
\lambda_{11} \\
\lambda'_{11}
\end{array}$$

such that $\lambda = (\lambda_{n1}, \ldots, \lambda_{nn})$, the remaining entries are all non-positive integers and the following inequalities hold

$$0 \geq \lambda'_{k1} \geq \lambda_{k1} \geq \lambda'_{k2} \geq \lambda_{k2} \geq \cdots \geq \lambda'_{k,k-1} \geq \lambda_{k,k-1} \geq \lambda'_{kk} \geq \lambda_{kk}$$

for $k = 1, \ldots, n$, and

$$0 \geq \lambda'_{k1} \geq \lambda_{k-1,1} \geq \lambda'_{k2} \geq \lambda_{k-1,2} \geq \cdots \geq \lambda'_{k,k-1} \geq \lambda_{k-1,k-1} \geq \lambda'_{kk}$$

for $k = 2, \ldots, n$.

**Theorem 3.5.** The vectors

$$\xi_\Lambda = \prod_{k=1}^n \left( \prod_{i=1}^{k-1} z_{\nu_i - \lambda_{k-1,i}}^{-\nu_i - \lambda_{ki}} \cdot \prod_{j=\nu_f}^\ell \cdot Z_{\nu_n - \nu}(j) \right) \xi$$

parametrized by the patterns $\Lambda$ form a basis of the representation $V(\lambda)$.  

D type case. The multiplicity $c(\mu)$ equals the number of $(n-1)$-tuples $v = (v_1, \ldots, v_{n-1})$ satisfying the inequalities

$$-|\lambda_1| \geq v_1 \geq \lambda_2 \geq v_2 \geq \lambda_3 \geq \cdots \geq \lambda_{n-1} \geq v_{n-1} \geq \lambda_n,$$

$$-|\mu_1| \geq v_1 \geq \mu_2 \geq v_2 \geq \mu_3 \geq \cdots \geq \mu_{n-1} \geq v_{n-1}$$

with all the $v_i$ being simultaneously integers or half-integers together with the $\lambda_i$. Set $v_0 = \max\{\lambda_1, \mu_1\}$.

**Lemma 3.6.** The vectors

$$\xi_v = \prod_{i=1}^{n-1} z_{v_i-\mu_i} z_{v_i-\lambda_i} \cdot \prod_{k=n}^{n-2} Z_{n,-n}(k)\xi$$

form a basis of the space $V(\lambda)^+$. 

Define the D type pattern $\Lambda$ associated with $\lambda$ as an array of the form

$$\begin{array}{cccc}
\lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \\
\lambda'_{n-1,1} & \cdots & \lambda'_{n-1,n-1} \\
\lambda_{n-1,1} & \cdots & \lambda_{n-1,n-1} \\
\cdots & \cdots & \cdots \\
\lambda_{21} & \lambda_{22} & \cdots & \lambda_{2n} \\
\lambda'_{11} & \cdots & \cdots & \cdots \\
\lambda_{11} & \cdots & \cdots & \cdots
\end{array}$$

such that $\lambda = (\lambda_{n1}, \ldots, \lambda_{nn})$, the remaining entries are all nonpositive integers or nonpositive half-integers together with the $\lambda_i$, and the following inequalities hold

$$-|\lambda_{k1}| \geq \lambda'_{k-1,1} \geq \lambda_k 2 \geq \lambda'_{k-2,2} \geq \cdots \geq \lambda_{k,k-1} \geq \lambda'_{k-1,k-1} \geq \lambda_{kk},$$

$$-|\lambda_{k-1,1}| \geq \lambda'_{k-1,1} \geq \lambda_{k-1,2} \geq \lambda'_{k-2,2} \geq \cdots \geq \lambda_{k-1,k-1} \geq \lambda'_{k-1,k-1}$$

for $k = 2, \ldots, n$. Set $\lambda'_{k-1,0} = \max\{\lambda_{k1}, \lambda_{k-1,1}\}$.

**Theorem 3.7.** The vectors

$$\xi_\Lambda = \prod_{k=2,\ldots,n} \left( \prod_{i=1}^{k-1} z_{k_{i,j}} z_{k_{i,j}} \lambda'_{k-1,j,i} \lambda'_{k-1,j,i} \lambda_{k_{i,j}} \lambda_{k_{i,j}} \right) \cdot \prod_{j=l_k}^{l_{k-1,k-1}-2} Z_{k,-k}(j)\xi$$

parametrized by the patterns $\Lambda$ form a basis of the representation $V(\lambda)$. 


Proofs of Theorems 3.3, 3.5 and 3.7 will be outlined in the next two sections. These are based on the application of the representation theory of the twisted Yangians. Clearly, due to the branching rules, it is sufficient to construct a basis in the multiplicity space $V(\lambda)^{\mu}$. 

3.3. Yangians and their representations

We start by introducing the Yangian $Y(2)$ for the Lie algebra $\mathfrak{gl}_2$. In what follows it will be convenient to use the indices $-n, n$ to enumerate the rows and columns of $2 \times 2$-matrices.

The Yangian $Y(2)$ is the complex associative algebra with the generators $t^{(1)}_{ab}, t^{(2)}_{ab}, \ldots$ where $a, b \in \{-n, n\}$, and the defining relations

$$ (u - v)\left[t_{ab}(u), t_{cd}(v)\right] = t_{cb}(u)t_{ad}(v) - t_{cb}(v)t_{ad}(u), $$

where

$$ t_{ab}(u) = \delta_{ab} + t^{(1)}_{ab}u^{-1} + t^{(2)}_{ab}u^{-2} + \cdots \in Y(2)[[u^{-1}]]. $$

Introduce the series $s_{ab}(u)$, $a, b \in \{-n, n\}$ by

$$ s_{ab}(u) = \theta_{n, -n}(u)t_{-b, -n}(u) + \theta_{-n, n}(u)t_{b, n}(u) $$

with $\theta_{ij}$ defined in (3.4). Write

$$ s_{ab}(u) = \delta_{ab} + s^{(1)}_{ab}u^{-1} + s^{(2)}_{ab}u^{-2} + \cdots. $$

The twisted Yangian $Y^{\pm}(2)$ is defined as the subalgebra of $Y(2)$ generated by the elements $s^{(1)}_{ab}, s^{(2)}_{ab}, \ldots$ where $a, b \in \{-n, n\}$. Also, $Y^{\pm}(2)$ can be viewed as an abstract algebra with generators $s_{ab}^{(r)}$ and quadratic and linear defining relations which have the following form

$$ (u^2 - v^2)\left[s_{ab}(u), s_{cd}(v)\right] = (u + v)(s_{cb}(u)s_{ad}(v) - s_{cb}(v)s_{ad}(u)) $$

$$ - (u - v)(\theta_{c, -b}s_{a, -c}(u)s_{-b, d}(v) - \theta_{a, -d}s_{c, -a}(v)s_{-d, b}(u)) $$

$$ + \theta_{a, -b}(s_{c, -a}(u)s_{-b, d}(v) - s_{c, -a}(v)s_{-b, d}(u)) $$

and

$$ \theta_{ab}s_{-b, -a}(-u) = s_{ab}(u) \pm \frac{s_{ab}(u) - s_{ab}(-u)}{2u}. $$

Whenever the double sign $\pm$ or $\mp$ occurs, the upper sign corresponds to the orthogonal case and the lower sign to the symplectic case. In particular, we have the relation

$$ [s_{n, -n}(u), s_{n, -n}(v)] = 0. $$
The Yangian $Y(2)$ is a Hopf algebra with the coproduct
\[
\Delta(t_{ab}(u)) = t_{an}(u) \otimes t_{bn}(u) + t_{a, -n}(u) \otimes t_{-n, b}(u). \tag{3.16}
\]

The twisted Yangian $Y^\pm(2)$ is a left coideal in $Y(2)$ with
\[
\Delta(s_{ab}(u)) = \sum_{c, d \in \{-n, n\}} \theta_{bd} t_{ac}(u) t_{-b, -d}(-u) \otimes s_{cd}(u). \tag{3.17}
\]

Given a pair of complex numbers $(\alpha, \beta)$ such that $\alpha - \beta \in \mathbb{Z}_+$ we denote by $L(\alpha, \beta)$ the irreducible representation of the Lie algebra $\mathfrak{gl}_2$ with highest weight $(\alpha, \beta)$ with respect to the upper triangular Borel subalgebra. Then $\dim L(\alpha, \beta) = \alpha - \beta + 1$. We equip $L(\alpha, \beta)$ with a $Y(2)$-module structure by using the algebra homomorphism $Y(2) \to U(\mathfrak{gl}_2)$ given by
\[
t_{ab}(u) \mapsto \delta_{ab} u - 1 + t(1)_{ab} u - 2 + \cdots + t(k)_{ab}. \tag{3.19}
\]

Any finite-dimensional irreducible $Y(2)$-module is isomorphic to a representation of this type twisted by an automorphism of $Y(2)$ of the form
\[
t_{ab}(u) \mapsto (1 + \varphi_1 u^{-1} + \varphi_2 u^{-2} + \cdots) t_{ab}(u), \quad \varphi_i \in \mathbb{C}.
\]

There is an explicit irreducibility criterion for the $Y(2)$-module $L$. To formulate the result, with each $L(\alpha, \beta)$ associate the string
\[
S(\alpha, \beta) = \{\beta, \beta + 1, \ldots, \alpha - 1\} \subset \mathbb{C}.
\]

We say that two strings $S_1$ and $S_2$ are in general position if

- either $S_1 \cup S_2$ is not a string, or $S_1 \subseteq S_2$, or $S_2 \subseteq S_1$.

**Theorem 3.8.** *The representation (3.18) of $Y(2)$ is irreducible if and only if the strings $S(\alpha_i, \beta_i)$, $i = 1, \ldots, k$, are pairwise in general position.*

Note that the generators $t_{ab}^{(r)}$ with $r > k$ act as zero operators in $L$. Therefore, the operators $T_{ab}(u) = u^k t_{ab}(u)$ are polynomials in $u$:
\[
T_{ab}(u) = \delta_{ab} u^k + t_{ab}^{(1)} u^{k-1} + \cdots + t_{ab}^{(k)}. \tag{3.19}
\]

Let $\xi_i$ denote the highest vector of the $\mathfrak{gl}_2$-module $L(\alpha_i, \beta_i)$. Suppose that the $Y(2)$-module $L$ given by (3.18) is irreducible and the strings $S(\alpha_i, \beta_i)$ are pairwise disjoint. Set
\[
\eta = \xi_1 \otimes \cdots \otimes \xi_k. \tag{3.20}
\]
Then using (3.16) we easily check that $\eta$ is the highest vector of the $Y(2)$-module $L$. That is, $\eta$ is annihilated by $T_{-n,n}(u)$, and it is an eigenvector for the operators $T_{n,n}(u)$ and $T_{n,-n}(u)$. Explicitly,

$$
T_{-n,-n}(u)\eta = (u + \alpha_1) \cdots (u + \alpha_k)\eta, \\
T_{n,n}(u)\eta = (u + \beta_1) \cdots (u + \beta_k)\eta.
$$

Let a $k$-tuple $\gamma = (\gamma_1, \ldots, \gamma_k)$ satisfy the following conditions: for each $i$

$$
\alpha_i - \gamma_i \in \mathbb{Z}_+, \quad \gamma_i - \beta_i \in \mathbb{Z}_+.
$$

Set

$$
\eta_{\gamma} = \prod_{i=1}^{k} T_{n,-n}(-\gamma_i + 1) \cdots T_{n,-n}(-\beta_i - 1)T_{n,-n}(-\beta_i). \tag{3.21}
$$

The following theorem provides a Gelfand–Tsetlin type basis for representations of the Yangian $Y(2)$ associated with the embedding $Y(1) \subset Y(2)$. Here $Y(1)$ is the (commutative) subalgebra of $Y(2)$ generated by the elements $t_{nn}^{(r)}$, $r \geq 1$.

**Theorem 3.9.** Let the $Y(2)$-module $L$ given by (3.18) be irreducible and let the strings $S(\alpha_i, \beta_i)$ be pairwise disjoint. Then the vectors $\eta_{\gamma}$ with $\gamma$ satisfying (3.22) form a basis of $L$. Moreover, the generators of $Y(2)$ act in this basis by the rules

$$
T_{n,n}(u)\eta_{\gamma} = (u + \gamma_1) \cdots (u + \gamma_k)\eta_{\gamma}, \\
T_{n,-n}(-\gamma_i)\eta_{\gamma} = \eta_{\gamma+\delta_i}, \\
T_{n,-n}(\gamma_i)\eta_{\gamma} = - \prod_{m=1}^{k} (\alpha_m - \gamma_i + 1)(\beta_m - \gamma_i)\eta_{\gamma-\delta_i}, \\
T_{n,-n}(u)\eta_{\gamma} = \prod_{i=1}^{k} (u + \alpha_i + 1)(u + \beta_i) \frac{1}{u + \gamma_i + 1}\eta_{\gamma} \\
+ \prod_{i=1}^{k} \frac{1}{u + \gamma_i + 1}T_{n,n}(u)T_{n,-n}(u + 1)\eta_{\gamma}. \tag{3.23}
$$

These formulas are derived from the defining relations for the Yangian (3.14) with the use of the quantum determinant

$$
d(u) = T_{n,-n}(u + 1)T_{n,n}(u) - T_{n,-n}(u + 1)T_{n,n}(u) \\
= T_{n,-n}(u)T_{n,n}(u + 1) - T_{n,n}(u)T_{n,-n}(u + 1). \tag{3.24}
$$
The coefficients of the quantum determinant belong to the center of $Y(2)$ and so $d(u)$ acts in $L$ as a scalar which can be found by the application of (3.24) to the highest vector $\eta$. Indeed, by (3.21)

$$d(u)\eta = (u + \alpha_1 + 1) \cdots (u + \alpha_k + 1)(u + \beta_1) \cdots (u + \beta_k)\eta.$$ 

This allows us to derive the last formula in (3.23) from (3.25). The operators $T_{-n,n}(u)$ and $T_{n,-n}(u)$ are polynomials in $u$ of degree $\leq k - 1$; see (3.19). Therefore, their action can be found from (3.23) by using the Lagrange interpolation formula.

We can regard (3.18) as a module over the twisted Yangian $Y^- (2)$ obtained by restriction. An irreducibility criterion for such a module is provided by the following theorem.

**Theorem 3.10.** The representation (3.18) of $Y^- (2)$ is irreducible if and only if each pair of strings

$$S(\alpha_i, \beta_i), S(\alpha_j, \beta_j) \quad \text{and} \quad S(\alpha_i, \beta_i), S(-\beta_j, -\alpha_j)$$

is in general position for all $i < j$.

The defining relations (3.14) allow us to rewrite formula (3.15) for $s_{n,-n}(u)$ in the form

$$s_{n,-n}(u) = \frac{u + 1/2}{u} \left( t_{n,-n}(u) t_{nn}(-u) - t_{n,-n}(-u) t_{nn}(u) \right).$$

Therefore the operator in $L$ defined by

$$S_{n,-n}(u) = \frac{u^{2k}}{u + 1/2} s_{n,-n}(u)$$

$$= \frac{(-1)^k}{u} \left( T_{n,-n}(u) T_{nn}(-u) - T_{n,-n}(-u) T_{nn}(u) \right)$$

(3.26)

is an even polynomial in $u$ of degree $\leq 2k - 2$. Its action in the basis of $L$ provided in Theorem 3.9 is given by

$$S_{n,-n}(\gamma_i)\eta_\gamma = 2 \prod_{a=1, a \neq i}^k (-\gamma_i - \gamma_a)\eta_{\gamma + \delta_i}, \quad i = 1, \ldots, k.$$ 

We have thus proved the following corollary.

**Corollary 3.11.** Suppose that the $Y^- (2)$-module $L$ is irreducible and we have

$$S(\alpha_i, \beta_i) \cap S(\alpha_j, \beta_j) = \emptyset \quad \text{and} \quad S(\alpha_i, \beta_i) \cap S(-\beta_j, -\alpha_j) = \emptyset$$
for all $i < j$. Then the vectors

$$\xi_\gamma = \prod_{i=1}^{k} S_{n,-n}(\gamma_i - 1) \cdots S_{n,-n}(\beta_i + 1)S_{n,-n}(\beta_i)\eta$$

with $\gamma$ satisfying (3.22) form a basis of $L$.

Let us now turn to the orthogonal twisted Yangian $Y^+(2)$. For any $\delta \in \mathbb{C}$ denote by $W(\delta)$ the one-dimensional representation of $Y^+(2)$ spanned by a vector $w$ such that

$$s_{n,n}(u)w = \frac{u + \delta}{u + 1/2}w, \quad s_{-n,-n}(u)w = \frac{u - \delta + 1}{u + 1/2}w,$$

and $s_{n,-a}(u)w = 0$ for $a = -n, n$. By (3.17) we can regard the tensor product $L \otimes W(\delta)$ as a representation of $Y^+(2)$. The representations of $Y^+(2)$ of this type, and the representations of $Y^-(2)$ of type (3.18) essentially exhaust all finite-dimensional irreducible representations of $Y^\pm(2)$, [97].

The following is an analog of Theorem 3.10.

**THEOREM 3.12.** The representation $L \otimes W(\delta)$ of $Y^+(2)$ is irreducible if and only if each pair of strings

$$S(\alpha_i, \beta_i), S(\alpha_j, \beta_j) \quad \text{and} \quad S(\alpha_i, -\beta_j), S(-\beta_j, -\alpha_j)$$

is in general position for all $i < j$, and none of the strings $S(\alpha_i, \beta_i)$ or $S(-\beta_i, -\alpha_i)$ contains $-\delta$.

Using the vector space isomorphism

$$L \otimes W(\delta) \to L, \quad v \otimes w \mapsto v, \quad v \in L,$$  \hspace{1cm} (3.27)

we can regard $L$ as a $Y^+(2)$-module. Accordingly, using the defining relations (3.14) and the coproduct formula (3.17) we can write $s_{n,-n}(u)$, as an operator in $L$, in the form

$$s_{n,-n}(u) = \frac{u - \delta}{u}l_{n,-n}(u)l_{nn}(-u) + \frac{u + \delta}{u}l_{n,-n}(-u)l_{nn}(u).$$

Therefore the operator in $L$ defined by

$$S_{n,-n}(u) = u^{2k}s_{n,-n}(u)$$

$$= \frac{(-1)^k}{u}((u - \delta)T_{n,-n}(u)T_{nn}(-u) + (u + \delta)T_{n,-n}(-u)T_{nn}(u))$$  \hspace{1cm} (3.28)

The second condition was erroneously omitted in the formulation of [98, Proposition 4.2] although it is implicit in the proof.
is an even polynomial in $u$ of degree $\leq 2k - 2$. Its action in the basis of $L$ provided in Theorem 3.9 is given by

$$S_{n,-n}(\gamma_i)\eta_{\gamma} = 2(-\delta - \gamma_i) \prod_{a=1, a\neq i}^{k} (-\gamma_i - \gamma_a)\eta_{\gamma + \delta_i}, \quad i = 1, \ldots, k.$$ 

We have thus proved the following corollary.

**COROLLARY 3.13.** Suppose that the $Y^+(2)$-module $L \otimes W(\delta)$ is irreducible and we have

$$S(\alpha_i, \beta_i) \cap S(\alpha_j, \beta_j) = \emptyset \quad \text{and} \quad S(\alpha_i, \beta_i) \cap S(-\beta_j, -\alpha_j) = \emptyset$$

for all $i < j$. Then the vectors

$$\xi_{\gamma} = \prod_{i=1}^{k} S_{n,-n}(\gamma_i - 1) \cdots S_{n,-n}(\beta_i + 1)S_{n,-n}(\beta_i)\eta$$

with $\gamma$ satisfying (3.22) form a basis of $L$.

### 3.4. Yangian action on the multiplicity space

Now we construct an algebra homomorphism $Y^\pm(2) \to \mathbb{Z}(\mathfrak{g}_n, \mathfrak{g}_{n-1})$ and then use it to define an action of $Y^\pm(2)$ on the multiplicity space $V(\lambda)^+_{\mu}$.

For $a, b \in \{-n, n\}$ and a complex parameter $u$ introduce the elements $Z_{ab}(u)$ of the Mickelsson–Zhelobenko algebra $\mathbb{Z}(\mathfrak{g}_n, \mathfrak{g}_{n-1})$ by

$$Z_{ab}(u) = -\left(\delta_{ab}(u + \rho_n + \frac{1}{2}) + F_{ab}\right) \prod_{i=-n+1}^{n-1} (u + g_i)$$

$$+ \sum_{i=-n+1}^{n-1} \sum_{j=-n+1, j\neq i}^{n-1} z_{ai}z_{ib} \prod_{j=-n+1, j\neq i}^{n-1} \frac{u + g_j}{g_i - g_j}$$

(3.29)

in the $B$ case,

$$Z_{ab}(u) = \left(\delta_{ab}(u + \rho_n + \frac{1}{2}) + F_{ab}\right) \prod_{i=-n+1}^{n-1} (u + g_i)$$

$$- \sum_{i=-n+1}^{n-1} \sum_{j=-n+1, j\neq i}^{n-1} z_{ai}z_{ib} \prod_{j=-n+1, j\neq i}^{n-1} \frac{u + g_j}{g_i - g_j}$$

(3.30)
in the $C$ case, and

$$
Z_{ab}(u) = -\left(\delta_{ab}(u + \rho_n + \frac{1}{2}) + F_{ab}\right) \prod_{i=-n+1}^{n-1} (u + g_i) - \sum_{i=-n+1}^{n-1} z_{ai} z_{ib}(u + g_{-i}) \prod_{j=-n+1, j \neq \pm i}^{n-1} \frac{u + g_j}{g_i - g_j} \frac{1}{2u + 1}
$$

(3.31)

in the $D$ case, where $g_i = f_i + 1/2$ for all $i$. In particular, it can be verified that each $Z_{n,-n}(u)$ coincides with the corresponding interpolation polynomial given in (3.10) or (3.11).

Consider now the three cases separately. We shall assume $\mu_n = -\infty$ in the notation below.

**B type case.**

**Theorem 3.14.**

(i) The mapping

$$
s_{ab}(u) \mapsto -u^{-2n}Z_{ab}(u), \quad a, b \in \{-n, n\},
$$

(3.32)

defines an algebra homomorphism $Y^+(2) \to Z(g_n, g_{n-1})$.

(ii) The $Y^+(2)$-module $V(\lambda)_{\mu}^+$ defined via the homomorphism (3.32) is isomorphic to the direct sum of two irreducible submodules, $V(\lambda)_{\mu}^+ \simeq U \oplus U'$, where

$$
U = L(0, \beta_1) \otimes L(\alpha_2, \beta_2) \otimes \cdots \otimes L(\alpha_n, \beta_n) \otimes W(1/2),
$$

$$
U' = L(-1, \beta_1) \otimes L(\alpha_2, \beta_2) \otimes \cdots \otimes L(\alpha_n, \beta_n) \otimes W(1/2),
$$

if the $\lambda_i$ are integers (it is supposed that $U' = \{0\}$ if $\beta_1 = 0$); or

$$
U = L(-1/2, \beta_1) \otimes L(\alpha_2, \beta_2) \otimes \cdots \otimes L(\alpha_n, \beta_n) \otimes W(0),
$$

$$
U' = L(-1/2, \beta_1) \otimes L(\alpha_2, \beta_2) \otimes \cdots \otimes L(\alpha_n, \beta_n) \otimes W(1),
$$

if the $\lambda_i$ are half-integers, and the following notation is used

$$
\alpha_i = \min\{\lambda_{i-1}, \mu_{i-1}\} - i + 1, \quad i = 2, \ldots, n,
$$

$$
\beta_i = \max\{\lambda_i, \mu_i\} - i + 1, \quad i = 1, \ldots, n.
$$
C type case.

**Theorem 3.15.**

(i) The mapping

\[ s_{ab}(u) \mapsto (u + 1/2)u^{-2n}Z_{ab}(u), \quad a, b \in \{-n, n\}, \quad (3.33) \]

defines an algebra homomorphism \( Y^- (2) \to Z(g_n, g_{n-1}) \).

(ii) The \( Y^- (2) \)-module \( V(\lambda) \) defined via the homomorphism (3.33) is irreducible and isomorphic to the tensor product

\[ L(\alpha_1, \beta_1) \otimes \cdots \otimes L(\alpha_n, \beta_n), \]

where \( \alpha_1 = -1/2 \) and

\[
\begin{align*}
\alpha_i &= \min\{\lambda_{i-1}, \mu_{i-1}\} - i + 1/2, \quad i = 2, \ldots, n, \\
\beta_i &= \max\{\lambda_i, \mu_i\} - i + 1/2, \quad i = 1, \ldots, n.
\end{align*}
\]

D type case.

**Theorem 3.16.**

(i) The mapping

\[ s_{ab}(u) \mapsto -2u^{-2n+2}Z_{ab}(u), \quad a, b \in \{-n, n\} \quad (3.34) \]

defines an algebra homomorphism \( Y^+ (2) \to Z(g_n, g_{n-1}) \).

(ii) The \( Y^+ (2) \)-module \( V(\lambda) \) defined via the homomorphism (3.34) is irreducible and isomorphic to the tensor product

\[ L(\alpha_1, \beta_1) \otimes \cdots \otimes L(\alpha_{n-1}, \beta_{n-1}) \otimes W(-\alpha_0), \]

where \( \alpha_1 = \min\{-|\lambda_1|, -|\mu_1|\} - 1/2, \alpha_0 = \alpha_1 + |\lambda_1 + \mu_1|, \)

\[
\begin{align*}
\alpha_i &= \min\{\lambda_i, \mu_i\} - i + 1/2, \quad i = 2, \ldots, n - 1, \\
\beta_i &= \max\{\lambda_{i+1}, \mu_{i+1}\} - i + 1/2, \quad i = 1, \ldots, n - 1.
\end{align*}
\]

**Outline of the Proof.** Part (i) of Theorems 3.14–3.16 is verified by using the composition of homomorphisms

\( Y^\pm (2) \to C_n \to Z(g_n, g_{n-1}) \),

where \( C_n \) is the centralizer \( U(g_n)_{g_{n-1}} \). The first arrow is the homomorphism provided by the centralizer construction (see [105,118]) while the second is the natural projection.
By the results of [97], every irreducible finite-dimensional representation of the twisted Yangian is a highest weight representation. It contains a unique, up to a constant factor, vector which is annihilated by $s_{-n,n}(u)$ and which is an eigenvector of $s_{nn}(u)$. The corresponding eigenvalue (the highest weight) uniquely determines the representation. The vectors in $V(\lambda)_\mu^+$ annihilated by $s_{-n,n}(u)$ can be explicitly constructed by using the lowering operators. One of these vectors is given by

$$
\xi_\mu = \prod_{i=1}^{n-1} (z_{\max\{\lambda_i,\mu_i\} - \mu_i} - z_{\max\{\lambda_i,\mu_i\} - \lambda_i}) \xi,
$$

where $\xi$ is the highest vector of $V(\lambda)$. This is the only vector in the C, D cases, while in the B case there is another one defined by

$$
\xi'_\mu = z_{n0} \xi_\mu.
$$

Calculating the eigenvalues of these vectors we conclude that they respectively coincide with the eigenvalues of the tensor product of the highest vectors of the modules $L(\alpha_i, \beta_i)$; see (3.20).

REMARK 3.17. Theorems 3.14–3.16 can be proved without using the branching rules for the reductions $sp_{2n} \downarrow sp_{2n-2}$ and $o_N \downarrow o_{N-2}$. Therefore, the reduction multiplicities can be found by calculating the dimension of the space $V(\lambda)_\mu^+$. For instance, in the symplectic case, Theorem 3.15 gives

$$
c(\mu) = \prod_{i=1}^{n} (\alpha_i - \beta_i + 1)
$$

which, of course, coincides with the value provided by the C type branching rule; see Section 3.2.

While keeping $\lambda$ and $\mu$ fixed we let $\nu$ run over the values determined by the branching rules; see Section 3.2. Using the homomorphisms of Theorems 3.14–3.16 we conclude from (3.26) and (3.28) that the element $S_{-n,-n}(u)$ acts in the representation $V(\lambda)_\mu^+$ precisely as the operator $-Z_{-n,-n}(u)$, $Z_{-n,-n}(u)$, or $-2Z_{-n,-n}(u)$, in the B, C or D cases, respectively. Thus, by Corollaries 3.11 and 3.13, the following vectors $\xi_\nu$ form a basis of the space $V(\lambda)_\mu^+$, where

$$
\xi_\nu = z_{n0}^\nu \prod_{i=1}^{n} Z_{-n,-n}(\gamma_i - 1) \cdots Z_{-n,-n}(\beta_i + 1) Z_{-n,-n}(\beta_i) \xi_\mu
$$

in the B case,

$$
\xi_\nu = \prod_{i=1}^{n} Z_{-n,-n}(\gamma_i - 1) \cdots Z_{-n,-n}(\beta_i + 1) Z_{-n,-n}(\beta_i) \xi_\mu
$$

(3.35)
in the $C$ case, and

$$
\xi_v = \prod_{i=1}^{n-1} Z_{n,-n}(\gamma_i - 1) \cdots Z_{n,-n}(\beta_i + 1) Z_{n,-n}(\beta_i) \xi_\mu
$$

in the $D$ case. Applying the interpolation properties of the polynomials $Z_{n,-n}(u)$ we bring the above formulas to the form given in Lemmas 3.2, 3.4 and 3.6, respectively. Clearly, Theorems 3.3, 3.5 and 3.7 follow.

3.5. Calculation of the matrix elements

Without writing down all explicit formulas we shall demonstrate how the matrix elements of the generators of $g_n$ in the basis $\xi_\Lambda$ provided by Theorems 3.3, 3.5 and 3.7 can be calculated. The interested reader is referred to the papers [98–100] for details. We choose the following generators

$$
F_{k-1,-k}, \quad F_{k-1,k}, \quad k = 1, \ldots, n,
$$
in the $B$ case,

$$
F_{k-1,-k}, \quad k = 2, \ldots, n, \quad \text{and} \quad F_{-k,k}, \quad F_{k,-k}, \quad k = 1, \ldots, n,
$$
in the $C$ case, and

$$
F_{k-1,-k}, \quad F_{k-1,k}, \quad k = 2, \ldots, n, \quad \text{and} \quad F_{21}, \quad F_{-2,1}
$$
in the $D$ case.

In the symplectic case the elements $F_{kk}, F_{k,-k}, F_{-k,k}$ commute with the subalgebra $g_{k-1}$ in $U(g_k)$. Therefore, these operators preserve the subspace of $g_{k-1}$-highest vectors in $V(\lambda)$. So, it suffices to compute the action of these operators with $k = n$ in the basis $\{\xi_v\}$ of the space $V(\lambda)^+_{\mu}$; see Lemma 3.4. For $F_{nn}$ we immediately get

$$
F_{nn} \xi_v = \left( 2 \sum_{i=1}^{n} \nu_i - \sum_{i=1}^{n} \lambda_i - \sum_{i=1}^{n-1} \mu_i \right) \xi_v.
$$

Further, by (3.35)

$$
Z_{n,-n}(\gamma_i) \xi_v = \xi_{v+i+\delta_i}, \quad i = 1, \ldots, n.
$$

However, $Z_{n,-n}(u)$ is a polynomial in $u^2$ of degree $n - 1$ with the highest coefficient $F_{n,-n}$. Applying the Lagrange interpolation formula with the interpolation points $\gamma_i, i = 1, \ldots, n$, we obtain

$$
Z_{n,-n}(u) \xi_v = \sum_{i=1}^{n} \prod_{a=1}^{n} \frac{u^2 - \gamma_a^2}{\gamma_i^2 - \gamma_a^2} \xi_{v+i+\delta_i}.
$$
Taking here the coefficient at \( u^{2n-2} \) we get
\[
F_{n,-n}\xi_v = \sum_{i=1}^{n} \prod_{a=1, a \neq i}^{n} \frac{1}{\gamma_i^2 - \gamma_a^2} \xi_{v+i}, \tag{3.36}
\]

The action of \( F_{-n,n} \) is found in a similar way with the use of Theorem 3.9.

In the orthogonal case the action of \( F_{nn} \) is found in the same way. However, the elements \( F_{n,-n} \) and \( F_{-n,n} \) are zero. We shall use second-order elements of the enveloping algebra instead. These are given by
\[
\Phi_{-a,a} = \frac{1}{2} \sum_{i=-n+1}^{n-1} F_{-a,i} F_{ia}
\]
with \( a \in \{-n, n\} \). The elements \( \Phi_{-a,a} \) commute with the subalgebra \( g_{n-1} \) so that, like in the symplectic case, they preserve the space \( V(\lambda)^{g_+} \) and their action in the basis \( \{\xi_i\} \) is given by formulas similar to those for \( F_{-a,a} \).

The calculation of the matrix elements of the generators \( F_{k-1, -k} \) is similar in all three cases. We may assume \( k = n \). The operator \( F_{n-1, -n} \) preserves the subspace of \( g_{n-2} \) highest vectors in \( V(\lambda) \). Consider the symplectic case as an example. Suppose that \( \mu' \) is a fixed \( g_{n-2} \) highest weight, \( \nu' \) is an \((n-1)\)-tuple of integers such that the inequalities (3.13) are satisfied with \( \lambda, \nu, \mu \) respectively replaced by \( \mu, \nu', \mu' \), and set \( \gamma_i' = \nu_i' + \rho_i + 1/2 \). It suffices to calculate the action of \( F_{n-1, -n} \) on the basis vectors of the form
\[
\xi_{\nu'\mu'\nu'} = X_{\nu'\mu'} \xi_{\nu'\mu'},
\]
where \( \xi_{\nu'\mu'} = \xi_v \) and \( X_{\nu'\mu'} \) denotes the operator
\[
X_{\nu'\mu'} = \prod_{i=1}^{n-2} \frac{\nu_i' - \mu_i}{\gamma_{n-1,i} - \gamma_{n-1,-i}}, \prod_{a=m_{n-1}} Z_{n-1, -n+1}(a),
\]
where we have used the notation \( m_i = \mu_i + \rho_i + 1/2 \). The operator \( F_{n-1, -n} \) is permutable with the elements \( z_{n-1,i} \) and \( Z_{n-1, -n+1}(n) \). Hence, we can write
\[
F_{n-1, -n}\xi_{\nu'\mu'} = X_{\nu'\mu'} F_{n-1, -n}\xi_{\nu'\mu'}.
\]
Now we apply Lemma 3.1. It remains to calculate \( z_{n}\xi_{\nu'\mu'} \) and \( X_{\nu'\mu'} z_{n-1, -i} \). Using the relations between the elements of the Mickelsson–Zhelobenko algebra \( Z(\mathfrak{g}_n, \mathfrak{g}_{n-1}) \) given in (3.9), we find that
\[
z_{n}\xi_{\nu'\mu'} = \xi_{\nu', \mu - \delta_i}
\]
if \( i > 0 \). Otherwise, if \( i = -j \) with positive \( j \), write
\[
z_{n, -j}\xi_{\nu'\mu'} = z_{n, -j} z_{a} \xi_{\nu', \mu + \delta_j} = Z_{n, -n}(m_j) \xi_{\nu', \mu + \delta_j}, \tag{3.37}
\]
where we have used the interpolation properties (3.12) of the polynomials $Z_{n,n}(u)$. Finally, we use the expression (3.35) of the basis vectors and Theorem 3.9 to present (3.37) as a linear combination of basis vectors. The same argument applies to calculate $X_{\mu\nu} Z_{n-1,-i}$.

The final formulas for the matrix elements of the generators $F_{n-1,-n}$ in all the three cases are given by multiplicative expressions in the entries of the patterns which exhibit some similarity to the formulas of Theorem 2.3.

In the orthogonal case we also need to find the action of the generators $F_{n-1,-n}$. Unlike the case of the generators $F_{n-1,-n}$, the corresponding matrix elements will be given by certain combinations of multiplicative expressions for which it does not seem to be possible to bring them into a product form. There are two alternative ways to calculate these combinations which we briefly outline below. First, as in the previous calculation, we can write

$$F_{n-1,n} z_{\nu\mu} = X_{\mu\nu} F_{n-1,n} z_{\nu\mu}.$$  

Applying again Lemma 3.1, we come to the calculation of $z_{\nu\mu}$. This time the interpolation property of $Z_{n-1,-n}(u)$ (see (3.29) and (3.31)) allows us to write, e.g., for $i > 0$

$$z_{\nu\mu} = z_{\nu n} z_{\nu,\mu+\delta_i} = z_{n,-i} z_{i,-n} z_{\nu,\mu+\delta_i} = Z_{n,-n}(m_i) z_{\nu,\mu+\delta_i}.$$  

Now, as $Z_{n,-n}(u)$ is, up to a multiple, the image of $S_{n,-n}(u)$ under the homomorphism $Y^+(2) \to Z(g_n, g_{n-1})$, we can express this operator in terms of the Yangian operators $T_{ab}(u)$ and then apply Theorem 3.9 to calculate its action.

Alternatively, the generator $F_{n-1,n}$ can be written modulo the left ideal $J$ of $U'(g_n)$ as

$$F_{n-1,n} = \Phi_{n-1,-n}(2) \Phi_{n,n} - \Phi_{n,n} \Phi_{n-1,-n}(0),$$  

where

$$\Phi_{n-1,n}(u) = \sum_{i=-n+1}^{n-1} z_{n-1,i} z_{i,-n} \prod_{a=-n+1, a \neq i}^{n-1} \frac{1}{f_i - f_a} \cdot \frac{1}{u + f_i + F_{nn}}$$  

in the $B$ case, and

$$\Phi_{n-1,n}(u) = \sum_{i=-n+1}^{n-1} z_{n-1,i} z_{i,-n} \prod_{a=-n+1, a \neq \pm i}^{n-1} \frac{1}{f_i - f_a} \cdot \frac{1}{u + f_i + F_{nn}}$$  

in the $D$ case. The action of $\Phi_{n-1,-n}(u)$ is found exactly as that of $F_{n-1,-n}$ and the matrix elements have a similar multiplicative form. Note, however, that formula (3.38), regarded as an equality of operators acting on $V(\lambda)^+$, is only valid provided the denominators in (3.39) or (3.40) do not vanish. Therefore, in order to use (3.38), we first consider $V(\lambda)$ with ‘generic’ entries of $\lambda$ and calculate the matrix elements of $F_{n-1,n}$ as functions in the entries of the patterns $\Lambda$. The final explicit formulas can be written in a singularity-free form and they are valid in the general case.
Bibliographical notes

The exposition here is based upon the author’s papers [98–100]. Slight changes in the notation were made in order to present the results in a uniform manner for all the three cases. The branching rules for all classical reductions $\mathfrak{o}_N \downarrow \mathfrak{o}_{N-1}$ and $\mathfrak{sp}_{2n} \downarrow \mathfrak{sp}_{2n-2}$ are due to Zhelobenko, [167]; see also Hegerfeldt, [58], King, [68], Proctor, [138], Okounkov, [111], Goodman and Wallach, [45]. The lowering operators for the symplectic Lie algebras were first constructed by Mickelsson, [91]; see also Bincer, [9]. The explicit relations in the algebra $\mathbb{Z}(\mathfrak{sp}_{2n}, \mathfrak{sp}_{2n-2})$ were calculated by Zhelobenko, [170].

The algebra $Y(n)$ was first studied in the work of Faddeev and the St.-Petersburg school in relation with the inverse scattering method; see, for instance, Takhtajan and Faddeev, [147], Kulish and Sklyanin, [74]. The term “Yangian” was introduced by Drinfeld in [28]. In that paper he defined the Yangian $Y(\mathfrak{a})$ for each simple finite-dimensional Lie algebra $\mathfrak{a}$. Finite-dimensional irreducible representations of $Y(\mathfrak{a})$ were classified by Drinfeld, [29], with the use of a previous work by Tarasov, [148,149]. Theorem 3.9 goes back to this work of Tarasov; see also [96,110]. The criterion of Theorem 3.8 is due to Chari and Pressley, [13]. It can also be deduced from the results of [148,149]; see [97]. The twisted Yangians were introduced by Olshanski, [118]; see also [104]. Their finite-dimensional irreducible representations were classified in the author’s paper [97] which, in particular, contains the criteria of Theorems 3.10 and 3.12. For more details on the (twisted) Yangians and their applications in classical representation theory see the expository papers [104,103] and the recent work of Nazarov, [107,108], where, in particular, the skew representations of twisted Yangians were studied.

In some particular cases, bases in $V(\lambda)$ were constructed, e.g., by Wong and Yeh, [165], Smirnov and Tolstoy, [143].

Weight bases for the fundamental representations of $\mathfrak{o}_{2n+1}$ and $\mathfrak{sp}_{2n}$ were independently constructed by Donnelly, [24–26], in a different way. He also demonstrated that these bases of his coincide with those of Theorems 3.3 and 3.5, up to a diagonal equivalence.

Harada, [57], employed the results of [98] to construct a new integrable (Gelfand–Tsetlin) system on the coadjoint orbits of the symplectic groups. This provides an analog of the Guillemin–Sternberg construction, [55], for the unitary groups.

4. Gelfand–Tsetlin bases for representations of $\mathfrak{o}_N$

In this section we sketch the construction of the bases proposed originally by Gelfand and Tsetlin in [42]. It is based upon the fact that the restriction $\mathfrak{o}_N \downarrow \mathfrak{o}_{N-1}$ is multiplicity-free. This makes the construction similar to the $\mathfrak{gl}_N$ case. We shall be applying the general method of Mickelsson algebras outlined in Section 2.2. In particular, the corresponding branching rules can be derived from Theorem 2.11; cf. Section 2.3.

It will be convenient to change the notation for the elements of the orthogonal Lie algebra $\mathfrak{o}_N$ used in Section 3. We shall now use the standard enumeration of the rows and columns of $N \times N$-matrices by the numbers $\{1, \ldots, N\}$. Define an involution of this set of indices by setting $i' = N - i + 1$. The Lie algebra $\mathfrak{o}_N$ is spanned by the elements

$$F_{ij} = E_{ij} - E_{j'i'}, \quad i, j = 1, \ldots, N.$$  (4.1)
We shall keep the notation $g_n$ for $\mathfrak{o}_N$ with $N = 2n + 1$ or $N = 2n$.

The finite-dimensional irreducible representations of $g_n$ are now parametrized by $n$-tuples $\lambda = (\lambda_1, \ldots, \lambda_n)$ where the numbers $\lambda_i$ satisfy the conditions

$$\lambda_i - \lambda_{i+1} \in \mathbb{Z}_+ \quad \text{for} \quad i = 1, \ldots, n - 1,$$

and

$$2\lambda_n \in \mathbb{Z}_+ \quad \text{for} \quad g_n = \mathfrak{o}_{2n+1},$$

$$\lambda_{n-1} + \lambda_n \in \mathbb{Z}_+ \quad \text{for} \quad g_n = \mathfrak{o}_{2n}.$$  \hfill (4.2)

Such an $n$-tuple $\lambda$ is called the highest weight of the corresponding representation which we shall denote by $V(\lambda)$. It contains a unique, up to a constant factor, nonzero vector $\xi$ (the highest vector) such that

$$F_{ii} \xi = \lambda_i \xi \quad \text{for} \quad i = 1, \ldots, n,$$

$$F_{ij} \xi = 0 \quad \text{for} \quad 1 \leq i < j \leq N.$$  \hfill (4.3)

4.1. Lowering operators for the reduction $\mathfrak{o}_{2n+1} \downarrow \mathfrak{o}_{2n}$

Taking $N = 2n + 1$ in the definition (4.1), we shall consider $\mathfrak{o}_{2n}$ as the subalgebra of $\mathfrak{o}_{2n+1}$ spanned by the elements (4.1) with $i, j \neq n + 1$. In accordance with the branching rule, the restriction of $V(\lambda)$ to the subalgebra $\mathfrak{o}_{2n}$ is given by

$$V(\lambda)|_{\mathfrak{o}_{2n}} \simeq \bigoplus_{\mu} V'(\mu),$$

where $V'(\mu)$ is the irreducible finite-dimensional representation of $\mathfrak{o}_{2n}$ with highest weight $\mu$ and the sum is taken over the weights $\mu$ satisfying the inequalities

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \lambda_{n-1} \geq \mu_{n-1} \geq \lambda_n \geq |\mu_n|,$$  \hfill (4.4)

with all the $\mu_i$ being simultaneously integers or half-integers together with the $\lambda_i$.

The elements $F_{n+1,i}$ span the $\mathfrak{o}_{2n}$-invariant complement to $\mathfrak{o}_{2n}$ in $\mathfrak{o}_{2n+1}$. Therefore, by the general theory of Section 2.2, the Mickelsson–Zhelobenko algebra $Z(\mathfrak{o}_{2n+1}, \mathfrak{o}_{2n})$ is generated by the elements

$$p F_{n+1,i}, \quad i = 1, \ldots, n, n', \ldots, 1',$$  \hfill (4.5)

where $p$ is the extremal projector for the Lie algebra $\mathfrak{o}_{2n}$. Let $\{\varepsilon_1, \ldots, \varepsilon_n\}$ be the basis of $\mathfrak{h}^*$ dual to the basis $\{F_{11}, \ldots, F_{nn}\}$ of the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{o}_{2n}$. Set $\varepsilon_{i'} = -\varepsilon_i$ for $i = 1, \ldots, n$. Denote by $p_{ij}$ the element $p_{\alpha}$ given by (2.13) for the positive root $\alpha = \varepsilon_i - \varepsilon_j$. 
Choosing an appropriate normal ordering on the positive roots, for any \( i = 1, \ldots, n \) we can write the elements (4.5) in the form

\[ p F_{n+1,i} = p_{i,i+1} \cdots p_{in} p_{n'} \cdots p_{i} F_{n+1,i}, \]

(4.6)

where the factor \( p_{i} \) is skipped in the product. Therefore the right denominator of this fraction is

\[ \pi_i = f_{i,i+1} \cdots f_{in} f_{n'} \cdots f_{i}, \]

where

\[ f_{ij} = \begin{cases} F_{ii} - F_{jj} + j - i & \text{if } j = 1, \ldots, n, \\ F_{ii} - F_{jj} + j - i - 2 & \text{if } j = 1', \ldots, n'. \end{cases} \]

Hence, the elements \( s'_{n,i} = p F_{n+1,i} \pi_i \) with \( i = 1, \ldots, n \) belong to the Mickelsson algebra \( S(\mathfrak{o}_{2n+1}, \mathfrak{o}_{2n}) \). One can verify that they are pairwise commuting.

Denote by \( V(\lambda)^+ \) the subspace of \( \mathfrak{o}_{2n} \)-highest vectors in \( V(\lambda) \). Given a \( \mathfrak{o}_{2n} \)-highest weight \( \mu = (\mu_1, \ldots, \mu_n) \) we denote by \( V(\lambda)^+_{\mu} \) the corresponding weight subspace in \( V(\lambda)^+ \):

\[ V(\lambda)^+_{\mu} = \{ \eta \in V(\lambda)^+ \mid F_{ii} \eta = \mu_i \eta, \ i = 1, \ldots, n \}. \]

By the branching rule, the space \( V(\lambda)^+_{\mu} \) is one-dimensional if the condition (4.4) is satisfied. Otherwise, it is zero.

**Theorem 4.1.** Suppose that the inequalities (4.4) hold. Then the space \( V(\lambda)^+_{\mu} \) is spanned by the vector

\[ s'_{n,1}^{\lambda_1-\mu_1} \cdots s'_{n,n}^{\lambda_n-\mu_n} \xi. \]

4.2. **Lowering operators for the reduction** \( \mathfrak{o}_{2n} \downarrow \mathfrak{o}_{2n-1} \)

Taking \( N = 2n \) in the definition (4.1), we shall consider \( \mathfrak{o}_{2n-1} \) as the subalgebra of \( \mathfrak{o}_{2n} \) spanned by the elements (4.1) with \( i, j \neq n, n' \) together with

\[ \frac{1}{\sqrt{2}} (F_{ii} - F_{i'i'}), \quad i = 1, \ldots, n-1, (n-1)', \ldots, 1'. \]

In accordance with the branching rule, the restriction of \( V(\lambda) \) to the subalgebra \( \mathfrak{o}_{2n-1} \) is given by

\[ V(\lambda) |_{\mathfrak{o}_{2n-1}} \simeq \bigoplus \mu V'(\mu), \]
where $V'(\mu)$ is the irreducible finite-dimensional representation of $\mathfrak{o}_{2n-1}$ with the highest weight $\mu$ and the sum is taken over the weights $\mu$ satisfying the inequalities

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \lambda_{n-1} \geq \mu_{n-1} \geq |\lambda_n|,$$

(4.7)

with all the $\mu_i$ being simultaneously integers or half-integers together with the $\lambda_j$.

The elements

$$F_{nn}, \quad F'_{ni} = \frac{1}{\sqrt{2}}(F_{ni} + F_{n'i}), \quad i = 1, \ldots, n-1, (n-1)', \ldots, ',$$

(4.8)

span the $\mathfrak{o}_{2n-1}$-invariant complement to $\mathfrak{o}_{2n-1}$ in $\mathfrak{o}_{2n}$. Therefore, by the general theory of Section 2.2, the Mickelsson–Zhelobenko algebra $Z(\mathfrak{o}_{2n}, \mathfrak{o}_{2n-1})$ is generated by the elements

$$pF_{nn}, \quad p F'_{ni}, \quad i = 1, \ldots, n-1, (n-1)', \ldots, ',$$

(4.9)

where $p$ is the extremal projector for the Lie algebra $\mathfrak{o}_{2n-1}$. Let $\{\varepsilon_1, \ldots, \varepsilon_{n-1}\}$ be the basis of $\mathfrak{h}^*$ dual to the basis $\{F_{11}, \ldots, F_{n-1,n-1}\}$ of the Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{o}_{2n-1}$. Set $\varepsilon_i' = -\varepsilon_i$ for $i = 1, \ldots, n-1$. Denote by $p_{ij}$ and $p_i$ the elements $p_{\alpha}$ given by (2.13) for the positive roots $\alpha = \varepsilon_i - \varepsilon_j$ and $\alpha = \varepsilon_i$, respectively. Choosing an appropriate normal ordering on the positive roots, for any $i = 1, \ldots, n-1$ we can write the elements (4.9) in the form

$$p F'_{ni} = p_{i,i+1}' \cdots p_{i,n-1}' p_i p_{i,(n-1)'} \cdots p_{i,'} F'_{ni},$$

(4.10)

where the factor $p_{i,'}$ is skipped in the product. Therefore the right denominator of this fraction is

$$\pi_i = f_{i,i+1}' \cdots f_{i,n-1}' f_i f_i' f_{i,(n-1)'} \cdots f_{i,'},$$

where

$$f_{ij} = \begin{cases} F_{ij} - F_{jj} + j - i & \text{if } j = 1, \ldots, n-1, \\ F_{ij} - F_{jj} + j - i - 2 & \text{if } j = (n-1)', \ldots, (n-1)' \end{cases}$$

and $f_i = f_i' - 1 = 2(F_{ii} + n - i)$. Hence, the elements $s_{ni} = p F'_{ni} \pi_i$ with $i = 1, \ldots, n-1$ belong to the Mickelsson algebra $S(\mathfrak{o}_{2n}, \mathfrak{o}_{2n-1})$. One can verify that they are pairwise commuting.

Denote by $V(\lambda)^+\mathfrak{o}_{2n-1}$ the subspace of $\mathfrak{o}_{2n-1}$-highest vectors in $V(\lambda)$. Given a $\mathfrak{o}_{2n-1}$-highest weight $\mu = (\mu_1, \ldots, \mu_{n-1})$ we denote by $V(\lambda)_{\mu}^+$ the corresponding weight subspace in $V(\lambda)^+$:

$$V(\lambda)_{\mu}^+ = \{ \eta \in V(\lambda)^+ \mid F_{ii} \eta = \mu_i \eta, \ i = 1, \ldots, n-1 \}.$$  

By the branching rule, the space $V(\lambda)_{\mu}^+$ is one-dimensional if the condition (4.7) is satisfied. Otherwise, it is zero.
THEOREM 4.2. Suppose that the inequalities (4.7) hold. Then the space $V(\lambda)_{\mu}^+$ is spanned by the vector

$$\lambda_{1-\mu_1} \cdots \lambda_{n-\mu_{n-1}} \xi.$$ \(\text{Note that the generator } pF_{nn} \text{ of the algebra } Z(\sigma_{2n}, \sigma_{2n-1}) \text{ does not occur in the formula for the basis vector as it has zero weight with respect to } \mathfrak{h}.\)

4.3. Basis vectors

The representation $V(\lambda)$ of the Lie algebra $\mathfrak{g}_n = \sigma_{2n+1}$ or $\sigma_{2n}$ is equipped with a contravariant inner product which is uniquely determined by the conditions

$$\langle \xi, \xi \rangle = 1 \quad \text{and} \quad \langle F_{ij} u, v \rangle = \langle u, F_{ji} v \rangle$$

for all $u, v \in V(\lambda)$ and any indices $i, j$.

Combining Theorems 4.1 and 4.2 we can construct another basis for each representation $V(\lambda)$ of $\mathfrak{g}_n$; cf. Section 3.2.

B type case. We need to modify the definition of the $B$ type pattern $\Lambda$ introduced in section 3.2. Here $\Lambda$ is an array of the form

$$\begin{array}{cccc}
\lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \\
\lambda'_{n1} & \lambda'_{n2} & \cdots & \lambda'_{nn} \\
\lambda_{n-1,1} & \cdots & \lambda_{n-1,n-1} \\
\lambda'_{n-1,1} & \cdots & \lambda'_{n-1,n-1} \\
& \cdots & \cdots \\
& & \lambda_{11} \\
& & \lambda'_{11}
\end{array}$$

such that $\lambda = (\lambda_{n1}, \ldots, \lambda_{nn})$, the remaining entries are all integers or half-integers together with the $\lambda_i$, and the following inequalities hold

$$\lambda_{k1} \geq \lambda'_{k1} \geq \lambda_{k2} \geq \lambda'_{k2} \geq \cdots \geq \lambda'_{kk} \geq |\lambda'_{kk}|$$

for $k = 1, \ldots, n$, and

$$\lambda'_{k1} \geq \lambda_{k-1,1} \geq \lambda'_{k2} \geq \lambda_{k-1,2} \geq \cdots \geq \lambda'_{k,k-1} \geq |\lambda'_{kk}|$$

for $k = 2, \ldots, n$. 
**Theorem 4.3.** The vectors  
\[ \eta_A = s_{11}^{\lambda_{11}' - \lambda_{11}} \prod_{k=2,\ldots,n} (s_{k1}^{\lambda_{k1}' - \lambda_{k1}} \cdots s_{kk}^{\lambda_{kk}' - \lambda_{kk}})\xi \]  
parametrized by the patterns \( \Lambda \) form an orthogonal basis of the representation \( V(\lambda) \).

**D type case.** Here we define the \( D \) type patterns \( \Lambda \) as arrays of the form  
\[
\begin{array}{cccc}
\lambda_{11} & \lambda_{12} & \cdots & \lambda_{nn} \\
\lambda_{11}' & \cdots & \lambda_{n-1,n-1}' \\
\lambda_{n-1,1} & \cdots & \lambda_{n-1,n-1} \\
\cdots & \cdots & \cdots & \cdots \\
\lambda_{11}' & \cdots & \cdots & \cdots \\
\lambda_{11}
\end{array}
\]  
such that \( \lambda = (\lambda_1,\ldots,\lambda_{nn}) \), the remaining entries are all integers or half-integers together with the \( \lambda_i \), and the following inequalities hold  
\[ \lambda_{k1} \geq \lambda_{k-1,1}' \geq \lambda_{k2} \geq \lambda_{k-1,2}' \geq \cdots \geq \lambda_{k,k-1} \geq \lambda_{k-1,k-1}' \geq |\lambda_{kk}| \]  
for \( k = 2,\ldots,n \), and  
\[ \lambda_{k1}' \geq \lambda_{k1}' \geq \lambda_{k2}' \geq \cdots \geq \lambda_{k,k-1}' \geq \lambda_{kk}' \geq |\lambda_{kk}| \]  
for \( k = 1,\ldots,n-1 \).

**Theorem 4.4.** The vectors  
\[ \eta_A = \prod_{k=1,\ldots,n-1} (s_{k+1,1}^{\lambda_{k+1,1}' - \lambda_{k+1,1}} \cdots s_{k+1,k}^{\lambda_{k+1,k}' - \lambda_{k+1,k}})\xi \]  
parametrized by the patterns \( \Lambda \) form an orthogonal basis of the representation \( V(\lambda) \).

The norms of the basis vectors \( \eta_A \) can be found in an explicit form. The formulas for the matrix elements of the generators of the Lie algebra \( \mathfrak{o}_N \) in the original paper by Gelfand and Tsetlin, [42], are given in the orthonormal basis  
\[ \xi_A = \eta_A/\|\eta_A\|, \quad \|\eta_A\|^2 = \langle \eta_A, \eta_A \rangle. \]
Bibliographical notes

The exposition of this section follows Zhelobenko, [171]. The branching rules were previously derived by him in [167]. The lowering operators for the reduction $\mathfrak{o}_N \downarrow \mathfrak{o}_{N-1}$ were constructed by Pang and Hecht, [133], and Wong, [164]; see also Mickelsson, [93]. They are presented in a form similar to (2.9) and (2.10) although more complicated. A derivation of the matrix element formulas of Gelfand and Tsetlin, [42], was also given in [133] and [164] which basically follows the approach outlined in Section 2.1. The defining relations for the algebra $Z(\mathfrak{o}_N, \mathfrak{o}_{N-1})$ were given in an explicit form by Zhelobenko, [170]. Gould’s approach based upon the characteristic identities of Bracken and Green, [12,54], for the orthogonal Lie algebras is also applicable; see Gould, [46,47,50]. It produces an independent derivation of the matrix element formulas. Although the quantum minor approach has not been developed so far for the Gelfand–Tsetlin basis for the orthogonal Lie algebras, it seems to be plausible that the corresponding analogs of the results outlined in Section 2.5 can be obtained.

Analogos of the Gelfand–Tsetlin bases, [42], for representations of a nonstandard deformation $U'_q(\mathfrak{o}_N)$ of $U(\mathfrak{o}_N)$ were given by Gavrilik and Klimyk, [38], Gavrilik and Iorgov, [37], and Iorgov and Klimyk, [62].

The Gelfand–Tsetlin modules over the orthogonal Lie algebras were studied by Mazorchuk, [89], with the use of the matrix element formulas from [42].

Acknowledgements

It gives me pleasure to thank I.M. Gelfand for his comment on the preliminary version of the chapter. My thanks also extend to V.K. Dobrev, V.M. Futorny, M.D. Gould, M. Harada, M.L. Nazarov, G.I. Olshanski, S.A. Ovsienko, T.D. Palev, V.S. Retakh, and V.N. Tolstoy who sent me remarks and references.

References

Gelfand–Tsetlin bases for classical Lie algebras


[54] H.S. Green, Characteristic identities for generators of $GL(n)$, $O(n)$ and $Sp(n)$, J. Math. Phys. 12 (1971), 2106–2113.


[57] M. Harada, The symplectic geometry of the Gel’fand–Cetlin–Molev basis for representations of Sp(2n, C), Preprint math.SG/0404485.


J.G. Nagel and M. Moshinsky, Operators that lower or raise the irreducible vector spaces of $U_{n-1}$ contained in an irreducible vector space of $U_n$, J. Math. Phys. 6 (1965), 682–694.


Section 4H
Rings and Algebras
with Additional Structure
This page intentionally left blank
Hopf Algebras

Miriam Cohen

Department of Mathematics, Ben Gurion University of the Negev, Beer Sheva, Israel
E-mail: mia@cs.bgu.ac.il

Shlomo Gelaki

Faculty of Mathematics, Technion, Haifa, Israel
E-mail: gelaki@tx.technion.ac.il

Sara Westreich

Interdisciplinary Department of the Social Sciences, Bar-Ilan University, Ramat-Gan, Israel
E-mail: swestric@mail.biu.ac.il

Contents
Introduction ........................................ 175
Part 1. Basic concepts .................................... 177
1.1. Coalgebras and comodules ........................................ 177
1.2. Hopf algebras ........................................ 182
   Definitions and basic examples ....................................... 182
   The finite dual ................................................. 184
1.3. Modules and comodules for Hopf algebras ......................... 185
1.4. Normal Hopf subalgebras, quotients and extensions ................ 187
1.5. Special Hopf algebras ........................................ 188
1.6. Twisting in Hopf algebras ........................................ 191
1.7. Constructing new Hopf algebras from known ones ................ 192
Part 2. Fundamental theorems ......................................... 193
2.1. The fundamental theorem for Hopf modules ....................... 193
2.2. Integrals .................................................... 194
2.3. Maschke’s theorem ........................................ 195
2.4. The antipode .............................................. 196
2.5. The Nichols–Zoeller theorem ................................... 196
2.6. Kac–Zhu theorem ........................................... 197

This research was supported by a grant from the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel.
This research was supported by The Israel Science Foundation founded by the Israel Academy of Sciences and Humanities.

HANDBOOK OF ALGEBRA, VOL. 4
Edited by M. Hazewinkel
© 2006 Published by Elsevier B.V.
Hopf algebras became an object of study from an algebraic standpoint only in the late 1960s. It soon became evident that applications of this theory are abundant in a wide variety of fields. These applications range from topology, knot theory, algebraic geometry, C*-algebras and combinatorics to statistical mechanics, quantum field theory, language theory in computer science, robotics, telecommunications and even chemistry.

The basic idea was developed in the work of H. Hopf, [108], on topological groups. The (co)homology of such groups form what is now termed: graded Hopf algebras. Algebraic properties of such Hopf algebras were first studied in Milnor and Moore’s [163] fundamental work in the mid 60s.

Let us start by introducing an operation termed “comultiplication” which Hopf algebras are endowed with and which are their main novelty. Comultiplication is in a sense going into the “opposite” direction of multiplication. When multiplying, one takes a pair of elements and gets a single element, while when comultiplying one starts with a single element which “opens up” to a sum of pairs.

Explicitly, if \((A, \mu, 1)\) is an algebra over the base field \(k\) then 1 can be considered as a map \(1 : k \to A\) satisfying \(\mu \circ (1 \otimes \text{id}) = \text{id}\) and the multiplication \(\mu : A \otimes A \to A\) satisfies associativity. Dualizing this: A coalgebra \((C, \Delta, \varepsilon)\) has a counit \(\varepsilon\) satisfying \((\varepsilon \otimes \text{id}) \circ \Delta = \text{id}\) and a comultiplication \(\Delta : C \to C \otimes C\) which is coassociative, namely, \((\Delta \otimes \text{id}) \circ \Delta = (\text{id} \otimes \Delta) \circ \Delta\). What coassociativity means is that after “opening up” \(C\) for the first time, one can either “open up” the left or the right tensorands and get the same result.

A bialgebra is an algebra over \(k\) which is also a coalgebra such that \(\Delta\) and \(\varepsilon\) are multiplicative. A Hopf algebra is a bialgebra with an additional special map called the antipode.

The special way in which coalgebras or better yet Hopf algebras arise in the study of the variety of fields is best displayed in combinatorics, where these notions serve as a valuable formal framework. The coproduct displays all ways of decomposing a structure into appropriate parts, while the antipode replaces the role usually played by Möbius inversion.

There are two basic examples of Hopf algebras. The group algebra \(kG\) of a group \(G\), where \(\Delta\) is the diagonal map \(g \mapsto g \otimes g\), \(\varepsilon(g) = 1\) and \(S(g) = g^{-1}\) for all \(g \in G\). The second example is \(U(g)\), the enveloping algebra of a Lie algebra \(g\) where \(\Delta(x) = x \otimes 1 + 1 \otimes x\), \(\varepsilon(x) = 0\) and \(S(x) = -x\) for each \(x \in g\) and extend these maps to \(U(g)\).

A geometrical motivation for the definition of a Hopf algebra is the following. Let \(G\) be a finite group and let \(k\) be any field. Consider the \(k\)-vector space \(H := \text{Fun}(G)\) of all \(k\)-valued functions on \(G\). Then \(H\), equipped with the pointwise multiplication, is a commutative and associative algebra with unit 1 (the function which assigns the value 1 to every \(g \in G\)). Note that the multiplication and unit in \(H\) can be regarded as \(k\)-linear maps \(\mu : H \otimes H \to H\) and \(\eta : k \to H\)

\[
\mu(\alpha \otimes \beta)(g) := \alpha(g)\beta(g) \quad \text{and} \quad \eta(a) := a1.
\]

The multiplication in \(G\) gives rise to a comultiplication map on \(H\):

\[
\Delta : H \to H \otimes H, \quad \Delta(\alpha)(g, h) = \alpha(gh)
\]
(here we identify $\text{Fun}(G \times G)$ with $H \otimes H$, which is allowed since $G$ is finite). Since the multiplication in $G$ is associative, one sees easily that $\Delta$ is coassociative. Note also that $\Delta$ is an algebra homomorphism.

Second, the unit element $e \in G$ gives rise to a counit map on $H$:

$$\varepsilon : H \to k, \quad \varepsilon(\alpha) = \alpha(e).$$

Note that $\varepsilon$ is an algebra homomorphism.

Finally, the inverse operation in $G$ gives rise to an antipode map on $H$:

$$S : H \to H, \quad S(\alpha)(g) = \alpha(g^{-1}).$$

The axioms of the inverse operation in $G$ translate to

$$(S \otimes \text{id}) \circ \Delta = \eta \circ \varepsilon = (\text{id} \otimes S) \circ \Delta.$$

Summarizing, we see that the concept ‘$G$ is a group’ can be expressed in terms of its algebra of functions by saying that this algebra admits three additional structure maps $\Delta$, $\varepsilon$, $S$ making it a commutative Hopf algebra.

More generally, if $G$ is an affine algebraic group, one can similarly define a structure of a (commutative) Hopf algebra on the coordinate ring $H := \text{Fun}(G)$ of all regular functions on $G$. The Hopf algebra $H$ is affine (i.e. finitely generated and commutative with 0 radical). Moreover, if $G$ is an affine variety, then its coordinate ring $H := \text{Fun}(G)$ of all regular functions on $G$ is an affine Hopf algebra if and only if $G$ is an affine algebraic group. This observation was used to develop the structure theory of algebraic groups (e.g., reductive, solvable) using the theory of Hopf algebras.

In the language of schemes introduced by Grothendieck, one may think of affine Hopf algebras as affine group schemes.

It is thus not surprising that much of the work on Hopf algebras in the 70s was inspired by group theory. Kaplansky’s conjectures are good examples of this approach. It was soon found that although quite a few properties can be generalized from group theory to the theory of Hopf algebras, some others are either false in general or hard to translate.

During the beginning of the 80s Hopf algebras have entered in a fundamental way as a unifying tool to analyze the theory of actions of various algebraic structures on algebras. Group actions, group-gradings and actions of Lie algebras are all examples of the theory. However the most striking boost to the theory was given in the beginning of the 80s with the introduction of quantum groups. These are Hopf algebras arising from solutions to the quantum Yang–Baxter equation from statistical mechanics. Another important connection to physics is the realization that standard notions of renormalization theory are derived from the Hopf algebra of rooted trees. This Hopf algebra acts on the Feynman diagrams.

During the 90s there has been a surge of interest in the structure of general Hopf algebras (mainly finite-dimensional) and fundamental examples of quantum groups. Techniques from other areas such as representation theory, algebraic geometry, category theory, Lie theory and ring theory were employed to answer some basic questions. The best results are attained for semisimple or pointed Hopf algebras.
Categorical considerations enter in a fundamental way in the general study of Hopf algebras since one of their striking properties is that they can be characterized by their categories of modules or comodules. Most of the basic Hopf algebraic concepts can be expressed in categorical terms. Furthermore, the theory of finite-dimensional (semisimple) Hopf algebras is one of the main motivations for the study of finite (fusion) categories, which is also motivated by physics (conformal field theory in the semisimple case and logarithmic conformal field theories in the non-semisimple case).

In fact the theory of Hopf algebras also motivated the definition and study of several other central algebraic objects; e.g., quasi-Hopf algebras and weak Hopf algebras.

Part 1. Basic concepts

Loosely speaking, a Hopf algebra is an algebra over a field $k$ also equipped with a “dual” structure, such that the two structures are compatible. In what follows we give precise definitions of the concepts involved.

Throughout, we let $k$ be a field. Vector spaces, algebras and tensor products are assumed to be over $k$ unless stated otherwise. Algebras $A$ are assumed to have a unit $1 = 1_A$ and $u : k \rightarrow A$ is the unit map $\alpha \mapsto \alpha 1_A$ for all $\alpha \in k$.

Our basic references are [1,166,221]. Another reference with more emphasis on the theory of coalgebras is [54].

1.1. Coalgebras and comodules

DEFINITION 1.1.1. A coalgebra over $k$ (or simply a coalgebra) is a vector space $C$ together with two linear maps, comultiplication $\Delta : C \rightarrow C \otimes C$ and counit $\varepsilon : C \rightarrow k$, such that:

1. $\Delta$ is coassociative: $(\Delta \otimes \text{id}) \circ \Delta = (\text{id} \otimes \Delta) \circ \Delta$.
2. $\varepsilon$ satisfies the counit property: $(\varepsilon \otimes \text{id}) \circ \Delta = \text{id} = (\text{id} \otimes \varepsilon) \circ \Delta$.

If $(C, \Delta, \varepsilon)$ is a coalgebra then $C^*$ is an algebra with multiplication given by $\Delta^*$, the dual map of $\Delta$. However, if we begin with an algebra $A$ and try to dualize then difficulties arise for infinite-dimensional $A$. But if $(A, \mu, u)$ is a finite-dimensional algebra then $(A^*, \mu^*, u^*)$ is a coalgebra with $\Delta = \mu^*$ and $\varepsilon = u^*$. Explicitly: $\Delta f(a \otimes b) = f(ab)$ and $\varepsilon(f) = f(1)$ for all $f \in A^*$, $a, b \in A$.

Inductively, one can apply $\Delta$ $n$ times to any one of the tensorands and get the same result. Trying to keep track of the multitude of indices involved in an $n$-fold application of $\Delta$ would be prohibitive. It is in order to simplify this that Heyneman and Sweedler, [106], introduced the extremely successful, so-called, sigma-notation.

SIGMA-NOTATION. For any $c \in C$ write: $\Delta(c) = \sum c_1 \otimes c_2$. The subscripts 1 and 2 are symbols and do not indicate particular elements. When $\Delta$ is applied again to the left tensorand this would symbolically be written as $\sum c_{11} \otimes c_{12} \otimes c_2$, while applying $\Delta$ to the right as: $\sum c_1 \otimes c_{21} \otimes c_{22}$. Coassociativity means that these two expressions are equal.
and hence it makes sense to write this element as \( \sum c_1 \otimes c_2 \otimes c_3 \). Iterating this procedure gives \( \Delta_{n-1}(c) = \sum c_1 \otimes \cdots \otimes c_n \) where \( \Delta_{n-1}(c) \) is the unique element obtained by applying coassociativity \((n-1)\) times. In this notation the counit property says that \( c = \sum \varepsilon(c_1)c_2 = \sum c_1 \varepsilon(c_2) \) for all \( c \in C \).

**EXAMPLE 1.1.2.** Let \( G \) be a group and \( C = kG \), the group algebra. Define \( \Delta, \varepsilon \) by \( \Delta(g) = g \otimes g \) and \( \varepsilon(g) = 1 \), for all \( g \in G \) and extend linearly. It is obvious that \( \Delta \) is coassociative (in fact \( \Delta_{n-1}(g) = g \otimes g \otimes \cdots \otimes g \)) and that \( \varepsilon \) is a counit. When referring to \( kG \) in the sequel we regard it as a coalgebra equipped with these \( \Delta \) and \( \varepsilon \).

This inspired a general terminology for coalgebras:

**DEFINITION 1.1.3.** Let \( (C, \Delta, \varepsilon) \) be a coalgebra. Then \( 0 \neq c \in C \) is called a *grouplike element* if \( \Delta(c) = c \otimes c \). Denote by \( G(C) \) the set of all grouplike elements of \( C \).

**REMARK 1.1.4** [221, p. 55]. \( G(C) \) is a linearly independent set over \( k \).

Let \( V, W \) be vector spaces. Then the *flip* (twist) map \( \tau : V \otimes W \to W \otimes V \) is defined by \( v \otimes w \mapsto w \otimes v \).

**DEFINITION 1.1.5.** Let \( (C, \Delta, \varepsilon) \) be a coalgebra. An element \( c \in C \) is *cocommutative* if \( \Delta(c) = \tau \circ \Delta(c) \). The coalgebra \( C \) is cocommutative if all its elements are cocommutative. (Compare: an algebra \( A \) with multiplication \( \mu \) is commutative if \( \mu = \mu \circ \tau \).)

Observe that the group algebra \( kG \) is cocommutative, while the following are examples of a non-cocommutative coalgebras.

**EXAMPLE 1.1.6.** Let \( C = \text{sp}_k\{1, g, x\} \) with coalgebra structure given by \( \Delta(1) = 1 \otimes 1 \), \( \Delta(g) = g \otimes g \), \( \Delta(x) = x \otimes 1 + g \otimes x \) and \( \varepsilon(1) = 1, \varepsilon(g) = 1, \varepsilon(x) = 0 \).

**EXAMPLE 1.1.7.** Let \( V \) be a finite-dimensional vector space and let \( \text{End}(V) = M_n(k) \) with \( \{e_{ij}\} \) its standard basis. Then the coalgebra structure on \( \text{End}(V)^\ast \) which is dual to the algebra structure of \( \text{End}(V) \) is given explicitly by

\[
\Delta(T^j_i) = \sum_{k=1}^{n} T^k_i \otimes T^j_k \quad \text{and} \quad \varepsilon(T^j_i) = \delta_{ij},
\]

where \( T^j_i \) are the coordinate functions given by \( T^j_i (e_{kl}) = \delta_{ik} \delta_{jl} \).

**DEFINITION 1.1.8.** Let \( (C, \Delta_C, \varepsilon_C) \) and \( (D, \Delta_D, \varepsilon_D) \) be coalgebras. A (linear) map \( f : C \to D \) is a *coalgebra map* if

\[
\Delta_D \circ f = (f \otimes f) \circ \Delta_C \quad \text{and} \quad \varepsilon_C = \varepsilon_D \circ f.
\]
DEFINITION 1.1.9. Let $C \neq 0$ be a coalgebra.

- $C$ is simple if it has no proper subcoalgebras.
- $C$ is irreducible if any two non-zero subcoalgebras of $C$ have a non-zero intersection.

It will be evident from local finiteness, described next, that $C$ is irreducible if and only if it contains a unique simple subcoalgebra. This is the reason for alternatively calling such a coalgebra “colocal”.

- $C$ is indecomposable if $C$ can not be expressed as a non-trivial coalgebra direct sum.

Coalgebras $C$ have the following striking properties.

1. Local finiteness: Every element of $C$ is contained in a finite-dimensional subcoalgebra. Thus, in particular, $C$ contains a simple subcoalgebra and every simple subcoalgebra of $C$ is finite-dimensional.

2. Distributivity: If $C = \bigoplus C_i$, where each $C_i$ is a subcoalgebra of $C$, and $D$ is a subcoalgebra of $C$, then $D = \bigoplus (D \cap C_i)$.

Consequences of the above properties are the following:

THEOREM 1.1.10 [116, 101]. Every coalgebra is uniquely a direct sum of indecomposable subcoalgebras.

Moreover, [116], if $C$ is a cocommutative coalgebra then every indecomposable subcoalgebra is irreducible. Hence $C$ is a unique direct sum of irreducible subcoalgebras. These irreducible subcoalgebras are maximal with respect to the irreducibility property, they are the so-called irreducible components of $C$.

The second part of the theorem above is not necessarily true if $C$ is not cocommutative. For example, the coalgebra defined in Example 1.1.6, is not a sum of irreducible subcoalgebras since its only irreducible components are $k 1$ and $kg$.

REMARK 1.1.11. If $g \in G(C)$, then $kg$ is a simple subcoalgebra of $C$ of dimension 1.

Conversely, any 1-dimensional subcoalgebra of $C$ is of the form $kg$ where $g \in G(C)$.

If $k$ is algebraically closed and $C$ is cocommutative then every simple subcoalgebra is 1-dimensional (this follows easily from considering the dual of $C$).

DEFINITION 1.1.12. Let $C$ be a coalgebra.

- The coradical $C_0$ of $C$ is the sum of all simple subcoalgebras of $C$.
- $C$ is cosemisimple if $C_0 = C$.

- $C$ is pointed if every simple subcoalgebra of $C$ is 1-dimensional (by Remark 1.1.11 this means that $C_0 = kG(C)$).

- $C$ is connected if $C_0$ is one-dimensional.

The coradical $C_0$ induces a filtration on $C$ by the so-called wedge product as follows: for each $n \geq 1$ define inductively

$$ C_n = \Delta^{-1}(C \otimes C_{n-1} + C_0 \otimes C) := C_0 \wedge C_{n-1}. $$

In fact $C_n = (J^n)^\perp$ where $J$ is the Jacobson radical of the algebra $C^*$. The following hold:
THEOREM 1.1.13. \( \{C_n \mid n \geq 0\} \) is a family of subcoalgebras of \( C \) satisfying

1. \( C_n \subseteq C_{n+1} \) and \( C = \bigcup_{n \geq 0} C_n \),
2. \( \Delta(C_n) \subseteq \sum_{i=0}^{n} C_i \otimes C_{n-i} \).

The filtration \( \{C_n\} \) is called the coradical filtration of \( C \).

The following theorem has important implications:

THEOREM 1.1.14 [105]. Let \( C \) and \( D \) be coalgebras and \( f : C \to D \) a coalgebra morphism so that the restriction of \( f \) to \( C_1 \) is injective. Then \( f \) is injective.

DEFINITION 1.1.15. Let \( C \) be a coalgebra with a distinguished group like element \( 1 \) (which will be the unit element in the case \( C \) is the underlying coalgebra of a bialgebra or Hopf algebra). Then \( x \in C \) is called a primitive element of \( C \) if \( \Delta(x) = x \otimes 1 + 1 \otimes x \). The set of primitive elements of \( C \) is denoted by \( P(C) \). More generally, an element \( x \in C \) is called \((\sigma, \tau)\)-primitive if \( \Delta(x) = x \otimes \sigma + \tau \otimes x \) for some \( \sigma, \tau \in G(C) \).

The set of \((\sigma, \tau)\)-primitive elements is denoted by \( P_{\sigma, \tau}(C) \). Such elements are also called skew-primitive elements. By definition, skew-primitive elements of \( C \) are in \( C_1 \).

The dual of the notion of an ideal is that of a coideal.

DEFINITION 1.1.16. A subspace \( I \subseteq C \) is a coideal if

\[ \Delta(I) \subseteq I \otimes C + C \otimes I \quad \text{and} \quad \varepsilon(I) = 0. \]

A subspace \( I \subseteq C \) is a left coideal if

\[ \Delta(I) \subseteq I \otimes C. \]

Right coideals are defined analogously.

When \( I \) is a coideal of \( C \), \( C/I \) is a coalgebra in a natural way.

Just as coalgebras and algebras are dual concepts so are comodules and modules.

DEFINITION 1.1.17. For a coalgebra \((C, \Delta, \varepsilon)\), a (right) \( C \)-comodule is a vector space \( M \) with a \( k \)-linear map \( \rho : M \to M \otimes C \) such that

\[ (\rho \otimes \text{id}) \circ \rho = (\text{id} \otimes \Delta) \circ \rho \quad \text{and} \quad \text{id} = (\text{id} \otimes \varepsilon) \circ \rho. \]

Left \( C \)-comodules are defined analogously.

We sometimes write \((M, \rho)\) or \((M, \rho_M)\) for a \( C \)-comodule. There is also a sigma-notation for right \( C \)-comodules. One writes

\[ \rho(m) = \sum m_0 \otimes m_1 \in M \otimes C. \]

EXAMPLE 1.1.18.

1. Every coalgebra \((C, \Delta, \varepsilon)\) is a right \( C \)-comodule by choosing \( \rho = \Delta \). The right (left) \( C \)-subcomodules of \( C \) are precisely the right (left) coideals.
(2) A somewhat less obvious example is the following: Let $G$ be a group and $M = \bigoplus_{g \in G} M_g$ be a $G$-graded vector space. Then $M$ is a right $kG$-comodule by setting $\rho(m) = m \otimes g$ for each $m \in M_g$.

(3) Recall Example 1.1.7 that if $V = \text{Sp}_k \{v_1, \ldots, v_n\}$ is a finite-dimensional vector space then $(\text{End}(V))^*$ is a coalgebra. Now, $V$ is a right $\text{End}(V)^*$-comodule via $\rho(v_i) = \sum_j v_j \otimes T^j_i$.

Just as for coalgebras we have local-finiteness for comodules $M$. That is, a comodule $M$ contains a finite-dimensional submodule; in particular, $M$ contains a simple submodule and every simple submodule of $M$ is finite-dimensional. Moreover, if $(M, \rho)$ is a simple $C$-comodule then $\rho^{-1}(M \otimes C)$ is a simple subcoalgebra of $C$.

**Definition 1.1.19.** Let $(M, \rho_M)$ and $(N, \rho_N)$ be right $C$-comodules. Then a map $f : M \to N$ is a comodule-map if

$$\rho_N \circ f = (f \otimes \text{id}) \circ \rho_M.$$

**Definition 1.1.20 [163].** Let $C$ be a coalgebra and let $(M, \rho_M)$ and $(N, \rho_N)$ be right and left $C$-comodules respectively. Then the cotensor product $M \Box_C N$ of $M$ and $N$ over $C$ is the equalizer

$$M \otimes N \xrightarrow{id_M \otimes \rho_N} M \otimes C \otimes N \xrightarrow{\rho_M \otimes \text{id}_N} M \otimes C \otimes N.$$

That is

$$M \Box_C N = \{x \in M \otimes N \mid (\text{id}_M \otimes \rho_N)(x) = (\rho_M \otimes \text{id}_N)(x)\}.$$

**Example 1.1.21.** Let $C$ be a coalgebra and $M$ a right $C$-comodule then $M \Box_C C = \rho(M) \cong M$.

Let $V$ be a vector space $M$, $N$ and $C$ as above. Then there exists a canonical isomorphism

$$V \otimes (M \Box_C N) = (V \otimes M) \Box_C N$$

(where $V \otimes M$ is a right $C$-comodule via $\text{id} \otimes \rho_M$).

The dual notion of a bi-module over two algebras is a bi-comodule over two coalgebras.

**Definition 1.1.22.** Let $H, L$ be coalgebras and let $M$ be a left $H$-comodule via $\rho$ and a right $L$-comodule via $\eta$. Then $M$ is an $(H, L)$-bicomodule if

$$(\rho \otimes \text{id}) \circ \eta = (\text{id} \otimes \eta) \circ \rho.$$
DEFINITION 1.1.23. Let \((C, \Delta, \varepsilon)\) be a coalgebra and \(A\) an algebra. Then \(\text{Hom}_k(C, A)\) becomes an algebra under the convolution product \((f \ast g)(c) = \sum f(c_1)g(c_2)\) for all \(f, g \in \text{Hom}_k(C, A)\) and \(c \in C\). The unit element is \(u_\varepsilon\).

In particular, as was already mentioned, \(C^* = \text{Hom}(C, k)\) is an algebra. In fact the convolution product in \(C^*\) is \(\Delta^*\).

We denote the evaluation \(p(c)\) by \(\langle p, c \rangle\) for any \(p \in C^*\) and \(c \in C\).

Observe that if \(I\) is a coideal of \(C\) then \(I^\perp = \{ f \in C^* \mid \langle f, I \rangle = 0 \}\) is a subalgebra of \(C^*\). While if \(J\) is a right coideal of \(C\) then \(J^\perp\) is a right ideal of \(C^*\). If \(D\) is a subcoalgebra of \(C\) then \(D^\perp\) is a two-sided ideal of \(C^*\).

1.2. Hopf algebras

Definitions and basic examples

DEFINITION 1.2.1. Let \(H\) be an algebra with multiplication \(\mu\) and unit map \(u\), and a coalgebra with comultiplication \(\Delta\) and counit \(\varepsilon\). Then \(H\) is a bialgebra if \(\Delta\) and \(\varepsilon\) are algebra maps or equivalently \(\mu, u\) are coalgebra maps.

DEFINITION 1.2.2. Let \(H\) be a bialgebra. Then \(H\) is a Hopf algebra if there exists an element \(S \in \text{End}_k(H)\) so that

\[
\sum S(h_1)h_2 = \varepsilon(h)1_H = \sum h_1S(h_2)
\]

for all \(h \in H\). The map \(S\) is called an antipode for \(H\). It is the inverse of the identity map under convolution.

A map \(f : H \rightarrow H'\) of bialgebras (Hopf algebras) \(H, H'\) is called a bialgebra (Hopf algebra) homomorphism if it is both an algebra and a coalgebra homomorphism (and \(f \circ S_H = S_{H'} \circ f\)).

The kernel of a bialgebra (Hopf algebra) map is a biideal (Hopf ideal). That is, it is both an ideal and a coideal (and stable under \(S\)).

\(H\) is a cocommutative Hopf algebra if it is cocommutative as a coalgebra.

\(H\) is a pointed Hopf algebra if it is pointed as a coalgebra.

\(H\) is a semisimple Hopf algebra if it is semisimple as an algebra.

\(H\) is a cosemisimple Hopf algebra if it is cosemisimple as a coalgebra.

It is easy to see that if \(H\) is either a commutative or a cocommutative Hopf algebra then \(S^2 = \text{id}\).

EXAMPLE 1.2.3.

(1) The group algebra \(kG\) with antipode \(S\) defined by \(S(g) = g^{-1}\) for all \(g \in G\). It is a pointed cosemisimple cocommutative Hopf algebra. If \(G\) is finite then \((kG)^*\) is a commutative semisimple Hopf algebra.
(2) Let $\mathfrak{g}$ be a Lie algebra and $U(\mathfrak{g})$ its enveloping algebra. For each $x \in \mathfrak{g}$ define
\[ \Delta(x) = 1 \otimes x + x \otimes 1 \quad \text{and} \quad S(x) = -x. \]
Then $U(\mathfrak{g})$ is a connected, thus pointed, cocommutative Hopf algebra.

(3) Let $G$ be a group. Let $H = R(G)$ be the Hopf algebra of all real-valued representative functions on $G$ with pointwise multiplication, coproduct given by $\Delta(f)(x, y) = f(xy)$, counit given by $\varepsilon(f) = f(e)$ and the antipode is given by $S(f)(x) = f(x^{-1})$ where $f \in R(G)$, $x, y \in G$ and $e$ is the identity of $G$ (see [1, 2.2] for details).

**Remark 1.2.4.**

(1) If $H$ is a Hopf algebra then $G(H)$ is in fact a group, where $g^{-1} = S(g)$.

(2) A combinatorial calculation implies that if $x$ is a primitive element of $H$ and $k$ is of characteristic zero then the set $\{x^i \mid i > 0\}$ is linearly independent. Hence if $H$ is finite-dimensional over $k$ then $P(H) = 0$.

The following Hopf algebra is the smallest non-commutative, non-cocommutative Hopf algebra. It is again pointed. This is Sweedler’s 4-dimensional Hopf algebra:

**Example 1.2.5.** Let $\text{ch} k \neq 2$ and
\[ H_4 := k[1, g, x, gx \mid g^2 = 1, x^2 = 0, xg = -gx] \]
with coalgebra structure as in Example 1.1.6, $S(g) = g$ and $S(x) = -gx$. Then $H_4$ is a Hopf algebra. Note that $S$ has order 4.

Here are basic properties of the antipode:

**Theorem 1.2.6.** Let $H$ be a Hopf algebra with antipode $S$. Then

1. $S$ is an anti-algebra morphism; that is, $S(hh') = S(h')S(h)$, for all $h, h' \in H$ and $S(1_H) = 1_H$.
2. $S$ is an anti-coalgebra morphism; that is, $\Delta(S(h)) = \sum S(h_2) \otimes S(h_1)$ and $\varepsilon(S(h)) = \varepsilon(h)$, for all $h \in H$.
3. If $H$ is finite-dimensional then $S$ is bijective.

**Remark 1.2.7.** Let $H$ be a Hopf algebra with bijective antipode $S$ and let $H^{\text{cop}} = H$ as an algebra with comultiplication $\Delta^{\text{cop}} := \tau \circ \Delta$ (where $\tau$ is the standard flip map). Then $H^{\text{cop}}$ is a Hopf algebra with antipode $S^{-1}$.

**Integrals.** Integrals for Hopf algebras $H$ are classically defined as certain elements in $H^*$. The name being motivated by the following example:

**Example 1.2.8.** Let $G$ be a compact topological group and $H = R(G)$ as in Example 1.2.3(3). Suppose $\eta$ is a Haar measure on $G$ and set $T(f) = \int_G f \, d\eta$, $f \in H$. Then $T \in H^*$ with an invariance property induced from the left invariance of the Haar measure. Specifically: $xT = \langle x, 1 \rangle T$ for all $x \in H^*$. 

The classical definition of a left integral for $H$ is an element $T \in H^*$ so that for each $x \in H^*$, $xT = \langle x, 1 \rangle T$. If $H$ is finite-dimensional one can thus define an integral for $H^*$ (which is an element of $(H^*)^* = H$). By an abuse of notation it is called a left integral in $H$.

**Definition 1.2.9.** Let $H$ be a finite-dimensional Hopf algebra. A left integral in $H$ is an element $t \in H$ such that $ht = \varepsilon(h)t$ for all $h \in H$. A right integral is an element $t' \in H$ such that $t'h = \varepsilon(h)t'$ for all $h \in H$. The space of left (right) integrals is denoted by $\int_H^l$ ($\int_H^r$) respectively. $H$ is called unimodular if $\int_H^l = \int_H^r$.

Note that $kt$ is a left ideal of $H$ (it is in fact an ideal of $H$ as seen in Theorem 2.2.1).

**Example 1.2.10.** A prime example of an integral is the “averaging element”. Specifically, if $G$ is a finite group and $H = kG$, then $t = \sum_{g \in G} g$ generates the space of left and right integrals in $H$.

**Example 1.2.11.** Let $H = H_4$ of Example 1.2.5. Then $x + gx \in \int_H^l$ and $x - gx \in \int_H^r$. Thus $H$ is not unimodular.

**The finite dual**

One of the important features of Hopf algebras is that its definition is in a sense self-dual. Namely, if $H$ is a finite-dimensional Hopf algebra then its linear dual $H^*$ has a canonical structure of a Hopf algebra with structure maps the transposes of the structure maps of $H$. It is called the dual Hopf algebra of $H$.

When $H$ is an infinite-dimensional Hopf algebra, $H^*$ is an algebra but no longer a Hopf algebra. However $H^*$ contains a Hopf algebra (which is maximal with respect to this property) which is called the finite-dual of $H$.

**Definition 1.2.12.** Let $H$ be a Hopf algebra. Then the finite-dual of $H$ is defined to be

$$H^0 = \{ p \in H^* \mid p \text{ vanishes on an ideal of } H \text{ of finite codimension} \}.$$ 

There are equivalent conditions for $p \in H^*$ to belong to $H^0$. Here is a typical one.

**Proposition 1.2.13** [221, p. 115]. Let $H$ be a Hopf algebra, then

$$H^0 = \{ p \in H^* \mid \dim(H \rightarrow p) < \infty \},$$

where $h \rightarrow p$ is defined by $\langle h \rightarrow p, h' \rangle = \langle p, h'h \rangle$ for all $h, h' \in H$ and $p \in H^*$.

**Theorem 1.2.14** [221, p. 122]. Let $H$ be a Hopf algebra, then $H^0$ is a Hopf algebra with structure maps dual to those in $H$. 
$H^*$ is a topological space with the finite discrete topology on $\text{Hom}(H,k)$, where $k$ has the discrete topology. A subspace $V$ of $H^*$ is dense in $H^*$ if it separates the points of $H$. The finite-dual can be dense in $H^*$, on the other hand it may be trivial.

**Example 1.2.15.**

1. If $H$ is an affine commutative Hopf algebra then $H^0$ is dense in $H$, [221, p. 121].
2. If $\mathfrak{g}$ is a finite-dimensional semisimple Lie algebra and $q$ is not a root of unity then $(U_q(\mathfrak{g}))^0$ is dense in $(U_q(\mathfrak{g}))^*$, [111].
3. If $K$ is an infinite field of cardinality greater than that of $k$ and $G = \text{PSL}_2(K)$, then $(kG)^0 = k\varepsilon$, [24].

An equivalent criterion for density of $H^0$ in $H^*$ is that $H$ is *residually finite-dimensional*. That is, there exists a family $\{\pi_\alpha\}$ of finite-dimensional $k$-representations of $A$ such that $\bigcap_{\alpha} \text{Ker} \pi_\alpha = \{0\}$, [166].

### 1.3. Modules and comodules for Hopf algebras

The representation and co-representation theories for Hopf algebras are particularly rich, as a result of the various structures involved. For any Hopf algebra $H$, there are several ways in which $H$ is acted upon by $H$ or by $H^*$. Here are some:

**Definition 1.3.1.** If $H$ is a Hopf algebra then:

1. $H$ is a left $H^*$-module by

   \[ p \rightarrow h = \sum (p, h_2) h_1 \]

   and a right $H^*$-module by

   \[ h \leftarrow p = \sum (p, h_1) h_2 \]

   for all $h \in H$ and $p \in H^*$.

2. $H$ is a left $H$-module via the left adjoint action $\text{ad}_l$:

   \[ h \text{~}_l \text{ad}_l x = \sum h_1 x S(h_2) \]

   and a right $H$-module via the right adjoint action $\text{ad}_r$:

   \[ x \text{~}_r \text{ad}_r h = \sum S(h_1) x h_2 \]

   for all $h, x \in H$.

Notice that the adjoint actions boil down to the usual adjoint actions of groups and Lie algebras for the Hopf algebras $kG$ and $U(\mathfrak{g})$, respectively.
There is more to say about the actions just described. The adjoint actions make $H$ into an $H$-module algebra while when $H$ is finite-dimensional then $\rightarrow (\leftarrow)$ makes $H$ into a left (right) $H^*$-module algebra, respectively.

**Definition 1.3.2.** Let $A$ be an algebra and $H$ a Hopf algebra. If $A$ is left $H$-module then $A$ is an $H$-module algebra if

$$h \cdot (ab) = \sum (h_1 \cdot a)(h_2 \cdot b)$$

for all $h \in H, a, b \in A$.

Analogously one can define module coalgebras and comodule (co)algebras. In the language of categories these mean that the structure maps of the algebras or the coalgebras are maps in the category of $H$-modules or in the category of $H$-comodules. (For more details see Part 4.)

**Example 1.3.3.** $H$ is a right $H$-comodule coalgebra via the right adjoint coaction: $\rho: H \rightarrow H \otimes H$ given by $h \mapsto \sum h_2 \otimes S(h_1)h_3$.

**Remark 1.3.4.** If $H$ is a finite-dimensional Hopf algebra then $(A, \rho)$ is a right $H$-comodule (algebra) if and only if $(A, \cdot)$ is a left $H^*$-module (algebra), where $\cdot$ and $\rho$ are transposes of each other. That is, if $(A, \rho)$ is a right $H$-comodule (algebra) then for $a \in A$ with $\rho(a) = \sum a_0 \otimes a_1 \in A \otimes H$ define

$$p \cdot a = \sum \langle p, a_1 \rangle a_0$$

for all $p \in H^*$. Conversely, given an action $\cdot$ of $H^*$ on $A$ define for $a \in A$,

$$\rho(a) = \sum (h_i^* \cdot a) \otimes h_i,$$

where $\{h_i\}, \{h_i^*\}$ are dual bases of $H$ and $H^*$.

**Example 1.3.5.** If $G$ is a finite group and $A = \sum_{g \in G} A_g$ is a $G$-graded algebra (that is, $A_g A_h \subset A_{gh}$) then as in Example 1.1.18(2), $A$ is a right $kG$-comodule algebra. By the remark above, $A$ becomes a left $(kG)^*$-module algebra by defining for all $a = \sum a_g \in A$, $p_g \cdot a = a_g$, where $\{p_g\}_{g \in G}$ is the basis of $(kG)^*$ dual to the basis $\{g\}_{g \in G}$ of $kG$.

**Definition 1.3.6.** Let $(M, \cdot)$ be a left $H$-module. Then the $H$-invariants are

$$M^H = \{m \in M \mid h \cdot m = \varepsilon(h)m, \text{ for all } h \in H\}.$$ 

Let $(M, \rho)$ be a right $H$-comodule. Then the $H$-coinvariants are

$$M^{\rho H} = \{m \in M \mid \rho(m) = m \otimes 1\}.$$ 

$^{\rho H}M$ is defined similarly for left $H$-comodules.
EXAMPLE 1.3.7.

(1) [24]. Let $H$ be a left $H$-module by the left adjoint action $\text{ad}$, then $H^H$ is the center of $H$.

(2) Let $H$ be a right $H$-comodule via $\Delta$. Then $H^{\text{co}H} = k$.

REMARK 1.3.8. Just as the averaging map for the theory of group actions so do integrals play a central role in the theory of actions of finite-dimensional Hopf algebras. For if $M$ is a left $H$-module then it is immediate that $t \cdot M \subset M^H$ for $t \in H^H$. If $\varepsilon(t) \neq 0$ then this is actually an equality (for then, $t \cdot m = \varepsilon(t)m$ for $m \in M^H$ implies that $m = t \cdot (\frac{1}{\varepsilon(t)}m)$).

1.4. Normal Hopf subalgebras, quotients and extensions

A basic concept in group theory is that of a normal subgroup, which is a subgroup stable under the adjoint action of the group on itself. It is characterized by the property that every right coset is a left coset as well. Equivalently, a normal subgroup is a kernel of a group homomorphism, and thus the quotient group is defined.

The Hopf algebra analogue of normality is given by:

DEFINITION 1.4.1. Let $K$ be a Hopf subalgebra of $H$. Then $K$ is a normal in $H$ if

$$(\text{ad}_{lH})(K) \subseteq K \quad \text{and} \quad (\text{ad}_{rH})(K) \subseteq K,$$

where $\text{ad}_{lH}, \text{ad}_{rH}$ are the left and right adjoint actions of $H$ on itself (Definition 1.3.1(2)).

It is straightforward to verify that if $K$ is a normal Hopf subalgebra of $H$ then $HK^+ = K + H$, where $K^+ = \{ h \in K \mid \varepsilon(h) = 0 \}$. Since $HK^+$ is a Hopf ideal of $H$ it follows that $H/K^+$ is a Hopf algebra (while $H/K$ is usually meaningless). Let $\overline{H} := H/K^+$ then we have the following exact sequence of Hopf algebras:

$$K \hookrightarrow H \twoheadrightarrow \overline{H}$$

and $H$ is called a Hopf extension (of $\overline{H}$ by $K$).

However, in the converse direction, if $\pi : H \rightarrow \overline{H}$ is a Hopf algebra epimorphism then $\text{Ker} \pi$ is not necessarily of the form $K^+ H$ for some normal Hopf subalgebra $K$. Furthermore, even if $K$ is a Hopf subalgebra of $H$ such that $HK^+ = K^+ H$ (and thus $H/K^+$ is a Hopf algebra) it is not always true that $K$ is a normal Hopf subalgebra of $H$. To discuss this (following [206,207]) we introduce the following notion:

Given any Hopf algebra epimorphism $\pi : H \rightarrow \overline{H}$, $H$ becomes both a left and a right $\overline{H}$-comodule via

$$\rho_l = (\pi \otimes \text{id}) \circ \Delta, \quad \rho_r = (\text{id} \otimes \pi) \circ \Delta.$$

Denote by $H^{\text{co}\overline{H}}$ and $^{\text{co}\overline{H}}H$ the algebras of coinvariants for those coactions (as defined in Definition 1.3.6).
If $K$ satisfies $HK^+ = K^+ H$ and we let $\overline{H} = H / K^+ H$, then clearly $K \subset H^{\text{co}\overline{H}}$ and $K \subset \text{co}\overline{H} H$.

**Proposition 1.4.2** [207]. If $H$ is faithfully flat over $K$ then

$$K = H^{\text{co}\overline{H}} = \text{co}\overline{H} H.$$  

In this case $K$ is indeed a normal Hopf subalgebra of $H$.

Examples for which $H$ is faithfully flat over any Hopf subalgebra $K$ are:

1. The Hopf algebra $H$ is finite-dimensional, [178] (see also Remark 3.2.8).
2. If $H$ is commutative, [59].
3. If $H_0$, the coradical of $H$, is cocommutative, [224]. In particular if $H$ is cocommutative or pointed.

Here is an example of a non-commutative non-cocommutative Hopf algebra which is an extension of a commutative and cocommutative Hopf algebra (in fact the unique non-commutative non-cocommutative semisimple Hopf algebra of dimension 8).

**Example 1.4.3** [115, Example 4.1]. Let $k$ be of characteristic 0 and let

$$H_8 := k\langle x, y, z \mid x^2 = y^2 = 1, xy = yx, zx = yz, xz = zy, z^2 = \frac{1}{2}(1 + x + y - xy) \rangle$$

with a coalgebra structure given by

$$\Delta(x) = x \otimes x, \quad \Delta(y) = y \otimes y,$$

$$\Delta(z) = \frac{1}{2}\left((1 + y) \otimes 1 + (1 - y) \otimes x\right)(z \otimes z).$$

Then $K := k(x, y) \cong k(\mathbb{Z}_2 \times \mathbb{Z}_2)$ is a normal Hopf subalgebra of $H$ and $\overline{H} = H / HK^+$ is isomorphic to $k\mathbb{Z}_2$. So $H$ is the Hopf extension

$$K \cong k(\mathbb{Z}_2 \times \mathbb{Z}_2) \hookrightarrow H \twoheadrightarrow k(\bar{z}) \cong k(\mathbb{Z}_2),$$

where $\pi(x) = \pi(y) = 1$, $\pi(z) = \bar{z}$ and $\Delta(\bar{z}) = \bar{z} \otimes \bar{z}$.

**1.5. Special Hopf algebras**

A generalization of cocommutative Hopf algebras are quasitriangular Hopf algebras, introduced by Drinfeld. The dual notion is that of coquasitriangular Hopf algebras (sometimes called braided Hopf algebras). Though quasitriangular Hopf algebras were introduced in the context of solutions of the quantum Yang–Baxter equation, they play an important role in the general theory of Hopf algebras.
DEFINITION 1.5.1 [72]. A quasitriangular Hopf algebra is a pair \((H, R)\), where \(H\) is a Hopf algebra and 
\[ R = \sum R^1 \otimes R^2 \in H \otimes H \]
is invertible, such that the following hold:

\begin{align*}
\text{(QT1)} \quad (\Delta \otimes \text{id})(R) &= R^{13} R^{23} = \sum R^1 \otimes r^1 \otimes R^2 r^2, \\
\text{(QT2)} \quad (\text{id} \otimes \Delta)(R) &= R^{13} R^{12} = \sum R^1 r^1 \otimes r^2 \otimes R^2, \\
\text{(QT3)} \quad (\tau \circ \Delta)(h) &= R_{\Delta(h)} R^{-1} \quad \text{for all } h \in H,
\end{align*}

where \( r = R \) and \( R^{13} = \sum R^1 \otimes 1 \otimes R^2 \in H^\otimes 3 \) etc.

\( R \) is sometimes called a universal \( R \)-matrix.

A consequence of the above is that 
\[ R^{-1} = \sum S(R^1) \otimes R^2, \sum \varepsilon(R^1) R^2 = \sum R^1 \varepsilon(R^2) = 1 \] and \((S \otimes S)(R) = R\).

If \( R^{-1} = R^t := \sum R^2 \otimes R^1 \) then \((H, R)\) is called a triangular Hopf algebra and \( R \) is called unitary.

It is property (QT3) which generalizes cocommutativity. In fact, Drinfeld termed Hopf algebras satisfying this property almost cocommutative.

EXAMPLE 1.5.2. The following are basic examples of quasitriangular Hopf algebras:

1. Every cocommutative Hopf algebra is triangular with \( R = 1 \otimes 1 \).

2. Let \( k \) be of characteristic \( \not= 2 \), then the group algebra \( kZ_2 = k \langle 1, g \rangle \) is triangular with
\[ R = \frac{1}{2}(1 \otimes 1 + 1 \otimes g + g \otimes 1 - g \otimes g). \]

3. [192]. Let \( H = H_4 \) be as in Example 1.2.5. Then \( H \) is quasitriangular with a family of quasitriangular structures \( R_\alpha, \alpha \in k \), given by
\[ R_\alpha = \frac{1}{2}(1 \otimes 1 + 1 \otimes g + g \otimes 1 - g \otimes g) \]
\[ + \frac{\alpha}{2}(x \otimes x + x \otimes gx - gx \otimes x + xg \otimes xg). \]

An important element in \((H, R)\) is the so-called Drinfeld element
\[ u = \sum S(R^2) R^1. \]

It is shown in [72] that \( u \) is an invertible element in \( H \), \( \Delta(u) = (u \otimes u)(R^t R)^{-1} \) and \( S^2 \) is an inner automorphism of \( H \) induced by \( u \); that is, \( S^2(h) = uhu^{-1} \), all \( h \in H \). Moreover, \( S^4 \) is induced by the grouplike element \( uS(u)^{-1} \).

Another important ingredient of a quasitriangular Hopf algebra \((H, R)\) is the Drinfeld map \( f : H^* \to H \) given by
\[ f(p) = (\text{id} \otimes p) R^t R. \]

Usually, it is not an algebra map, however its restriction to
\[ O(H^*) := \{ p \in H^* \mid \langle p, hh' \rangle = \langle p, S^2(h')h \rangle \forall h, h' \in H \} \]
is an algebra map to the center of $H$. If $H$ is semisimple and $k$ is algebraically closed of characteristic 0 then $O(H^*)$ coincides with the character ring of $H$ (see Section 5.1 for definition and properties) hence $O(H^*)$ is called the algebra of generalized characters.

**Definition 1.5.3.** A finite-dimensional quasitriangular Hopf algebra for which the Drinfeld map $f$ is injective (and thus bijective) is called *factorizable*.

It was proved in [194] that every factorizable Hopf algebra is unimodular.

We have:

**Theorem 1.5.4.** Let $(H, R)$ be a finite-dimensional quasitriangular Hopf algebra. Then the following are equivalent:

1. $H$ is factorizable.
2. $f(T') \neq 0$ where $T' \neq 0$ is a right integral for $H^*$.
3. $f$ restricted to $O(H^*)$ is injective.

**Example 1.5.5.** The Drinfeld double, $D(H)$, is defined for any finite-dimensional Hopf algebra $H$. It is a factorizable Hopf algebra which contains $H$ (see Section 4.3 for details).

When $(H, R)$ is finite-dimensional, $R \in H \otimes H \cong (H^* \otimes H^*)^*$, hence $R$ defines a bilinear form (or an $R$-form) $\langle \mid \rangle_R$ on $H^*$. The properties of this form define the axioms for coquasitriangular Hopf algebras. Thus if $(H, R)$ is a finite-dimensional Hopf algebra then $(H^*, \langle \mid \rangle_R)$ is a coquasitriangular Hopf algebra. Coquasitriangular Hopf algebras have been studied by a number of people (cf. [142, 136, 155, 202]).

**Definition 1.5.6.** A coquasitriangular Hopf algebra is a pair $(H, \langle \mid \rangle)$ where $H$ is a Hopf algebra over $k$ and $\langle \mid \rangle : H \otimes H \to k$ is a $k$-linear form (braiding) which is convolution invertible in $\text{Hom}_k(H \otimes H, k)$ such that the following hold for all $h, g, l \in H$:

(CQT1) $\langle h | gl \rangle = \sum \langle h_1 | g \rangle \langle h_2 | l \rangle$.

(CQT2) $\langle hg | l \rangle = \sum \langle g | l_1 \rangle \langle h | l_2 \rangle$.

(CQT3) $\sum \langle h_1 | g_1 \rangle g_2 h_2 = \sum h_1 g_1 \langle h_2 | g_2 \rangle$.

If $\sum \langle h_1 | g_1 \rangle \langle g_2 | h_2 \rangle = \varepsilon(g)\varepsilon(h)$ then $(H, \langle \mid \rangle)$ is called a cotriangular Hopf algebra.

A non-trivial example of a cotriangular Hopf algebra is $kG$ where $G$ is an Abelian group with a symmetric bicharacter $\langle \mid \rangle$. This Hopf algebra is commutative and cocommutative and it arises in the context of Lie color algebras (cf. [14]).

A special class of quasitriangular Hopf algebras is that of ribbon Hopf algebras.

**Definition 1.5.7.** A finite-dimensional ribbon Hopf algebra over $k$ is a triple $(H, R, v)$, where $(H, R)$ is a finite-dimensional quasitriangular Hopf algebra over $k$ and $v \in H$ satisfies the following axioms:

(R1) $v$ is in the center of $H$.

(R2) $S(v) = v$, and

(R3) $\Delta(v) = (v \otimes v)(R^{\tau}R)^{-1} = (R^{\tau}R)^{-1}(v \otimes v)$. 

It follows from these axioms that $\varepsilon(v) = 1$ and $v^2 = uS(u)$. Also, observe that $G := u^{-1}v$ is a grouplike element of $H$. It is called the special grouplike element of $H$. Ribbon Hopf algebras were introduced and studied by Reshetikhin and Turaev in connection with invariants of links and 3-manifolds, [198]. Here $u$ is the Drinfeld element.

**REMARK 1.5.8.** Any triangular Hopf algebra is ribbon with 1 as the ribbon element and $u^{-1}$ as the special grouplike element. Also, any semisimple quasitriangular Hopf algebra such that $S^2 = \text{id}$ (e.g., if the characteristic of $k$ is zero) is ribbon with $u$ as the ribbon element and 1 as the special grouplike element.

### 1.6. Twisting in Hopf algebras

Let $H$ be a Hopf algebra. It is possible to twist either the comultiplication or the multiplication of $H$ and thereby construct a new Hopf algebra. The following fundamental definition is due to Drinfeld, [71].

**DEFINITION 1.6.1.** A dual Hopf 2-cocycle ($=\tau$ twist) for $H$ is an invertible element $J \in H \otimes H$ which satisfies:

$$(\Delta \otimes \text{id})(J)(J \otimes 1) = (\text{id} \otimes \Delta)(J)(1 \otimes J),$$

$$(\varepsilon \otimes \text{id})(J) = (\text{id} \otimes \varepsilon)(J) = 1.$$

Given a twist $J$ for $H$, one can twist the comultiplication and define a new Hopf algebra structure $(H^J, m, 1, \Delta^J, S^J)$ on the algebra $(H, m, 1)$. The coproduct and antipode are determined by

$$\Delta^J(a) = J^{-1} \Delta(a) J, \quad S^J(a) = Q^{-1} S(a) Q$$

for every $a \in H$, where $Q := m \circ (S \otimes \text{id})(J)$.

Suppose $(H, R)$ is also (quasi)triangular. Then it is straightforward to verify that $(H^J, R^J)$ is quasi(triangular) with universal $R$-matrix $R^J := (J^\tau)^{-1} R J$ (where $J^\tau = \tau(J)$).

The tensor category of left (right) $H$-modules is equivalent to that of left (right) $H^J$-modules.

Dually, one can twist the multiplication on $H$.

**DEFINITION 1.6.2.** A linear form $\sigma : H \otimes H \to k$ is called a Hopf 2-cocycle for $H$ (see [63]) if it has an inverse $\sigma^{-1}$ under the convolution product $\ast$ in $\text{Hom}_k(H \otimes H, k)$, and satisfies the cocycle condition:

$$\sum \sigma(a_1 b_1, c) \sigma(a_2, b_2) = \sum \sigma(a, b_1 c_1) \sigma(b_2, c_2),$$

$$\sigma(a, 1) = \varepsilon(a) = \sigma(1, a)$$

for all $a, b, c \in H$. 
Observe that any 2-cocycle $\sigma$ on a group $G$ can be extended to a Hopf 2-cocycle for $kG$. Given a Hopf 2-cocycle $\sigma$ for $H$, one can construct a new Hopf algebra structure $(H^\sigma, m^\sigma, 1, \Delta, \varepsilon, S^\sigma)$ on the coalgebra $(H, \Delta, \varepsilon)$. The new multiplication is given by

$$m^\sigma(a \otimes b) = \sum \sigma^{-1}(a_1, b_1)a_2b_2\sigma(a_3, b_3)$$

for all $a, b \in H$. The new antipode is given by

$$S^\sigma(a) = \sum \sigma^{-1}(a_1, S(a_2))S(a_3)\sigma(S(a_4), a_5)$$

for all $a \in H$.

Suppose $H$ is also co(quasi)triangular with universal $R$-form $\langle | \rangle : H \otimes H \to k$. Then $H^\sigma$ is co(quasi)triangular with universal $R$-form $(\sigma \circ \tau)^{-1} * (|) * \sigma$.

Observe that a twist $J$ on $H$ yields a Hopf 2-cocycle $\sigma_J$ on $H^*$ by identifying $H \otimes H$ with a subalgebra of $(H^* \otimes H^*)^*$. If $H$ is finite-dimensional then the existence of a twist $J$ for $H$ is equivalent to the existence of a Hopf 2-cocycle $\sigma_J$ for $H^*$.

**Example 1.6.3.**

(1) If $(H, R)$ is quasitriangular then $R$ is a twist for $(H^\text{cop}, R^t)$ and $((H^\text{cop})^R, (R^t)^R) = (H, R)$. Dually, if $(H, \langle | \rangle)$ is coquasitriangular then $\langle | \rangle \circ \tau$ is a Hopf 2-cocycle for $H$.

(2) Let $G$ be a non-Abelian finite group, $H = kG$ and $J$ a twist for $H$. Since every cocommutative Hopf algebra is triangular with $R = 1 \otimes 1$ we have that $(kG)^J$ is a triangular Hopf algebra with non-trivial triangular structure $R^J := (J^t)^{-1}J$.

A dual Hopf 2-cocycle is a special case of what is known as a (dual) pseudo-cocycle. A (dual) pseudo-cocycle for $H$ is an invertible element $J \in H \otimes H$ satisfying necessary and sufficient conditions that make $\Delta J = J^{-1} \Delta J$ coassociative.

**Example 1.6.4 [180].** The Hopf algebra $H_8$ in Example 1.4.3 can be obtained from the group algebra of either $D_4$ or $Q$ by pseudo-twists (where $D_4$ is the dihedral group and $Q$ is the quaternion group). However $H_8$ has no non-trivial Hopf 2-cocycle or dual Hopf 2-cocycle, [229].

### 1.7. Constructing new Hopf algebras from known ones

Techniques in developing the structure theory for Hopf algebras entail constructing new Hopf algebras from known ones by using several methods. Some of these methods have been described already. We list them briefly.

(1) For two Hopf algebras $H_1$ and $H_2$ one can always define $H_1 \otimes H_2$, the tensor Hopf algebra with tensor structure maps.

(2) Given a Hopf algebra $H$ one can twist either the multiplication via a 2-cocycle or the comultiplication via a dual 2-cocycle and obtain another Hopf algebra $H_\sigma$ or $H^J$ (Section 1.6).
(3) Given Hopf algebras $H$ and $K$ one can sometimes define an extension $K^\#_\sigma H$ of $H$ by $K$ by using a 2-cocycle $\sigma$ and a dual 2-cocycle $\tau$ (Remark 3.2.8).

(4) Another possible product of two Hopf algebras $H_1$ and $H_2$ satisfying certain conditions is the double crossed product $H_1 \triangleleft \rhd H_2$. In particular $D(H) = H^{\text{cop}} \rhd H$, the Drinfeld double of $H$, is defined for any finite-dimensional Hopf algebra $H$ (Definition 4.3.8).

(5) In certain cases the smash product $A \# H$ can be turned into a Hopf algebra in a process called biproduct or bosonization (Theorem 4.4.6).

Part 2. Fundamental theorems

2.1. The fundamental theorem for Hopf modules

Definition 2.1.1. Let $H$ be a Hopf algebra. Then $(M, \cdot, \rho)$ is a right $H$-Hopf module if

(HM1) $(M, \cdot)$ is a right $H$-module.

(HM2) $(M, \rho)$ is a right $H$-comodule.

(HM3) The compatibility condition

$$\rho(m \cdot h) = \sum m_0 \cdot h_1 \otimes m_1 h_2$$

for all $m \in M$, $h \in H$, is satisfied.

Example 2.1.2. Let $V$ be a vector space and $M = V \otimes H$. $(M, \cdot)$ is a right $H$-module by $(m \otimes x) \cdot h = m \otimes xh$. $(M, \rho)$ is a right $H$-comodule by $\rho(m \otimes x) = \sum m \otimes x_1 \otimes x_2$ for all $m \in M$, $x \in H$. These make $M$ into a right $H$-Hopf module. It is called a trivial Hopf module.

The fundamental theorem says that all Hopf modules are trivial with $V = M^{\text{co}H}$. Explicitly:

Theorem 2.1.3 [135]. Let $M$ be a right $H$-Hopf module. Then $M \cong M^{\text{co}H} \otimes H$ as right $H$-Hopf modules, where $M^{\text{co}H} \otimes H$ is a trivial Hopf module.

The isomorphism $M \rightarrow M^{\text{co}H} \otimes H$ is given by:

$$m \mapsto \sum m_0 \cdot (S(m_1)) \otimes m_2.$$ 

The crucial steps are to prove that $\sum m_0 \cdot (S(m_1)) \in M^{\text{co}H}$ for all $m \in M$ and that this map is an $H$-module isomorphism.

This theorem is essential in the proof about the uniqueness of the integral and in the Nichols–Zoeller theorem (see Sections 2.2 and 2.5).
2.2. Integrals

The following important theorem about integrals is due to Larson and Sweedler:

**Theorem 2.2.1** [135]. Let $H$ be a finite-dimensional Hopf algebra. Then

1. $\int_H^l$ and $\int_H^r$ are one-dimensional ideals (with generators $0 \neq t$ and $0 \neq t'$ respectively).
2. $H^* \rightarrow t = H = t' \leftarrow H^*$. That is, $H$ is a free left (right) $H^*$-module of rank 1.

The major step in the proof of the theorem is to show that $H^*$ is a right $H$-Hopf module. Now, $H^*$ is a left $H^*$-module via left multiplication. The transpose of this action makes $H^*$ into a right $H$-comodule as in Remark 1.3.4. Moreover, $H^*$ is also a right $H$-module via $p \leftarrow h = \sum (S(h), p_2) p_1$ for all $h \in H$, $p \in H^*$. Once it is proved that this action and coaction satisfy the compatibility condition for Hopf modules (HM3), the fundamental Theorem 2.1.3 implies that $H^* \cong (H^*)^{\text{co}H} \otimes H$. However, $(H^*)^{\text{co}H} = \int_H^l$. Thus since $\dim H^* = \dim H$, this implies that $\int_H^l$ is one-dimensional.

It is quite surprising that the existence of $0 \neq t \in H$ so that $xt = \epsilon(x)t$ for all $x \in H$ implies that $H$ is finite-dimensional. This follows from the following theorem and its corollary.

**Theorem 2.2.2** [220]. Let $H$ be any Hopf algebra and $0 \neq I$ a right ideal in $H$. Then $H^* \rightarrow I = H$.

**Corollary 2.2.3.** If a Hopf algebra $H$ contains a non-zero finite-dimensional left (right) ideal then $H$ is finite-dimensional.

As we have seen in Example 1.2.8, when the Hopf algebra $H$ is infinite-dimensional there may still exist $0 \neq T \in H^*$ so that $xT = \langle x, 1 \rangle T$ for all $x \in H^*$. (Recall that $H^*$ is no longer a Hopf algebra.) It is straightforward to see that

$$xT = \langle x, 1 \rangle T \iff \langle T, h \rangle = \sum \langle T, h_2 \rangle h_1$$

for all $h \in H$, which is equivalent to $T$ being a left $H$-comodule map.

Here again the dimension of the space of such $T$ is 1, [219], which means uniqueness of the integral. Moreover,

**Theorem 2.2.4** [220]. Let $H$ be a Hopf algebra, then the following are equivalent:

1. $H$ is cosemisimple.
2. There exists $T \in H^*$ so that $\langle T, 1 \rangle \neq 0$ and $xT = \langle x, 1 \rangle T$ for all $x \in H^*$.

Some quantized coordinate algebras of algebraic groups $G$ are known to be cosemisimple (e.g., $O_q(\text{SL}_n)$ where $q$ is not a root of unity, [104]).

Back to finite-dimensional Hopf algebras $H$. The space $\int_H^l$ is a 2-sided 1-dimensional ideal of $H$. Hence for any $h \in H$, $th = \langle \alpha, h \rangle t$, where $\langle \alpha, h \rangle \in k$. Since the map $\alpha$ is an
algebra map it follows that $\alpha \in G(H^*)$. This element $\alpha$ is called the left distinguished grouplike or left modular element of $H$ (of course $H$ is unimodular if and only if $\alpha = \varepsilon$). The right modular element of $H$ is defined analogously and equals $\alpha^{-1}$.

Another consequence of Theorem 2.2.1 is that any finite-dimensional Hopf algebra is a Frobenius algebra.

**Theorem 2.2.5** [183]. Let $H$ be a finite-dimensional Hopf algebra, $T \in \int_l^H$ and $t' \in \int_r^H$ such that $\langle T, t' \rangle = 1$. Then $T$ is a Frobenius homomorphism with dual bases $(S(t'_1), t'_2)$ (that is, for all $h \in H$, $h = \sum S(t'_1)(T, t'_2 h)$).

An important application is the use of the integral in computations of traces ($\text{Tr}$) of linear endomorphisms.

**Theorem 2.2.6** [191]. Let $H$ be a finite-dimensional Hopf algebra, $t \in \int_l^H$ and $T' \in \int_r^H$ so that $\langle T', t \rangle = 1$. Then $\text{Tr}(f) = \sum \langle T', S(t_2) \rangle f(t_1)$ for all $f \in \text{End}_k(H)$.

As a corollary we have

**Theorem 2.2.7.** Let $H, T'$ and $t$ be as in Theorem 2.2.6, then

$$\text{Tr}(S^2) = \langle \varepsilon, t \rangle \langle T', 1 \rangle.$$ 

### 2.3. Maschke’s theorem

A classical result about finite groups $G$ is Maschke’s theorem: $kG$ is a semisimple algebra if and only if $|G|^{-1} \in k$. In Hopf algebraic terms: Let $t := \sum_{g \in G} g$, then $t \in \int_k^H$ and $\varepsilon(t) = |G|$. Thus $|G|^{-1} \in k$ if and only if $\varepsilon(t) \neq 0$ in $k$. Hence $kG$ is a semisimple algebra if and only if $\varepsilon(t) \neq 0$. Inspired by this, Larson and Sweedler showed the last statement to be true for any finite-dimensional Hopf algebra.

**Theorem 2.3.1** [135]. Let $H$ be any finite-dimensional Hopf algebra. Then $H$ is semisimple if and only if $\varepsilon(\int_l^H) \neq 0$ (if and only if $\varepsilon(\int_r^H) \neq 0$).

One direction is easily proved. Assume $H$ is a semisimple algebra. Since $\ker \varepsilon$ is an ideal of $H$, there exists a left ideal $0 \neq I$ of $H$ so that $H = I \oplus \ker \varepsilon$. Then it is shown directly that $I \subset \int_l^H$ and hence $\varepsilon(\int_l^H) \neq 0$.

For the converse choose $t \in \int_l^H$ so that $\varepsilon(t) = 1$. Let $M$ be any left $H$-module and $N$ be an $H$-submodule of $M$. The essence of the proof is to use the integral to produce an $H$-complement of $N$ from a mere $k$-complement of $N$. Let $\pi : M \to N$ be any $k$-linear projection and define $\bar{\pi} : M \to N$ by $\bar{\pi}(m) = \sum t_1 \cdot \pi(S(t_2) \cdot m)$ for all $m \in M$. Then $\bar{\pi}$ is an $H$-projection of $M$ onto $N$.

**Remark 2.3.2.** If $H$ is semisimple then $H$ is unimodular (since if $t \in \int_l^H$ with $\varepsilon(t) = 1$ and $t' \in \int_r^H$ with $\varepsilon(t') = 1$, then $t = t't = t'$).
2.4. The antipode

Let $H$ be a finite-dimensional Hopf algebra with antipode $S$. The antipode, intertwined with the integral and the modular elements play a central role in the theory, [131,193].

**Theorem 2.4.1** [189]. Let $\alpha \in G(H^*)$ be the left modular element of $H$ and $g \in G(H)$ be the right modular element of $H^*$. Then for all $h \in H$

$$S^4(h) = g(\alpha \to h \leftarrow \alpha^{-1})g^{-1} = \alpha \to (ghg^{-1}) \leftarrow \alpha^{-1}.$$ 

A simplified proof of this theorem appears in [209]. It depends on treating $H$ as a Frobenius algebra as in Theorem 2.2.5. It can also be derived from the trace formula (Theorem 2.2.6). Another way of proving this formula appears in [72] via the Drinfeld double.

Since $H$ is finite-dimensional and since grouplike elements are linearly independent, there exist only a finite number of powers of the grouplike elements $a$ and $\alpha$. Thus, a corollary of the previous theorem is:

**Theorem 2.4.2** [189]. Let $H$ be a finite-dimensional Hopf algebra, then $S$ has finite order (necessarily even).

It is worth mentioning that for any $n$ there exists a Hopf algebra with an antipode of order $2n$; they are called the Taft algebras $H_n$.

**Example 2.4.3** [222]. Let $n \geq 1$ and $\omega \in k$ be a primitive $n$-th root of unity. Then $H_{n,\omega}$ is a Hopf algebra defined as follows. As an algebra $H_{n,\omega}$ is generated by $g, x$ subject to the relations

$$g^n = 1, \quad x^n = 0, \quad \text{and} \quad xg = \omega gx.$$ 

The coalgebra structure of $H_{n,\omega}$ is determined by

$$\Delta(g) = g \otimes g \quad \text{and} \quad \Delta(x) = x \otimes g + 1 \otimes x.$$ 

Observe that $\dim H_{n,\omega} = n^2$; indeed $\{g^i x^j\}_{0 \leq i, j < n}$ is a linear basis for $H_{n,\omega}$. The antipode of $H_{n,\omega}$ is determined by $S(g) = g^{-1}$ and $S(x) = -xg^{-1}$. Hence $S^2$ is the algebra automorphism of $H_{n,\omega}$ determined by $S^2(g) = g$ and $S^2(x) = gxg^{-1} = \omega^{-1}x$ and so $S^{2n} = \text{id}$.

2.5. The Nichols–Zoeller theorem

One of the fundamental results in the theory of finite-dimensional Hopf algebras is the Nichols–Zoeller theorem which is a positive answer to one of Kaplansky’s conjectures, [116]. This conjecture was inspired by the famous “Lagrange theorem” in group theory, and boils down to it for $H = kG$. 

THEOREM 2.5.1 [178]. Let $H$ be a finite-dimensional Hopf algebra and $K$ a Hopf subalgebra of $H$. Then $H$ is free as a right $K$-module. In particular, $\dim K$ divides $\dim H$.

The theorem is actually more general. It says that every $M \in \mathcal{M}^H_K$ is free as a right $K$-module. Here $M \in \mathcal{M}^H_K$ if $M$ satisfies the conditions in Definition 2.1.1 with a modified (HM1) in which $K$ replaces $H$.

The essential part of the proof is based on the fact that any finite-dimensional Hopf algebra is a Frobenius algebra, on the fundamental theorem for Hopf modules and on the Krull–Schmidt theorem. By using these it is shown that if $W$ is a finitely generated right $K$-module and $V$ is a finitely generated faithful right $K$-module so that $W \otimes V \cong W^{\dim V}$ as right $K$-modules, then $W$ is free over $K$. Next it is proved that $M \otimes H \cong M^{\dim H}$ for any $M \in \mathcal{M}^H_K$. Both imply that $M$ is free over $K$.

A simple corollary of this theorem is

COROLLARY 2.5.2. Let $H$ be a finite-dimensional Hopf algebra over any field. Then the order of $G(H)$ divides $\dim(H)$.

Another corollary is

COROLLARY 2.5.3. Let $H$ and $K$ be as in Theorem 2.5.1 and assume $H$ is semisimple, then so is $K$.

2.6. Kac–Zhu theorem

The following theorem was conjectured in [116] and proved by Zhu using the Nichols–Zoeller theorem and extending work of G.I. Kac, [114], who worked in the framework of $C^*$-algebras.

THEOREM 2.6.1 [114,250]. Let $H$ be a Hopf algebra over an algebraically closed field $k$ of characteristic 0. If $\dim(H) = p$ is prime then $H$ is isomorphic to the group algebra $k\mathbb{Z}_p$.

If $G(H)$ or $G(H^*)$ is non-trivial then by the Nichols–Zoeller theorem $H = k\mathbb{Z}_p$ and we are done. Otherwise, $S^4 = \text{id}$ by the formula for $S^4$ (Theorem 2.4.1). If $p$ is odd then it is easy to see that $\text{Tr}(S^2) \neq 0$ which implies by Theorem 2.2.7 that $H$ is semisimple. If $p = 2$ then semisimplicity can be proved directly. Now, the character theory of semisimple Hopf algebras and the appropriate “class equation” described in Theorem 5.1.6 imply that $H = k\mathbb{Z}_p$.

Part 3. Actions and coactions

3.1. Smash products, crossed products and invariants

One of the important topics in ring theory during the 70s was, so-called, non-commutative Galois theory. Specifically, if $A$ is an algebra, $G$ is a group of automorphisms of $A$ and $A^G$
is the subalgebra of $G$-invariants, then the study concerned connections between the ideal structure of $A$ and $A^G$. Much of the information is encoded in a generalized semidirect product, the skew group algebra $A * G$, [165], and the connection $A^G \subset A \subset A * G$.

Another setup with a similar flavor is: $A$ is an algebra, $L$ is a Lie algebra of derivations of $A$ and $A^L = \{a \in A \mid l(a) = 0 \text{ for all } l \in L\}$. The analogue of $A * G$ is not so obvious.

A third setup of the same nature is that of group-graded algebras, [48,176], where $G$ is a finite group. The analogue of $A^G$ and $A^L$ is $A_1$ (where 1 is the identity element of $G$), but the analogue of $A * G$ is even less obvious.

It turns out that all three setups are unified by the fact that all the algebras $A$ are $H$-module algebras for appropriate $H$, [47,21], and once this is understood there exists a generalization of the semidirect product which is $A * G$ for $H = kG$. This is the smash product $A \# H$. It plays a central role in the theory as did $A * G$ for non-commutative Galois theory.

**DEFINITION 3.1.1.** Let $H$ be a Hopf algebra and let $A$ be a left $H$-module algebra. Then the smash product algebra $A \# H$ is defined as follows:

1. As vector spaces, $A \# H = A \otimes H$. Write $a \# h$ for $a \otimes h$.
2. Multiplication is given by:

\[(a \# h)(b \# g) = \sum a(h_1 \cdot b)h_2k\]

for all $a, b \in A, h, g \in H$.

The smash product $A \# H$ is an algebra which contains $A$ via $a \mapsto a \# 1$, for $a \in A$, and $H$ via $h \mapsto 1 \# h$, for $h \in H$.

A generalization of $A \# H$ is, as for group actions, the notion of a crossed product. It is not necessary for the algebra $A$ to be an $H$-module algebra, but only an $H$-measured algebra. That is, there exists a linear map $H \otimes A \rightarrow A$ given by $h \otimes a \mapsto h \cdot a$, such that $h \cdot 1 = \varepsilon(h)1$ and $h \cdot (ab) = \sum (h_1 \cdot a)(h_2 \cdot b)$ for all $a, b \in A, h \in H$.

**DEFINITION 3.1.2.** Let $H$ be a Hopf algebra and $A$ an $H$-measured algebra. Assume that $\sigma$ is an invertible map (under convolution) in $\text{Hom}_k(H \otimes H, A)$. The crossed product $A \#_{\sigma} H$ is $A \otimes H$ as a vector space. We write $a \# h$ for element in $A \otimes H$. Multiplication is given by:

\[(a \# h)(b \# g) = \sum a(h_1 \cdot b)\sigma(h_2, g_1) \# h_3g_2\]

for all $a, b \in A, h, g \in H$.

**THEOREM 3.1.3** [67,24]. $A \#_{\sigma} H$ is an associative algebra with identity element $1 \# 1$ if and only if:

1. $A$ is a twisted $H$-module; that is, $1 \cdot a = a$ and

\[h \cdot (g \cdot a) = \sum \sigma(h_1, g_1)(h_2g_2 \cdot a)\sigma^{-1}(h_3, g_3)\]

for all $h, g \in H, a \in A$. 

(2) \( \sigma \) is a 2-cocycle; that is, \( \sigma(h, 1) = \sigma(1, h) = \epsilon(h)1 \) and

\[
\sum h_1 \cdot \sigma(g_1, m_1) \sigma(h_2, g_2 m_2) = \sum \sigma(h_1, g_1) \sigma(h_2 g_2, m_2)
\]

for all \( h, g, m \in H \).

REMARK 3.1.4. Observe that for the special case of \( A = k \) (which is an \( H \)-module algebra via \( h \cdot \alpha = \epsilon(h) \alpha \) for all \( h \in H, \alpha \in k \)) the definition of a 2-cocycle coincides with Hopf 2-cocycle (see Definition 1.6.2). In this case we can construct the crossed product \( k\#_\sigma H \) which is denoted also by \( \sigma H \) and is called a twisted Hopf algebra (reminiscent of twisted group rings; it is not a Hopf algebra though).

One can similarly define crossed product for right actions and can obtain a right twist of \( H \). The twist for \( H \) defined in Section 1.6 has then the form \( \sigma_{H_{\alpha^{-1}}} \) (which is a Hopf algebra).

Let \( A \) be an \( H \)-module algebra. By Remark 1.3.8 if \( t \in \int^H \) then \( t \cdot A \subset A^H \) and equality holds when \( H \) is semisimple. \( A^H \) is connected to \( A\#H \) in several ways:

[20]. Let \( H \) be finite-dimensional. Then \( A^H \cong (\text{End}_{A\#H}(A))^\text{op} \) as algebra (\( \text{op} = \text{opposite algebra} \)).

[44]. Let \( H \) be finite-dimensional and assume \( t \cdot A = A^H \). Then there exists an idempotent \( e \in A\#H \) such that \( e(A\#H)e = AHe \cong A^H \) as algebras.

The close relationship between \( A^H \) and \( A\#H \) can be expressed in terms of a Morita context.

THEOREM 3.1.5 [43,44]. Let \( H \) be a finite-dimensional Hopf algebra, \( 0 \neq t \in \int^H \) and \( \alpha \) the distinguished grouplike element associated with it. Consider \( A \) as a left (right) \( A^H \)-module via left (right) multiplication, as a left \( A\#H \)-module via \( (a\#h) \cdot b = a(h \cdot b) \) and as a right \( A\#H \)-module via \( b \cdot (a\#h) = [S^{-1}(\alpha \rightarrow h)] \cdot (ba) \), for all \( a, b \in A, h \in H \). Then \( M := A^H A A\#H \) and \( N := A\#H A A^H \) together with the maps:

\[
[\cdot, \cdot] : A^H \otimes A \rightarrow A\#H \quad \text{given by} \quad [a, b] = (a\#t)(b\#1),
\]

\[
(\cdot, \cdot) : A\#H \otimes A \rightarrow A^H \quad \text{given by} \quad (a, b) = t \cdot (ab)
\]

give a Morita context for \( A^H \) and \( A\#H \).

This extends earlier work on group actions by [37]. Note that \( (A, A) = t \cdot A \).

COROLLARY 3.1.6. If both \( t \cdot A = A^H \) and \( (A\#t)(A\#1) = A\#H \) then \( A^H \) is Morita-equivalent to \( A\#H \).

The surjectivity of the Morita map \( [\cdot, \cdot] \) has strong implications as will be seen in the next section.

Using ring theoretic properties deduced from the theory of Morita contexts we have:
**COROLLARY 3.1.7** [44]. The following are equivalent:

1. $A \# H$ is a prime ring.
2. $A$ is a left and right faithful $A \# H$-module and $A$ is a prime ring.
3. If $A$ is left and right faithful $A \# H$-module then $A^H$ is a primitive ring if and only if $A \# H$ is a primitive ring.

As for group actions, semiprimeness of $A \# H$ insures that the ideal structures of $A$ and $A^H$ are closely related. If $A$ is semiprime it is known that $A \# kG$ is semiprime if $|G|^{-1} \in k$, [94], in this case $kG$ is semisimple. This suggests the still open generalized Maschke-type theorems:

**QUESTION 3.1.8.**

1. [43]. If $H$ is semisimple and $A$ is a semiprime $H$-module algebra, is $A \# H$ semiprime?
2. [25]. More generally: under the conditions of (1), is $A \# \sigma H$ semiprime for any 2-cocycle $\sigma$?
3. [42]. In a similar spirit, let $B \rtimes H$ be a double crossproduct of $B$ and $H$ ([153]). If $B$ and $H$ are semisimple Hopf algebras is $B \rtimes H$ semisimple?

The answer to the first and second questions is positive in the following cases: (a) If $A$ is also Artinian, [43,25]. (b) If $A$ is any $k$-affine PI algebra and the characteristic of $k = 0$ or with a suitable assumption on the PI degree if the characteristic of $k > 0$, [138]. (c) If the action of $H$ on $A$ is inner (see Definition 3.4.5), [24,26]. (d) If $H$ is commutative (for then $H$ is essentially $(kG)^*$ and the answer is positive in this situation by [47]). (e) If $H$ is pointed cocommutative, [39,172]. Question (3) has a positive answer when the field $k$ is algebraically closed of characteristic zero. It also has a positive answer over arbitrary fields for the Drinfeld double, namely, for $B = H^{\text{cop}}$, [192].

A dual notion of the surjectivity of the form $(,)$ in the Morita context is better suited for infinite-dimensional Hopf algebras $H$ and right $H$-comodule algebras $A$. This is the notion of a total integral, [61].

**DEFINITION 3.1.9** [61]. Let $A$ be a right $H$-comodule algebra. Then a (right) total integral for $A$ is a right $H$-comodule map $\Phi : H \to A$ such that $\Phi(1) = 1$.

**REMARK 3.1.10** [43,61]. When $H$ is finite-dimensional then surjectivity of the Morita map $(,)$ on an $H$-module algebra $A$ is equivalent to the existence of a total integral for $A$ considered as a right $H^*$-comodule algebra.

**REMARK 3.1.11.** Note that the existence of a total integral for $k$ is equivalent to $H$ being cosemisimple by Theorem 2.2.4.

The existence of a total integral is also related to Galois extensions (see next section) by the following:
Remark 3.1.12 [61]. If $A^\text{co} H \subset A$ is a faithfully flat $H$-Galois extension then there exists a total integral for $A$.

3.2. Galois extensions

Crossed products are examples of Hopf Galois extensions. The definition of these extensions has its roots in the Chase, Harrison and Rosenberg, [37], approach to Galois theory for groups acting on commutative rings. The definition is given in terms of coactions.

Definition 3.2.1. Let $(A, \rho)$ be a right $H$-comodule algebra. Then the extension $A^\text{co} H \subset A$ is right $H$-Galois if the map

$$\beta : A \otimes_{A^\text{co} H} A \to A \otimes H \quad \text{given by} \quad a \otimes b \mapsto (a \otimes 1) \rho(b)$$

is bijective.

Example 3.2.2.

1. Classical Galois field extensions are examples of Hopf–Galois extensions, [165].
2. Let a Hopf algebra $H$ be a right $H$-comodule algebra via $\Delta$. Then $H^\text{co} H = k$ and $k \subset H$ is right $H$-Galois with $\beta^{-1} : H \otimes H \to H \otimes H \otimes H$ given by $x \otimes y \mapsto \sum x(S(y_1)) \otimes y_2$.
3. Let $H$ be a finite-dimensional Hopf algebra, $K$ a normal Hopf subalgebra and $H = H/K + K$. Then $K \subset H$ is $H$-Galois. This is true since $H^\text{co} H = K$ by Proposition 1.4.2, and the Galois map $\beta : H \otimes_K H \to H \otimes H$ has an inverse defined similarly to $\beta^{-1}$ of part (2).
4. Let $B = A \#_H H$ be any crossed product. Then $B$ is an $H$-comodule algebra via $\rho = \text{id}_A \otimes \Delta$ and $B^\text{co} H = A \#_H k \cong A$. The fact that $A \subset B$ is Galois follows since

$$(A \#_H H) \otimes_A (A \#_H H) \cong (A \#_H H) \otimes H$$

and the Galois map has the form $\text{id}_A \otimes \beta$ with an inverse $\text{id}_A \otimes \beta^{-1} : (A \#_H H) \otimes H \to (A \#_H H) \otimes H$ where $\beta, \beta^{-1}$ are defined as in (2) above.

Recall Example 1.1.18(2). If $A = \sum_{g \in G} A_g$ then $A$ is a $kG$-comodule algebra by $x \mapsto x \otimes g$ for all $x \in A_g$. In this case:

Theorem 3.2.3 [234]. $A_1 \subset A$ is $kG$-Galois if and only if $A_g A_h = A_{gh}$ for all $g, h \in G$ ($A$ is then called strongly graded).

Of special interest is the case where $A^\text{co} H = k$.

Definition 3.2.4. If the extension $k \subset A$ is $H$-Galois then $A$ is called an $H$-Galois object (sometimes the extension is called an $H$-Galois extension).

There also exists a two-sided analogue which plays an important role in the representation theory of $H$. 

Hopf algebras
DEFINITION 3.2.5. Let $H$, $L$ be Hopf algebras and let $A$ be an $(H, L)$-bicomodule algebra. Then $A$ is called an $(H, L)$-bi-Galois object if $A$ is both a left $H$-Galois object and a right $L$-Galois object.

When $A^{\text{co}H} \subset A$ is right $H$-Galois then $C_A(A^{\text{co}H})$, the centralizer of $A^{\text{co}H}$ in $A$, becomes a right $H$-module algebra via the so-called Miyashita–Ulbrich action.

DEFINITION 3.2.6 [235,68]. Let $A^{\text{co}H} \subset A$ be a right $H$-Galois extension with Galois map $\beta$. Define the right Miyashita–Ulbrich action as follows:

For $h \in H$, write $\beta^{-1}(1 \otimes h) = \sum a_i \otimes b_i$ and then define

$$x \leftarrow h = \sum a_i x b_i,$$

for all $x \in C_A(A^{\text{co}H})$.

THEOREM 3.2.7. The above $\leftarrow$ defines a right action of $H$ on $C_A(A^{\text{co}H})$. The algebra of invariants of this action is $Z(A)$, the center of $A$.

Back to crossed products. While all crossed products are Galois extensions, in order for a Galois extension $A^{\text{co}H} \subset A$ to become a crossed product $A^{\text{co}H} \#_\sigma H$, the extension must have the normal basis property, that is

$$A \cong A^{\text{co}H} \otimes H$$

as a left $A^{\text{co}H}$-module and a right $H$-comodule. This is the essence of the work done by [129,67,25].

REMARK 3.2.8. Normal Hopf subalgebras $K$ of finite-dimensional $H$ (Example 3.2.2(3)) are not only Galois but $K \subset H$ satisfies the normal basis property, [207,66], hence $H = K \#_\sigma \overline{H}$ as algebras.

In fact, $H = K \#_\tau \overline{H}$ as Hopf algebras where $\tau$ is a dual cocycle which twists the coalgebra structure of $K \otimes H$.

Progress in the theory of such extensions has been made in the case that $K = (kG)^*$ and $\overline{H} = kG'$, where $G$ and $G'$ are finite groups. These extensions can be described in terms of actions, coactions, a cocycle and a dual cocycle relating the groups $G$ and $G'$ and satisfying certain compatibility conditions. Such groups were considered in [225] and were named matched pairs. The subject was further discussed in [146,148,5,3].

Other instances in which Galois extensions $A^{\text{co}H} \subset A$ are crossed products $A \cong A^{\text{co}H} \#_\sigma H$ are:

THEOREM 3.2.9.

(1) [128]. If $H$ is a finite-dimensional Hopf algebra and $A$ is an $H$-Galois object then $A$ is a crossed product over $k$.

(2) [196]. If $H$ is a finite-dimensional Hopf algebra such that $A$ and $A \otimes H^*$ satisfy Krull–Schmidt for projectives and $A_A^{\text{co}H}$ is free then $A$ is a crossed product over $A^{\text{co}H}$. 
As a corollary, if $H$ is a finite-dimensional Hopf algebra over an algebraically closed field $k$, any $H$-Galois extension $A^{co}H \subset A$ with $A^{co}H$ a local ring is a crossed product over $A^{co}H$.

(3) [19]. If $H$ and $A$ satisfy the equivalent conditions of the following Theorem 3.2.12 and $H$ is connected (as a coalgebra) then $A$ is a crossed product over $A^{co}H$.

When $H$ is finite-dimensional then a left $H$-action on $A$ gives rise to a right $H^*$-coaction on $A$ so that $A^H = A^{co}H^*$. Thus it makes sense to ask when is the extension $A^H \subset A$ right $H^*$-Galois. Here are some equivalences:

**THEOREM 3.2.10** [44,129]. Let $H$ be a finite-dimensional Hopf algebra and $A$ a left $H$-module algebra. Then the following are equivalent:

1. $A^H \subset A$ is right $H^*$ Galois.
2. The Morita map $[,]$ of Theorem 3.1.5 is surjective.
3. For any $M \in \mathcal{A}^H \text{Mod}$, $M \cong A \otimes A^H M^H$ as an $A^H$-module (via $a \otimes m \mapsto a \cdot m$).
4. $A$ is a generator for $\mathcal{A}^H \text{Mod}$.

When $A = D$ is a division algebra more can be said.

**THEOREM 3.2.11** [44]. If $A = D$ is a division algebra then the equivalent conditions of Theorem 3.2.10 are also equivalent to each of the following:

1. The right (left) dimension of $D$ over $D^H$ equals $\dim H$.
2. $D^H$ is a simple ring.
3. $D$ is a faithful left or right $D^H$-module.
4. $D \cong D^H \# H^*$.

An extension of Theorem 3.2.10, which also generalizes a theorem for algebraic groups on induction of modules and affine quotients is:

**THEOREM 3.2.12** [206]. Let $H$ be a Hopf algebra with bijective antipode and $A$ a right $H$-comodule algebra. Then the following are equivalent:

1. $A^{co}H \subset A$ is a right $H$-Galois extension and $A$ is a faithfully flat left (or right) $A^{co}H$-module.
2. The Galois map $\beta$ is surjective and $A$ is an injective $H$-comodule.
3. The map $\text{Mod}_{A^{co}H^*} \rightarrow \mathcal{M}_A^H$ given by $N \mapsto N \otimes A^{co}H$ is an equivalence (where $\mathcal{M}_A^H$ is the subcategory of $A$-modules in the category of $H$-comodules, [60]).

The ideal structures of $A^{co}H$ and $A$ are closely related for some Galois extensions.

**THEOREM 3.2.13** [172]. Let $A$ and $H$ be as in Theorem 3.2.12 and assume $A^{co}H \subset A$ is a right $H$-Galois extension and $A$ is a faithfully flat $A^{co}H$-module. Then

1. There exists a bijection between the following sets of ideals:
   \[
   \{ I \subset A^{co}H \text{ satisfying } IA = AI \} \overset{\Psi}{=} \{ J \subset A \text{ which are } H \text{-subcomodules} \},
   \]
   \[\Phi : I \mapsto IA \quad \text{and} \quad \Psi : J \mapsto J \cap A^{co}H.\]
If $H$ is also finite-dimensional then there is a bijective correspondence between $H$-equivalent primes of $A^{\text{co} H}$ and $H^*$-equivalent primes of $A$ (the equivalence relation is quite natural; see [172] for the exact definition).

If $A^{\text{co} H} \subseteq Z(A)$, more can be said.

**Theorem 3.2.14.** Let $H$ be a finite-dimensional Hopf algebra and $A$ an $H$-module algebra. If $A/A^H$ is $H^*$-Galois and $A^H \subseteq Z(A)$ then:

1. By [13,129,61], $A$ is a faithfully flat right $A^{\text{co} H^*}$-module and the Morita map $(,)$ is surjective. In particular $A^H$ and $A^{\# H}$ are Morita equivalent, [44].
2. [49]. For every ideal $I$ of $A^{\# H}$

\[ I = (I \cap A)^{\# H} = ((I \cap A^H)A)^{\# H}.\]

Consequently, if $A$ is an $H$-Galois object then $A^{\# H}$ is a simple ring.

3. [172]. If $H^*$ is pointed and $G = G(H^*)$, then there exists a bijection

\[ \text{Spec}(A)/G \rightarrow \text{Spec}(A^{\text{co} H^*}) \]

given by $[P] \mapsto P \cap A^{\text{co} H^*}$ (where $\text{Spec}(A)/G$ is the set of $G$-orbits $[P]$ in $\text{Spec}(A)$).

### 3.3. Duality theorems

Let $A$ be a $G$-graded algebra, where $G$ is a finite group; then $A$ is a left $(kG)^*$-module algebra (Example 1.3.5) and $A^{\# (kG)^*}$ is a left $kG$-module where $kG$ acts trivially on $A$ and by $\rightarrow$ (Definition 1.3.1) on $kG$. The first “duality” theorem was proved in this setup and considerably generalized independently by [25] and [239].

**Theorem 3.3.1 [47].** Let $G$ be a finite group and $A$ be a $G$-graded algebra. Then

\[(A^{\# (kG)^*})^{\# kG} \cong M_n(A),\]

where $M_n(A)$ is the algebra of $n \times n$ matrices over $A$.

The following theorem deals with a general $H$. Again $H^*$ acts on $A^{\# H}$ as for the $G$-graded case.

**Theorem 3.3.2 [25,239].** Let $H$ be a finite-dimensional Hopf algebra and $A$ an $H$-module algebra. Then

\[(A^{\# H})^{\# H^*} \cong M_n(A).\]

This theorem is in fact a corollary of the most general result:
THEOREM 3.3.3 [25]. Let $H$ be a Hopf algebra and $U$ a Hopf subalgebra of $H^0$ such that both $H$ and $U$ have bijective antipodes, and assume that $U$ satisfies the RL-condition with respect to $H$ (see [25] for the definition). Let $A$ be a $U$-comodule algebra and define an action of $H$ on $A$ by: $h \cdot a = \sum (a_1, h) a_0$. Then

$$(A\#H)\#U \cong A \otimes (H\#U).$$

In particular:

COROLLARY 3.3.4 [25]. Specifying Theorem 3.3.3 to a residually finite-dimensional Hopf algebra $H$ and $U$ a dense Hopf subalgebra of $H^0$, then

$$(A\#H)\#U \cong A \otimes L,$$

where $L$ is a dense subring of $\text{End}_k(H)$.

The duality theorem has been reproved using other methods. For example, by using the right smash product of a comodule algebra with a Hopf algebra, [16]. It has been generalized by using other constructions. For example, by using “opposite smash products” for right $H$-comodule algebras, [127].

There also exist duality theorems for crossed coproducts of Hopf algebras coacting (weakly) on coalgebras, [55].

The duality theorem has its origin in operator algebra theory for actions and coactions of locally compact groups on von Neumann algebras, [175], and Kac algebras, [75]. It has been generalized to weak Kac algebras, [181].

3.4. Analogues of two theorems of E. Noether; inner and outer actions

A classical theorem of E. Noether on invariants states that if $A$ is a commutative $k$-affine algebra and $G$ is a finite group of automorphisms on $A$ then $A^G$ is $k$-affine. This theorem has the following generalizations. Part (1) of the generalization is a consequence of a result of Grothendieck, [59, p. 309]. A more explicit proof which uses determinants is due to [92].

THEOREM 3.4.1 [92]. Let $H$ be a finite-dimensional cocommutative Hopf algebra and let $A$ be a commutative $H$-module algebra. Then

1. $A$ is integral over $A^H$.
2. If $A$ is $k$-affine then so is $A^H$.

This was generalized as follows:

THEOREM 3.4.2 [51]. Let $(H, R)$ be a triangular semisimple Hopf algebra in characteristic 0. Let $A$ be an $H$-commutative $H$-module algebra (see Definition 4.3.3), then:

1. $A$ is integral over $A^H$. 


(2) \( A \) is a PI algebra.

(3) If \( A \) is \( k \)-affine then so is \( A^H \).

Another proof of the above theorem follows from [77] who proved that \( H \) in Theorem 3.4.2 is a twisting of a group algebra \( kG \), and from [170] who showed various algebra properties invariant under twisting.

A result similar in flavor is:

**Theorem 3.4.3** [166]. Let \( H \) be a finite-dimensional Hopf algebra and \( A \) a left Noetherian \( H \)-module algebra so that the Morita map \( (, ) \) is surjective. If \( A \) is \( k \)-affine then so is \( A^H \).

The following is an infinite-dimensional Noether-type theorem for coactions.

**Theorem 3.4.4** [74]. Let \( A = A_0 \oplus A_1 \oplus \cdots \) be a right Noetherian \( \mathbb{N} \)-graded algebra with \( A_0 = k \). Let \( H \) be a cosemisimple Hopf algebra and suppose \( A \) is a right \( H \)-comodule so that each \( A_i \) is a subcomodule of \( A \). Then the subalgebra \( A^H \) is \( k \)-affine.

The classical Noether–Skolem theorem asserts that if \( A \) is a simple Artinian ring with center \( Z \) and \( B \supseteq Z \) is a simple subalgebra of \( A \) with \( Z \)-finite dimension then any isomorphism of \( B \) into \( A \) extends to an inner automorphism of \( A \). This can be generalized to Hopf algebras as in the next theorem. First a definition.

**Definition 3.4.5.** Let \( C \) be a coalgebra and \( B \subset A \) be algebras. Consider a left action \( C \otimes B \to A \), given by \( c \otimes b \mapsto c \cdot b \) which measures \( B \) to \( A \) (that is \( x \cdot (ab) = \sum (x_1 \cdot a)(x_2 \cdot b) \) and \( x \cdot 1 = \varepsilon(x)1 \) for all \( x \in C, a, b \in B \)). Then the measuring is inner if there exists a convolution invertible map \( u \in \text{Hom}(C, A) \) such that for all \( x \in C, b \in B \),

\[
x \cdot b = \sum u(x_1)bu^{-1}(x_2).
\]

This definition boils down to the usual definition of inner automorphisms and inner derivations in the appropriate settings. For if \( \sigma \) is an automorphism of \( A \) and there exists \( a \in A \) such that \( \sigma \cdot x = axa^{-1} \) for all \( x \in A \) then define \( u(\sigma) = a \) and extend the definition to \( C := k(\sigma) \). Similarly, if a derivation \( \delta \) satisfies \( \delta(x) = ax - xa \) for some \( a \in A \) all \( x \in A \), then define \( u(\delta) = a, u(1) = 1 \) and extend it to \( C := k(1, \delta) \).

**Example 3.4.6.** The left adjoint action of \( H \) on \( A \) is inner (with \( u = \text{id} \) and \( u^{-1} = S \)).

**Theorem 3.4.7** [126]. Let \( A \) be a simple Artinian algebra with center \( k \) and let \( B \) be a finite-dimensional simple subalgebra of \( A \). Let \( C \) be a coalgebra which measures \( B \) to \( A \). Assume also that \( B^{op} \otimes D^* \) is a simple algebra for each simple subcoalgebra \( D \) of \( C \). Then the measuring is inner.

Important applications of this theorem are:
COROLLARY 3.4.8. Let $A$ be a simple Artinian algebra with center $k$ and $B$ be a finite-dimensional simple subalgebra of $A$. Let $C$ be a pointed coalgebra. Then any measuring of $B$ to $A$ by $C$ is inner.

The following was proved independently in [144].

COROLLARY 3.4.9. Let $A$ be a simple algebra which is finite-dimensional over its center. Let $C$ be a coalgebra. Then any measuring of $A$ by $C$ is inner.

If $H$ is a Hopf algebra and $A$ is an $H$-module algebra we say that $H$ is inner on $A$ if the measuring of $A$ by $H$ is inner.

If $G$ is a group of automorphisms of $A$ then the set $N$ of all $g \in G$ which are inner automorphisms is a normal subgroup of $G$ and $kN$ is the maximal sub-Hopf algebra of $kG$ which is inner on $A$. Moreover, $G/N$ acts on $A^N$. This can be generalized as follows:

THEOREM 3.4.10. Let $H$ be a finite-dimensional pointed Hopf algebra and $A$ an $H$-module algebra. Then

1. [145]. There exists a unique sub-Hopf algebra $H_{\text{inn}}$ of $H$ which is inner on $A$ and is maximal with respect to this property.

2. [208]. If $H$ is also cocommutative then $H_{\text{inn}}$ is a normal sub-Hopf algebra and $H = H/H(H_{\text{inn}})^+$ acts on $A^H_{\text{inn}}$.

When a group $G$ acts by automorphisms on a prime ring $A$ then some $g \in G$ may fail to be inner on $A$, but extending the action of $G$ to $Q$, the symmetric Martindale ring of fractions of $A$, it may be inner on $Q$. When this happens $g$ is called $X$-inner (where $X$ stands for Kharchenko, see [123]). Extending these ideas to Hopf algebra actions on algebras, [41], constructed $H$-quotient rings on which the action may be inner.

The best results are obtained for pointed Hopf algebras $H$; it was proved that the action of $H$ can be extended to $Q$ as for group actions, [167,171]. In this case one can define $X$-inner actions to be actions which become inner when extended to $Q$.

While for actions of groups by automorphisms (or of restricted Lie algebras by derivations) we define $X$-outer actions as those for which the only inner automorphism (derivation) is trivial, it is rather unsatisfactory to extend this definition for general Hopf algebras.

The following definition for outer actions of pointed Hopf algebras is due to [161]. Let $H$ be a pointed Hopf algebra acting on a prime algebra $A$, let $Q$ be the symmetric Martindale ring of quotients of $A$ and let $K$ be the center of $Q$. Set $E := C_{Q\#H}(A)$, the centralizer of $A$ in $Q\#H$.

DEFINITION 3.4.11. Let $H$ be a pointed Hopf algebra acting on a prime algebra $A$. Then the action of $H$ on $A$ is $X$-outer if $E = K$.

Previous results obtained in [124] for $X$-outer actions of groups and restricted Lie algebras can be generalized for $X$-outer actions of pointed Hopf algebras. We list some of them below.
THEOREM 3.4.12 [161]. Let $H$ be a finite-dimensional pointed Hopf algebra acting on a prime algebra $A$. If the action of $H$ on $A$ is $X$-outer, then:

1. $Q\#H$ is a prime algebra, and if $Q$ is $H$-simple then $Q\#H$ is a simple algebra.
2. If $A\#H$ is a prime algebra and $A$ is $H$-simple then $A\#H$ is a simple algebra.
3. $A^H$ is a prime algebra.
4. $C_Q(A^H) = K$, where $C_Q(A^H)$ is the centralizer of $A^H$ in $Q$.
5. $A$ and $A^H$ satisfy the same multilinear identities.

Based on these results, a Galois type correspondence theory for $X$-outer actions of finite-dimensional pointed Hopf algebras on prime algebras was proved in [247,244,245].

THEOREM 3.4.13. Let $k$ be a field of characteristic zero and let $H$ be a finite-dimensional pointed Hopf algebra over $k$ acting on a prime algebra $A$ such that the action is $X$-outer. Consider $A$ and $K\#H$ as subalgebras of $Q\#H$. For a rationally complete intermediate subalgebra $U$ of $A$, let $\Phi(U)$ denote the centralizer of $U$ in $K\#H$, and for a right coideal subalgebra $\Lambda$ of $K\#H$ containing $K$, let $R^{\Lambda}$ denote the centralizer of $\Lambda$ in $A$. Then $U = R^{\Phi(U)}$ and $\Lambda = \Phi(R^{\Lambda})$. Thus $U \to \Phi(U)$ determines a one to one correspondence between the set of rationally complete intermediate subalgebras $U$ of $A$ and the set of right subcomodule algebras of $K\#H$ containing $K$.

Part 4. Categories of representations of Hopf algebras

The abundance of structures related to Hopf algebras give rise to a number of new constructions. Many constructions are related to quantum groups or to solutions of the quantum Yang–Baxter equation, some are related to Hopf algebras in categories and others are deformations of known objects.

4.1. Rigid tensor categories and Hopf algebras

An important point of view of Hopf algebras arises from categorical considerations.

Let $(A, m, 1)$ be a finite-dimensional unital associative algebra, and let $\mathcal{C} := \text{Rep}(A)$ be the category of finite-dimensional left $A$-modules. Clearly, $\mathcal{C}$ is a $k$-linear Abelian category. Also, the algebra $A$ can be reconstructed from $\mathcal{C}$ and the forgetful functor $\mathcal{C} \to \text{Vec}$.

Suppose that $A = H$ is a Hopf algebra. Then $\mathcal{C}$ turns out to have a rich structure as seen in the following:

1. Since $\varepsilon$ is an algebra map, $k$ becomes an object of $\mathcal{C}$ by
   
   $$a \cdot x := \varepsilon(a)x$$

   for any $a \in H$, $x \in k$.

2. Since $\Delta$ is an algebra map it follows that for any $V, W \in \mathcal{C}$, $V \otimes W$ becomes an object of $\mathcal{C}$ by:

   $$a \cdot (v \otimes w) := \sum a_1 \cdot v \otimes a_2 \cdot w$$

   for any $a \in H$, $v \in V$ and $w \in W$. 

(3) The coassociativity of $\Delta$ implies that the standard associativity isomorphism $(U \otimes V) \otimes W \to U \otimes (V \otimes W)$ is an $H$-module map.

(4) Since $S$ is an anti-algebra isomorphism it follows that for any $V \in \mathcal{C}$, its linear dual $V^*$ becomes an object of $\mathcal{C}$ by:

$$(a \cdot f)(v) := f(S(a) \cdot v)$$

for any $a \in H$, $v \in V$ and $f \in V^*$.

(5) Let $V \in \mathcal{C}$ then it is straightforward to verify that the maps

$${\text{ev}_V : V^* \otimes V \to k, \quad f \otimes v \mapsto f(v)}$$

and $${\text{coev}_V : k \to V \otimes V^* \text{ determined by} \quad 1 \mapsto \sum_{i} v_i \otimes f_i}$$

are in fact $H$-module maps, where $\{v_i\}$ and $\{f_i\}$ are dual bases of $V$, $V^*$ respectively.

The above indicate that $(\text{Rep}(H), \otimes, k, a, l, r)$ is a rigid tensor (also called monoidal) category where $k$ is the unit object and $a, l, r$ are the standard associativity and unit constrains as in Vec. (See [150,151,200] for definitions and properties of such categories.)

In the following Tannaka–Krein type theorem it is shown that corresponding conditions on $\text{Rep}(A)$ when $A$ is an algebra, induce a Hopf algebra structure on $A$ (see, e.g., [89]):

**THEOREM 4.1.1.** Let $(A, m, 1)$ be an algebra and $\mathcal{C} := \text{Rep}(A)$. Then there exists a bijection between:

1. rigid tensor structures on $\mathcal{C}$, together with a compatible tensor structure on the forgetful functor $\text{Forget} : \mathcal{C} \to \text{Vec}$, and
2. Hopf algebra structures on $(A, m, 1)$.

**REMARK 4.1.2.** If we omit the rigidity requirement from $\mathcal{C}$ then the corresponding structure on $A$ is that of a bialgebra.

Similarly, the category $H \mathcal{M}$ of left $H$-comodules has a structure of a rigid tensor category as well. It is thus useful to think of a Hopf algebra as an algebra (coalgebra) whose representation category (the category of $H$-comodules) has a structure of a rigid tensor category. The Hopf algebra structure is unique only up to twisting as discussed in the following Theorem 4.1.4.

**DEFINITION 4.1.3.** Let $H$ and $L$ be Hopf algebras. Then $H$ and $L$ are monoidally co-Morita equivalent (or Morita–Takeuchi equivalent) if the categories $L \mathcal{M}$ and $H \mathcal{M}$ are equivalent as tensor categories.

If $H$ can be obtained from $L$ by twisting the algebra structure then $L \mathcal{M}$ and $H \mathcal{M}$ are monoidally co-Morita equivalent. But more is true:
**Theorem 4.1.4** [204]. Let $H$ and $L$ be Hopf algebras. Then:

1. $H$ and $L$ are monoidally co-Morita equivalent ⇔ there exists an $(H, L)$-bi-Galois object $M$.
2. If a crossed product over $k$ is an $(H, L)$-bi-Galois object then $H$ can be obtained from $L$ by twisting the algebra structure.
3. As a corollary of (2) and Theorem 3.2.9(1) if $H$ and $L$ are finite-dimensional monoidally co-Morita equivalent then $H$ can be obtained from $L$ by twisting the algebra structure.

The following is a very useful criteria for co-Morita equivalence.

**Theorem 4.1.5** [149]. Suppose $K$ is a sub-Hopf algebra of a Hopf algebra $H$. If $I, J$ are Hopf ideals in $K$ so that $I = g^{-1}Jg$ for some $g \in G(K)^*$, then $H/(I)$ and $H/(J)$ are monoidally co-Morita equivalent (where $(I)$ and $(J)$ are the Hopf ideals in $H$ generated by $I$ and $J$ respectively).

The striking property of quasitriangular (coquasitriangular) Hopf algebras $H$ is that the tensor category $\mathcal{C} = \text{Rep}(H)$ of their category of representations (comodules) is also braided. Namely, just like in group representations, there is a natural isomorphism between $X \otimes Y$ and $Y \otimes X$ for any two representations (comodules) $X, Y$ (this is not necessarily true in general). Specifically, if $(H, R)$ is quasitriangular where $R = \sum R_1 \otimes R_2$, we can define for any $X, Y \in \mathcal{C}$ an isomorphism

$$c_{X,Y} : X \otimes Y \rightarrow Y \otimes X$$

given by:

$$x \otimes y \mapsto \tau(R \cdot (x \otimes y)) = \sum R_2 \cdot y \otimes R_1 \cdot x.$$ 

Analogously we can define for comodules over a coquasitriangular Hopf algebra $(H, \langle \rangle)$

$$x \otimes y \mapsto \sum \langle x_1 | y_1 \rangle y_0 \otimes x_0$$

for all $x \in X, y \in Y$.

The map $c_{X,Y}$ above is in fact an $H$-module map by property (QT2) of $R$ (see Definition 1.5.1). The collection $c := \{c_{X,Y} | X, Y \in \mathcal{C}\}$ determines a braided structure on $\mathcal{C}$. The meaning of (QT1) in the definition is that to move $Z$ to the left or to the right of $X \otimes Y$ is the same thing as to permute $X, Y$ separately with $Z$. (The reader is referred to [112,113] for a detailed discussion of braided tensor categories.)

If $(H, R)$ is triangular, then $R$ determines a symmetric structure on $\mathcal{C}$. The meaning of this is that the composition $X \otimes Y \xrightarrow{c_{X,Y}} Y \otimes X \xrightarrow{c_{Y,X}} X \otimes Y$ is the identity for every $X, Y$. This is a generalization of the standard flip $\tau$ in the category of representations of cocommutative Hopf algebras (e.g., the universal enveloping algebra $U(g)$ or the group algebra $kG$).
REMARK 4.1.6. If \( C \) is a braided tensor category then the braiding gives rise to a representation of the braid group \( \mathbb{B}_n \) on \( V \otimes^n \), \( V \in C \), in the following sense: Each generator \( \sigma_i \) of \( \mathbb{B}_n \) acts on \( V \otimes^n \) by applying the braiding \( c_{V,V} \) to the \((i, i + 1)\) component of \( V \otimes^n \). Explicitly, the representation \( \rho_n : \mathbb{B}_n \to \text{Aut}(V \otimes^n) \) is given by:

\[
\rho_n(\sigma_i) = \text{id} \otimes^{i-1} c_{V,V} \otimes \text{id} \otimes^n \otimes^{i-1}.
\]

If the category is symmetric then it gives rise to a non-trivial representation of the symmetric group \( S_n \) on the \( n \)-th tensor product of objects of \( C \).

REMARK 4.1.7. Theorem 4.1.1 extends to a bijective assignment between (quasi)triangular structures on \( H \), and (braided) symmetric rigid tensor structures on \( C \).

Going a step further, just as algebras can be reconstructed from their category of representations one can reconstruct a ((co)quasitriangular) Hopf algebra from a rigid tensor category which admits a fiber (\( = \) exact, faithful and tensor) functor to the category of vector spaces. This problem has been studied extensively by many authors who have considered various possible set-ups and have accordingly reconstructed various structures (cf. \([200,58,187,236,112,154,155,157,202]\)). Here is one example:

**Theorem 4.1.8** \([236]\). Let \( C \) be a small Abelian rigid tensor category and \( F \) a \( k \)-linear fiber functor to \( \text{Vec} \). Then there exists a Hopf algebra \( H \) such that \( C \) is equivalent to \( H \mathcal{M} \), the category of left \( H \)-comodules, and \( F \) is isomorphic to the forgetful functor.

The basic idea of the proof is the following: For each \( V \in \text{Vec} \) define a functor \( F_V : C \to \text{Vec} \) by \( X \mapsto F(X) \otimes V \) for all \( X \in C \). The finiteness assumptions on the category imply that the functor \( V \mapsto \text{Mor}(F, F_V) \) is representable, that is, there exists an object \( H \in \text{Vec} \) such that

\[
\text{Hom}_k(H, V) = \text{Mor}(F, F_V)
\]

for all \( V \in \text{Vec} \). Then \( H \) is our desired Hopf algebra where the structure maps of \( H \) are reconstructed as well.

To the notion of a ribbon Hopf algebra there corresponds the notion of a ribbon category. A ribbon category \( C \) is a rigid braided tensor category with the following extra structure: There exists an automorphism of the identity functor \( \text{id} \) of \( C \), which is compatible with the tensor product, braiding and taking duals in a certain natural sense. In ribbon categories it is possible to define dimensions of objects (sometimes called quantum dimensions), and more generally to define traces of endomorphisms. This allows to associate link invariants to any ribbon category. In particular, all classical polynomial invariants (e.g., Jones polynomial) can be constructed in this fashion. (The reader is referred to \([120–122,194]\) for an extensive study of ribbon Hopf algebras and their connections with Hennings’ and Kauffman’s invariants of knots, links and 3-manifolds.)

An important class of ribbon categories is the class of modular categories, \([15,232]\). A modular category \( C \) is a semisimple ribbon category with finitely many (up to isomor-
irreducible objects \( \{V_i \mid 0 \leq i \leq m\} \) with \( V_0 \) as the unit object, so that the matrix \( s := (s_{ij}) \), where \( s_{ij} := \text{tr}(c_{ij} V_i V_i^* V_j V_j^*) \), is invertible.

**Example 4.1.9** [77]. Let \( H \) be a semisimple Hopf algebra over an algebraically closed field \( k \) of characteristic 0. Then \( \text{Rep}(D(H)) \) is a modular category. This is true essentially since \( D(H) \) is factorizable.

Modular categories arise naturally in physics in the framework of quantum field theory, and in topology in the framework of invariants of 3-manifolds (see, e.g., [232]).

### 4.2. The FRT construction

Let \( G \) be an affine algebraic group over \( k \); we then associate to \( G \) in the usual way two \( k \)-Hopf algebras: \( A(G) \) (sometimes denoted by \( O(G) \)), whose elements are representative functions on \( G \), and \( U(g) \), whose underlying \( k \)-algebra is the enveloping algebra of the Lie algebra \( g \) of \( G \). Those Hopf algebras certified by workers in the field as being "quantum groups", fall into two main classes: those deforming the type \( U(g) \), the quantum enveloping algebras, and those deforming the type \( A(G) \).

Perhaps the earliest systematic construction of infinite families of these two types of Hopf algebras, was furnished by the following seminal work of Faddeev, Reshetikhin and Takhtajan, [90,91]. It was studied later extensively by many authors (e.g., [12,40,103,136, 230,202,218,226] and others in quantum group theory).

Let \( V \) be a finite-dimensional vector space and let \( R : V \otimes V \to V \otimes V \) be a linear isomorphism that satisfies the braid relation, namely:

\[
(R \otimes I_V) \circ (I_V \otimes R) \circ (R \otimes I_V) = (I_V \otimes R) \circ (R \otimes I_V) \circ (I_V \otimes R).
\]

Let \( \{v_1, \ldots, v_n\} \) be a basis of \( V \) and assume \( R \) as above is given by:

\[
R(v_i \otimes v_j) = \sum_{k,l=1}^{n} R_{k,l}^{i,j} v_k \otimes v_l.
\]

Recall, Example 1.1.18, that \( \text{End}(V)^* \) is a coalgebra and \( V \) is a right \( \text{End}(V)^* \)-comodule. Let \( T \) be the tensor algebra of \( \text{End}(V)^* \). Then \( T \) is a bialgebra by extending the coproduct on \( \text{End}(V)^* \) to \( T \) multiplicatively. We wish to construct a bialgebra \( A(R) = T/I \), for some biideal \( I \), so that the map \( R \) will induce a braiding in the category of right \( A(R) \)-comodules. This is given by the following defining relations which are precisely those needed to make \( R \) into a right \( A(R) \)-comodule map.

\[
\sum_{I,J=1}^{n} R_{I,J}^{i',j'} T_{i}^{I} T_{j}^{J} = \sum_{I,J=1}^{n} R_{I,J}^{i',j'} T_{i}^{J} T_{j}^{I},
\]

all \( 1 \leq i, i', j, j' \leq n \).
We have:

**Theorem 4.2.1** [90]. Let \( T \) be as above and let \( I \) be the ideal of \( T \) generated by formula (1). Then \( I \) is a biideal and thus \( A(R) \) is a bialgebra.

A Hopf algebra version \( H(R) \) of the FRT-construction is given in [202] with regard to a rigid tensor category. It was proved in [136] that \( A(R) \) is coquasitriangular with a braiding structure given on generators by

\[
\langle T^i_j | T^l_k \rangle_R = R^l_j_{ik}.
\]

The paper [90] goes on to construct inside \((A(R))^0\) a bialgebra, yet not a Hopf algebra. This result was improved by Faddeev, Reshetikhin and Takhtajan in a later paper [91], where they construct inside \((A(R))^0\) a Hopf algebra \( \hat{U}(R) \), properly containing their earlier construction. The Hopf algebra \( \hat{U}(R) \) is defined as follows:

Let \( l^+, l^-, r^+, r^- \) be the following maps from \( A(R) \) to \( A(R)^0 \):

\[
l^+(a) = \langle a | - \rangle_R,
\]

\[
r^+(a) = \langle - | a \rangle_R,
\]

\[
l^-(a) = \langle a | - \rangle^*,
\]

\[
r^-(a) = \langle - | a \rangle^*,
\]

where \( \langle | \rangle^* \) is the convolution inverse of \( \langle | \rangle \). Then the following Hopf algebra is constructed:

\[
\hat{U}(R) = \text{Im}(l^+) + \text{Im}(l^-) + \text{Im}(r^+) + \text{Im}(r^-) \subset (A(R))^0.
\]

The Hopf algebra \( (\hat{U}(R), \hat{R}) \) is essentially quasitriangular in the following sense: The braiding \( \langle | \rangle_R = \hat{R} \) is an element of \((A(R) \otimes A(R))^*\). But more is true, \( \hat{R} \) is an element of \( \hat{U}(R) \otimes \hat{U}(R) \) which is the topological completion of \( \hat{U}(R) \otimes \hat{U}(R) \) in the topological space \((A(R) \otimes A(R))^*\). (The topological aspects are discussed in details in [132].)

4.3. Yetter–Drinfeld categories and the Drinfeld double

One of the most significant categorical aspects of bialgebras \( H \) was introduced by Yetter in [248]. The coalgebra and algebra structure of \( H \) are taken into account simultaneously.

**Definition 4.3.1.** Let \( H \) be a bialgebra over \( k \). The “Yetter–Drinfeld” category \( H \mathcal{YD} \) \((H \mathcal{YD})^H\) is the category of objects which are left \( H \)-modules, left (right) \( H \)-comodules, and each \( M \in H \mathcal{YD} \) satisfies the compatibility condition, namely, for all \( h \in H \), \( m \in M \)

\[
\sum h_1m_{-1} \otimes h_2 \cdot m_0 = \sum (h_1 \cdot h_2)_{-1}m_0 \otimes (h_1 \cdot m)_0,
\]

where \( \langle | \rangle_R \) is the convolution inverse of \( \langle | \rangle \).
where $\rho(m) = \sum m_{-1} \otimes m_0$ (or respectively

$$\sum h_1 \cdot m_0 \otimes h_2 m_1 = \sum (h_2 \cdot m)_0 \otimes (h_2 \cdot m)_1 h_1,$$

where $\rho(m) = \sum m_0 \otimes m_1$).

Similar compatibility conditions are given for right–right and right–left Yetter–Drinfeld categories.

**Example 4.3.2.**

(1) A particular example of an object in $^H_H \mathcal{YD}$ is $H$ itself considered as a left (right) $H$-comodule via $\Delta$ and a left (right) $H$-module via the left (right) adjoint action.

(2) If $(H, R)$ is a quasitriangular Hopf algebra then every left $H$-module $M$ is in $^H_H \mathcal{YD}$ by defining $\rho : M \mapsto H \otimes M$ by

$$\rho(m) = \sum S(R^1) \otimes R^2 \cdot m.$$  

Similarly, if $(H, \langle | \rangle)$ is a coquasitriangular Hopf algebra then every right $H$-comodule $M$ is in $^H_H \mathcal{YD}$ by defining

$$h \cdot m = \sum \langle m_1 | h \rangle m_0$$

for all $h \in H, m \in M$.

The Yetter–Drinfeld category $^H_H \mathcal{YD}$ has the following natural pre-braiding structure: Given $M, N \in ^H_H \mathcal{YD}$, define $c_{M,N} : M \otimes N \mapsto N \otimes M$ by:

$$c_{M,N}(m \otimes n) = \sum m_{-1} \cdot n \otimes m_0$$

for $m \in M, n \in N$. When $H$ is a Hopf algebra with an invertible antipode then $^H_H \mathcal{YD}$ is a braided tensor category with a braiding structure defined by $c$.

The category $^H_H \mathcal{MD}^H_H$ of two-sided two-cosided Hopf-modules satisfying six compatibility relations (also called tetramodules) was considered in [246]. He discussed the interrelation between $^H_H \mathcal{YD}$ and $^H_H \mathcal{MD}^H_H$. An equivalence between the pre-braided categories of tetramodules and that of Yetter–Drinfeld modules over $H$ was stated in [203].

Another related category is the category of Doi–Hopf modules, [62]. This category includes a variety of modules as special cases; for example Hopf modules and graded modules are Doi–Hopf modules. Furthermore, it was proved in [33] that $^H_H \mathcal{YD}$ can be considered as a special case of Doi–Hopf modules and in [18] that the same holds for two-sided two-cosided Hopf modules, illustrating the “unifying” property of Doi’s concept. In fact, the above mentioned category equivalence between Yetter–Drinfeld modules and tetramodules can be described in terms of an adjoint pair of functors between categories of Doi–Hopf modules (in the sense of [34]).
In a Yetter–Drinfeld category we can consider commutativity in the category. This is defined in general as follows:

**Definition 4.3.3.** A left $H$-module and left $H$-comodule algebra $(A, \cdot, \rho)$ is called $H$-commutative (or quantum commutative) if

$$ab = \sum (a_{-1} \cdot b)a_0 \quad \text{for all } a, b \in A.$$ 

$H$-commutativity is defined similarly for a right $H$-module and right $H$-comodule algebras, etc.

**Remark 4.3.4.** If $(A, \cdot, \rho)$ is $H$-commutative then $A^{{\co}H}$ and $A^H$ are contained in $Z(A)$.

**Example 4.3.5.** The following are examples of $H$-commutative algebras.

1. Let $H$ be as in Example 4.3.2(1). Then $(H, \text{ad}_l, \Delta)$ $(H, \text{ad}_r, \Delta)$ resp.) is $H$-commutative.

2. [49,160]. Let $A$ be a commutative superalgebra, that is $A$ is a $\mathbb{Z}_2$-graded algebra and $ab = (-1)^{\deg a \deg b} ba$ for homogeneous elements $a, b \in A$. Let $G = \{1, g\} \cong \mathbb{Z}_2$ and $H = kG$. Consider $A$ as an $H$-comodule by the $\mathbb{Z}_2$-grading and an $H$-module by defining $g \cdot a = (-1)^{\deg a} a$ for homogeneous $a \in A$. Then $A$ is $H$-commutative.

3. [49]. Let $A := \mathbb{C}_q[x, y]$, the quantum plane, that is $A$ equals the free algebra $\mathbb{C}(x, y)$ modulo the relation $xy = qyx$, where $q$ is an $n$-th root of 1. Let $G := \mathbb{Z}_n \times \mathbb{Z}_n$ and $H = kG$, then $A$ is $H$-commutative for a certain action and coaction of $H$. If we localize and obtain $B := A[x^{-1}, y^{-1}]$ then $B/B^H$ is $H$-commutative and $H^*$-Galois.

Recall the Miyashita–Ulbrich action defined in Definition 3.2.6. The following theorem describes how $H$-commutativity is related to $H$-Galois extensions and to objects in $\mathcal{YD}_H^H$ via the Miyashita–Ulbrich action $\leftarrow$.

**Theorem 4.3.6.** Let $A^{{\co}H} \subset A$ be a right $H$-Galois extension, then:

1. [235,68]. $(A^{{\co}H}, \leftarrow, \rho)$ is $H$-commutative.

2. [35]. If $H$ has a bijective antipode then

$$(A^{{\co}H}, \leftarrow, \rho) \in \mathcal{YD}_H^H.$$ 

In particular, if $A^{{\co}H} \subset Z(A)$ then $(A, \leftarrow, \rho) \in \mathcal{YD}_H^H$.

3. [45]. If $A$ is a right $H$-commutative $H$-module algebra, then $C_A(A^{{\co}H}) = A$, the given action $\cdot$ coincides with $\leftarrow$ and $(A, \cdot, \rho) \in \mathcal{YD}_H^H$.

**Example 4.3.7.** Let $A = H$ be a right $H$-comodule with $\rho = \Delta$. Then by Example 3.2.2(2), $k \subset H$ is a right $H$-Galois with $h^{-1}(1 \otimes h) = \sum S(h_1) \otimes h_2$ and so the Miyashita–Ulbrich action $x \leftarrow h = \sum S(h_1)xh_2$ is the right adjoint action.
Theorem 4.3.6(3) generalizes now the right version of Example 4.3.2(1).

The Drinfeld double

If \( H \) is finite-dimensional then \( \mathcal{YD}^H \) has a nice realization related to the so-called Drinfeld double of \( H \) constructed by Drinfeld, \([71]\). The Drinfeld double is a double crossproduct, a construction described in \([153]\) and modified in \([192]\). The double crossproduct of the bialgebras \( H \) and \( B \) is defined when \( B \) is a left \( H \)-module coalgebra and \( H \) is a right \( B \)-module coalgebra satisfying some compatibility conditions. A special case is when \( H \) is a finite-dimensional Hopf algebra, \( B = H^* \) and the actions are given by the left coadjoint action of \( H \) on \( H^* \)

\[
h \longrightarrow p = \sum h_1 \rightarrow p \leftarrow S^{-1}(h_2)
\]

and the right coadjoint action of \( H^* \) on \( H \)

\[
h \longrightarrow p = \sum S^{s^{-1}}(p_1) \rightarrow h \leftarrow p_2
\]

for \( h \in H, \ p \in H^* \).

**Definition 4.3.8.** Let \( H \) be a finite-dimensional Hopf algebra. The Drinfeld double \( D(H) = H^{*\cop} \bowtie H \) is defined as the vector space \( H^{*\cop} \otimes H \) with multiplication defined by

\[
(p \bowtie h)(p' \bowtie h') = p(h_1 \rightarrow p'_2) \bowtie (h_2 \leftarrow p'_1)h'
\]

and comultiplication given by the tensor comultiplication in the tensor coalgebra \( H^{*\cop} \otimes H \), that is:

\[
\Delta_{D(H)}(p \bowtie h) = (p_2 \bowtie h_1) \otimes (p_1 \bowtie h_2)
\]

for all \( h \in H, \ p \in H^* \).

The antipode is given by

\[
S_{D(H)}(p \bowtie h) = (1 \bowtie S(h))(S^*(p) \bowtie 1).
\]

More precise formulas for the multiplication in \( D(H) \) are given by:

\[
(p \bowtie h)(p' \bowtie h') = \sum p(h_1 \rightarrow p' \leftarrow S^{-1}(h_3)) \bowtie h_2' p',
\]

\[
(p \bowtie h)(p' \bowtie h') = \sum pp' \bowtie (S^{s^{-1}}(p_1) \rightarrow h \leftarrow p_3')h'.
\]

Observe that the Hopf algebras \( H \) and \( H^{*\cop} \) are contained in \( D(H) \) hence a left \( D(H) \)-module \( M \) is in particular a left \( H \)-module and a left \( H^* \)-module. Thus, by Remark 1.3.4,
$M$ is also a right $H$-comodule. Now, a straightforward (long) verification shows that the definition of the multiplication in $D(H)$ implies that $h \cdot (p \cdot m) = ((h \triangleleft h)(p \triangleleft 1)) \cdot m$ for all $m \in M$, $h \in H$ is equivalent to $M$ being an object in $\mathcal{H} \mathcal{Y} \mathcal{D}^H$. We summarize:

**Theorem 4.3.9** [156]. Let $H$ be a finite-dimensional Hopf algebra. Then the Yetter–Drinfeld category $\mathcal{H} \mathcal{Y} \mathcal{D}^H$ is equivalent to the category of left modules over the Drinfeld double $D(H)$.

**Remark 4.3.10.** The process of taking the double is mostly effective if it is done just once; the double of $D(H)$ can be obtained from the tensor product $D(H) \otimes D(H)$ by twisting the comultiplication, [197,209].

The Drinfeld double $D(H)$ is naturally quasitriangular by letting

$$R := \sum (\varepsilon \triangleleft h_i) \otimes (h_i^* \triangleleft 1),$$

where $\{h_i\}$ and $\{h_i^*\}$ are any dual bases of $H$ and $H^*$. Consequently, $(D(H)^*)^o \langle \quad \rangle_R$ is coquasitriangular. Moreover, $D(H)$ is factorizable (hence unimodular), [71,197,194].

**Remark 4.3.11.**

1. [140]. Recall, Example 1.6.3, that $\sigma = \langle \quad \rangle_R \circ \tau$ is a Hopf 2-cocycle on $D(H)^*$. Then

$$\sigma D(H)^* \cong \hat{\tau}(H) = H \# H^*$$

as algebras, where $\sigma D(H)^*$ is a twisted Hopf algebra as defined in Remark 3.1.4 and $\hat{\tau}(H)$ is the so-called “Heisenberg double” of $H$; it is a simple algebra.

2. It is straightforward to check that

$$J := \sum_i (h_i^* \otimes 1) \otimes (\varepsilon \otimes h_i)$$

is a twist for the Hopf algebra $H^{op} \otimes H$ where $\{h_i\}$ and $\{h_i^*\}$ are dual bases in $H$ and $H^*$ respectively. Then $(H^{op} \otimes H)^J$, the Hopf algebra obtained by twisting the comultiplication via $J$ is isomorphic to $D(H)^{op}$, the (opposite) dual of the Drinfeld double of $H$.

For a braided monoidal category $\mathcal{C}$, a Brauer group $\text{Br}(\mathcal{C})$ was defined by [240] so that many Brauer group constructions are particular cases. For example, the classical Brauer group of a commutative ring $k$ is $\text{Br}(\mathcal{C})$ for $\mathcal{C}$ the category of $k$-modules; the Brauer group of a scheme $(X, O_X)$ is $\text{Br}(\mathcal{C})$ where is $\mathcal{C}$ is the category of $O_X$-module sheaves; the Brauer–Long group is $\text{Br}(\mathcal{C})$ for $\mathcal{C}$ the category of $H$-dimodules (where $H$ is a commutative cocommutative Hopf algebra).

The Brauer–Long group was generalized and denoted by $\text{BQ}(k, H)$ where $H$ is any Hopf algebra with a bijective antipode and the dimodules are replaced by Yetter–Drinfeld modules. It is proved:
THEOREM 4.3.12 [241]. Let \( H \) be a finite-dimensional Hopf algebra then there is an exact sequence

\[
1 \to G(D(H)^+ \to G(D(H) \to \text{Aut}_{\text{Hopf}}(H) \to BQ(k, H),
\]

where \( D(H) \) is the Drinfeld double of \( H \).

As a consequence, it follows that the Brauer group of Sweedler’s four-dimensional Hopf algebra \( H_4 \) contains \( k^*/\{-1, 1\} \) as a subgroup and thus \( BQ(k, H) \) is in general highly non-torsion.

4.4. Hopf algebras in braided categories, biproducts and bosonizations

One of the first comprehensive steps to abstract the notion of a Hopf algebra from the work of Hopf in topology was carried out by Milnor and Moore, [163]. Though their notion of a Hopf algebra is not the one used in this chapter, it turns out to be a Hopf algebra in the category of \( k\mathbb{Z} \)-comodules. Specifically,

EXAMPLE 4.4.1 [163]. Let \( A = A_0 \oplus A_1 \oplus \cdots \) be a \( \mathbb{N} \)-graded vector space over a field \( k \). If \( A \) and \( B \) are \( \mathbb{N} \)-graded then the “twisting morphism” \( T: A \otimes B \to B \otimes A \) is the morphism defined by:

\[
T(a \otimes b) = (-1)^{pq} b \otimes a
\]

for \( a \in A_p, b \in B_q \).

\( A \) is called a graded Hopf algebra over \( k \) if

1. \((A, \mu, 1)\) is a graded \( k \)-algebra (as usual).
2. \((A, \Delta, \varepsilon)\) is a graded \( k \)-coalgebra, viz., \( \Delta(A_n) \subset \sum_{i=0}^n A_i \otimes A_{n-i} \) and \( \sum_{i>0} A_i \subset \ker \varepsilon \).
3. \( \Delta \circ \mu = (\mu \otimes \mu) \circ (\text{id} \otimes T \otimes \text{id}) \circ (\Delta \otimes \Delta) \) and \( \varepsilon \) is an algebra map.
4. The identity map \( \text{id}_A \) is invertible under convolution. In particular, its inverse \( S \) satisfies \( S \circ \mu = \mu \circ T \circ (S \otimes S) \).

It is condition (3) in the above example that reflects the basic idea of a bialgebra in a braided category, where the braiding plays the role of the “twisting operator”.

Additional examples in the same spirit are enveloping algebras of Lie superalgebras and more generally, of Lie color algebras, [205].

These examples were known without the formalism of category theory. The more general notion of Hopf algebras in braided tensor categories was introduced in [159] and have been since extensively studied by many authors. A comprehensive survey is given in [228].

Notions like algebras, coalgebras and bialgebras can be considered categorically, that is, all structure maps are required to be maps in the category. \((H\)-module algebras and \( H \)-comodule algebras are examples that have already been mentioned.) To define a bialgebra in a category requires an appropriate braiding. Explicitly:
Given a braided tensor category with a braiding structure \( c \), one can define an algebra structure on the tensor product of two algebras as follows: For any \((A, \mu_A)\) and \((B, \mu_B)\) in the category define \( \mu_{A \otimes B} : (A \otimes B) \otimes (A \otimes B) \mapsto A \otimes B \) by:

\[
\mu_{A \otimes B} = (\mu_A \otimes \mu_B)(\text{id} \otimes c_{B,A} \otimes \text{id}).
\]

Once the tensor product of two algebras is an algebra we can define:

**Definition 4.4.2.** A bialgebra in a braided tensor category is a 5-tuple \((A, \mu, 1, \Delta, \varepsilon)\) where \((A, \mu, 1)\) is an algebra in the category and \((A, \Delta, \varepsilon)\) is a coalgebra in the category so that

\[
\Delta \circ \mu_A = \mu_{A \otimes A} \circ (\Delta \otimes \Delta).
\]

A is a Hopf algebra in the category if moreover the identity \( \text{id}_A \) is invertible under convolution; its inverse is the antipode of \( A \).

Hopf algebras in the category are also called braided Hopf algebras.

**Example 4.4.3** [11]. Let \( H \) be any Hopf algebra and let \( V \in H \mathcal{YD} \). Then the tensor algebra \( T(V) = \bigoplus_{n \geq 0} T(V)(n) \), where \( T(V)(n) = V^\otimes n \) is also an object in \( H \mathcal{YD} \). If we define \( \Delta_V(v) = 1 \otimes v + v \otimes 1 \) for all \( v \in V \), then there is a unique extension of \( \Delta_V \) to a map \( \Delta : T(V) \rightarrow T(V) \otimes T(V) \) which is an algebra map in \( H \mathcal{YD} \). The counit \( \varepsilon \) is defined by \( \varepsilon(v) = 0 \) for all \( v \in V \) and thus \( T(V) \) is a (graded) bialgebra in the category. Moreover, it can be proved that \( T(V) \) is actually a Hopf algebra in \( H \mathcal{YD} \).

Let \( I \) be the largest Hopf ideal generated by homogeneous elements of degree \( > 1 \). Then \( B(V) = T(V)/I \) is a (graded) Hopf algebra in \( H \mathcal{YD} \). The graded braided Hopf algebra \( B(V) \) is unique with respect to the following properties: \( B(V) \) is connected as a coalgebra, generated as an algebra by elements of degree 1 and \( V = B(V)(1) = P(B(V)) \), the space of primitive elements of \( B(V) \).

The braided Hopf algebra \( B(V) \) was termed the Nichols algebra of \( V \) honoring Nichols who described \( B(V) \) in a different setting, [177].

Nichols algebras were rediscovered and studied independently by several authors. They were treated as the invariant parts of “algebras of quantum differential forms” in [246] and as “quantum symmetric algebras” in [199].

It has been proved that most of the fundamental properties of ordinary finite-dimensional Hopf algebras can be generalized to braided Hopf algebra theory, even in a more generalized form, [228]. We summarize:

**Theorem 4.4.4.** The following is true for Hopf algebras in a category:

1. [143]. The fundamental theorem for Hopf modules (Theorem 2.1.3).
2. [143,227]. The bijectivity of the antipode (Theorem 2.2.6).
3. [143,227]. The uniqueness of the integral (Theorem 2.2.1).
4. [95]. Frobenius property (Theorem 2.2.5).
Starting from a Hopf algebra $B$ in the braided tensor category $H_{H}YD$ it is possible to “lift” $B$ to an ordinary Hopf algebra. The process was given in [190] (without using the notion of a Hopf algebra in the category) and in [158].

To describe the process we need first to introduce the smash coproduct.

Let $H$ be a bialgebra and $A$ a left $H$-comodule coalgebra. To avoid confusion we write for all $a \in A$:

$$\Delta_{A}(a) = \sum a^{1} \otimes a^{2} \quad \text{and} \quad \rho_{H}(a) = \sum a_{-1} \otimes a_{0}.$$ 

Then the tensor product $A \otimes H$ can be equipped with a coalgebra structure via the smash coproduct as follows:

**Proposition 4.4.5** [164]. Let $H$ be a Hopf algebra and let $A$ be a left $H$-comodule coalgebra. Let $A\#H$ be the vector space $A \otimes H$ with coproduct given by

$$\Delta(a\#h) = \sum a^{1}\#(a^{2}_{-1}h_{1} \otimes (a^{2})_{0}\#h_{2}$$

and counit given by $\varepsilon(a\#h) = \varepsilon_{A}(a)\varepsilon_{H}(h)$. Then the above structure maps make $A\#H$ into a coalgebra.

Let $A \otimes H$ be equipped with the smash product and the smash coproduct. We call it a biproduct and denote it by $A \ast H$. Necessary and sufficient conditions for the biproduct to be a Hopf algebra were given by Radford in [190]. We give a categorical version of the theorem.

**Theorem 4.4.6.** Let $H$ be a bialgebra and let $A$ be an algebra in $H_{H}Mod$ and a coalgebra in $H_{H}Com$. Then $A \ast H$ is a bialgebra if and only if $A$ is a bialgebra in $H_{H}YD$.

If $H$ is a Hopf algebra with a bijective antipode $S_{H}$ and $A$ is a Hopf algebra in $H_{H}YD$ with an antipode $S_{A}$ then $A \ast H$ is a Hopf algebra with antipode given by:

$$S(a \ast h) = (1 \ast S_{H}^{-1}(a_{-1}h))(S_{A}(a_{0}) \ast 1).$$

A similar process was named bosonization, [158]. He considers the braided category $H_{H}Mod$ over a quasitriangular Hopf algebra $H$ and proves moreover that if $H$ is triangular and $A$ is quasitriangular in the category, then $A \ast H$ is quasitriangular.

**Example 4.4.7.** Let $H = H_{4}$ be Sweedler’s 4-dimensional Hopf algebra (Example 1.2.5) then $H = A \ast kG$ where $G = \{1, g\} \cong \mathbb{Z}_{2}$ and $A = sp_{k}\{1, x\}$. 

The following is a version of a structure theorem about biproducts. It is most useful in the classification theory of finite-dimensional Hopf algebras.

**Theorem 4.4.8** [190]. If $H \overset{i}{\hookrightarrow} B \overset{\pi}{\rightarrow} H$ is a sequence of finite-dimensional Hopf algebra maps where $i$ is injective, $\pi$ is surjective and $\pi \circ i = \text{id}_H$, then there exists a coideal subalgebra $A \subset B$ such that:

1. $A$ is a left $H$-module algebra via the adjoint action.
2. $A$ is a left $H$-comodule algebra via $\rho(b) = \sum a^{(1)} \otimes a^{(2)} = \sum \pi(a^{(1)}) \otimes a^{(2)}$.
3. $A \cong B/BH^+$ as Hopf algebras (where $H^+ = \text{Ker} \, \varepsilon$).
4. $A$ is a Hopf algebra in the category $H$-YD.
5. $B \cong A \ast H$ as a bialgebra.

Bosonizations and biproducts were used in [46,93] to prove a generalized Schur double centralizer theorem for Lie algebras in certain symmetric tensor categories $C$. Explicitly, for a finite-dimensional object $V \in C$ one can define the $C$-analogue, $U(gl_C(V))$, of the enveloping Lie algebra $U(gl(V))$ by using the braiding structure. Then $U(gl_C(V))$ is a Hopf algebra in the category $C$ and thus by Theorem 4.4.6, $\hat{H} = U(gl_C(V)) \ast H$ is an ordinary Hopf algebra. The Hopf algebra $\hat{H}$ acts on $V^{\otimes n}$ via $\Delta$ while the symmetric group $S_n$ acts on $V^{\otimes n}$ via the usual flip map. It was proved:

**Theorem 4.4.9** [46]. Let $C$, $V$, $U(gl_C(V))$ and $\hat{H}$ be as above and assume that the characteristic of $k$ is 0. Then the actions of $kS_n$ and $\hat{H}$ on $V^{\otimes n}$ centralize each other.

[93] have proved in the same spirit a double centralizer theorem for Lie color algebras.

**Part 5. Structure theory for special classes of Hopf algebras**

**5.1. Semisimple Hopf algebras**

There are several surveys regarding semisimple Hopf algebras, the reader is referred to [168,169,4].

Observe first that since $\text{Ker} \, \varepsilon$ is a non-zero ideal of $H$, it follows that there exist no Hopf algebras which are simple as algebras. Thus the simplest objects are the semisimple Hopf algebras.

A consequence of Corollary 2.2.3 is that all semisimple Hopf algebras are finite-dimensional. For if $H$ is semisimple then $H = I \oplus \text{Ker} \, \varepsilon$ where $I$ is a 1-dimensional left ideal of $H$ (since $\text{Ker} \, \varepsilon$ has codimension 1). Moreover, if $H$ is a semisimple Hopf algebra then it is a separable algebra (i.e. for any field extension $E \supseteq k$, $H \otimes E$ is semisimple). This is easily seen from Maschke’s theorem (Theorem 2.3.1) and from the fact that the extensions of $\Delta$, $\varepsilon$, $S$ to $\overline{H} = H \otimes E$ make $\overline{H}$ a Hopf algebra over $E$ with integral $\int^I_H \otimes E$.

Semisimple Hopf algebras in characteristic 0 are close in spirit to $kG$, $G$ a finite group as will be seen in this section. Some of Kaplansky’s conjectures, [116], are inspired by this resemblance (see [213] for a detailed exposition).
The square of the antipode

Kaplansky’s 5th Conjecture. Let $H$ be a semisimple Hopf algebra. Then $S^2 = \text{id}$. In [131] it is proved that if $H$ is semisimple over an algebraically closed field then $S^2$ is an inner automorphism, and in [183] the same is proved over any base field $k$. Observe that from Theorems 2.2.7 and 2.3.1, we deduce:

**Theorem 5.1.1.** Let $H$ be a finite-dimensional Hopf algebra then $H$ is semisimple and cosemisimple if and only if $\text{Tr}(S^2) \neq 0$.

By using Theorem 5.1.1, a positive answer to Kaplansky’s 5-th conjecture in characteristic 0 was given by Larson and Radford:

**Theorem 5.1.2** [133,134]. Let $H$ be a finite-dimensional Hopf algebra and assume that the base field $k$ has characteristic 0. Then the following are equivalent:

1. $S^2 = \text{id}$.
2. $H$ and $H^*$ are semisimple.
3. $H$ is semisimple.
4. $H^*$ is semisimple.

When the characteristic of $k$ is positive then (1) $\Rightarrow$ (3) is trivially false, for example, if $H = kG$ and the characteristic of $k$ divides the order of $G$.

In positive characteristic it is thus natural to consider Hopf algebras which are both semisimple and cosemisimple. This was manifested in [76] where it was proved that any such Hopf algebra in positive characteristic can be lifted to a semisimple (and hence cosemisimple) Hopf algebra in characteristic 0 of the same dimension. This implies that, essentially, it is enough to study semisimple Hopf algebras in characteristic 0, and lift the results to positive characteristic. The proof of the lifting theorem uses Witt vectors (see, e.g., [211]), and the so-called Gerstenhaber–Schack cohomology, [99]. In particular it allowed to prove:

**Theorem 5.1.3** [77]. Let $H$ be a semisimple and cosemisimple Hopf algebra over any field $k$. Then $S^2 = \text{id}$. Moreover, if $H$ is any finite-dimensional Hopf algebra over any field $k$, then $H$ is semisimple and cosemisimple $\Leftrightarrow S^2 = \text{id}$ and $\text{dim}(H) \neq 0$ in $k$.

If $H$ is a semisimple Hopf algebra and the characteristic of $k$ is large enough in comparison to the dimension of $H$ then $H$ is also cosemisimple and hence $S^2 = \text{id}$ (see [212,76]).

Character theory

Motivated by group representation theory, a basic tool in the theory of semisimple Hopf algebras $H$ over an algebraically closed field of characteristic 0 is the character ring of $H$. 


Let $V$ be a finite-dimensional left $H$-module and let $\rho_V : H \to \text{End}(V)$ be the corresponding representation. Then $\chi_V \in H^*$ is defined by

$$
\chi_V(h) := \text{Tr}(\rho_V(h))
$$

for any $h \in H$. Since $\text{Tr}(\rho_V(hh')) = \text{Tr}(\rho_V(h'h))$ for all $h, h' \in H$ it follows that $\chi_V$ is a cocommutative element of $H^*$.

If $V$ is an irreducible module we say that $\chi_V$ is an irreducible character. It is easily seen that

1. $\chi_{V \oplus W} = \chi_V + \chi_W$,
2. $\chi_{V \otimes W} = \chi_V \ast \chi_W$,
3. $\chi_V^* = S(\chi_V)$.

Define the character ring $R(H)$ of $H$ to be the $k$-span in $H^*$ of all the characters on $H$. Since $H$ is semisimple it follows that $R(H)$ is generated over $k$ by the finite set of its irreducible characters. Actually, $R(H)$ is the subalgebra of all cocommutative elements of $H^*$.

Let $t \in \int_H$ be such that $\varepsilon(t) = 1$. Define a form $\langle \cdot, \cdot \rangle$ on $R(H)$ by:

$$
\langle \varphi | \psi \rangle := \langle \varphi \ast S(\psi), t \rangle = \sum \langle \varphi, t_1 \rangle \langle S(\psi), t_2 \rangle
$$

for any characters $\varphi, \psi \in R(H)$.

As for characters of finite groups, there are orthogonality relations for irreducible characters via this form. Let $\{V_0, V_1, \ldots, V_m\}$ be a complete set of irreducible left $H$-modules, where $V_0$ is the trivial module. Let $n_i = \text{dim}(V_i)$ and let $\chi_i$ denote the character $\chi_{V_i}$. We have:

**Theorem 5.1.4 [131].** Let $H$ be a semisimple Hopf algebra over an algebraically closed field and let $\{\chi_0, \ldots, \chi_m\}$ be the set of irreducible characters of $H$. Then $\langle \chi_i | \chi_j \rangle = \delta_{ij}$.

A consequence of the above theorem is that $R(H)$ is a semisimple algebra.

Let $R_\mathbb{Z}(H) = \sum_i \mathbb{Z}\chi_i \subset R(H)$ where the $\{\chi_i\}$ are the irreducible characters on $H$. Since the $\{\chi_i\}$ are $\mathbb{Z}$-independent by orthogonality, $R_\mathbb{Z}(H)$ is a finite free $\mathbb{Z}$-module. In fact $R_\mathbb{Z}(H) \cong K_0(H)$, the Grothendieck ring of $H$.

**Theorem 5.1.5 [180].** Two semisimple Hopf algebras have isomorphic Grothendieck rings if and only if they are pseudo-twists of each other.

An important generalization from group theory is the class equation for semisimple Hopf algebras.
THEOREM 5.1.6 [114,250]. Let $H$ be a semisimple Hopf algebra over an algebraically close field of characteristic 0. Let $\{e_0, e_1, \ldots, e_m\}$ be a complete set of primitive orthogonal idempotents in $R(H)$, where $e_0$ is an integral for $H^*$. Then

$$\dim(H) = 1 + \sum_{i=1}^{m} \dim(e_i H^*)$$

and $\dim(e_i H^*)$ divides $\dim(H)$ for all $0 \leq i \leq m$.

Many results in the classification theory of semisimple Hopf algebras are due to the class equation. An immediate one is the Kac–Zhu theorem (Theorem 2.6.1). When $H = kG$, the theorem boils down to the usual class equation for finite groups. When $H = (kG)^*$ then $R(H) = H^* = kG$ and the class equation says that the dimension of an irreducible $G$-module divides the order of $G$. This is the classical theorem of Frobenius for finite groups that motivated Kaplansky’s 6th conjecture:

KAPLANSKY’S 6TH CONJECTURE [116]. Let $H$ be a semisimple Hopf algebra. Then the dimension of any irreducible $H$-module divides the dimension of $H$.

We say that $H$ is of Frobenius type if it satisfies this conjecture. Let the base field $k$ be algebraically closed of characteristic 0, then $H$ is of Frobenius type in the following cases:

1. If $(H, R)$ is quasitriangular, [78].
2. If $H$ is semisolvable, that is, $H$ has a normal series of Hopf subalgebras such that each Hopf quotient is either commutative or cocommutative, [173].
3. If $R(H)$ is central in $H^*$, [249].
4. If $H$ is cotriangular, [81]. This was proved using Theorem 5.1.7 below.

In fact, for the case of semisimple quasitriangular Hopf algebra it is proved that the dimension of any irreducible $D(H)$-module divides the dimension of $H$. The proof uses the theory of modular categories (the representation category of $D(H)$ is modular), and in particular the Verlinde formula, [242], applies. See [210,231] for later proofs in the quasitriangular case.

Other important results in this direction are that if $H$ has an irreducible module of dimension 2 then $H$ has even dimension, [179], and more generally that if $H$ has an even-dimensional irreducible module then $H$ has even dimension, [119].

Semisimple triangular Hopf algebras

The structure of triangular Hopf algebras is far from trivial, and yet is more tractable than that of general Hopf algebras, due to their proximity to groups and Lie algebras.

THEOREM 5.1.7 [77]. Any semisimple triangular Hopf algebra over an algebraically closed field $k$ of characteristic 0 is isomorphic to $(kG)^J$ for a unique (up to isomorphism) finite group $G$ and a unique (up to gauge equivalence) twist $J$. 

The proof of this theorem is based on a deep theorem of Deligne on Tannakian categories, [56]. The idea is that if \((H, R)\) is a semisimple triangular Hopf algebra then one can modify \(R\) to get an element \(\tilde{R}\) such that the category \(\text{Rep}(H, \tilde{R})\) is not only semisimple and symmetric but also has the property that the categorical dimensions of its objects are non-negative integers. Thus by the Deligne theorem, \(\text{Rep}(H, \tilde{R})\) is Tannakian; that is, it is equivalent to \(\text{Rep}(G)\) for a unique finite group \(G\), [58].

This theorem is the key step in the complete classification of triangular semisimple Hopf algebras described in the following theorem.

THEOREM 5.1.8 [80]. Triangular semisimple Hopf algebras of dimension \(N\) over \(k\) are in one to one correspondence with quadruples \((G, H, V, u)\), where \(G\) is a finite group of order \(N\), \(H < G\), \(V\) is an irreducible projective representation of \(H\) of dimension \(|H|^{1/2}\), and \(u \in G\) a central element of order \(\leq 2\).

By the previous theorem one needs to classify twists for a given finite group \(G\), up to gauge equivalence. It turns out that any twist \(J\) for \(G\) is gauge equivalent to a “minimal” twist coming from a subgroup \(H\) of \(G\). Finally, using Movshev’s theory, [174], one shows that equivalence classes of minimal twists for a finite group \(H\) are in bijection with isomorphism classes of irreducible projective representation of \(H\) of dimension \(|H|^{1/2}\). It is interesting to note that \(H\) is a central type group, so in particular solvable, [109].

For non-semisimple finite-dimensional triangular Hopf algebras \(H\) over an algebraically closed field \(k\) of characteristic 0 it is no longer true that the categorical dimensions of objects in \(\text{Rep}(H)\) are non-negative integers; so the Deligne theorem, [56], cannot be applied. Nevertheless using a recent theorem of Deligne, [57], it was proved in [84] that any finite-dimensional triangular Hopf algebra has the Chevalley property; namely, the semisimple part of \(H\) is itself a Hopf algebra. This leads to the complete and explicit classification of finite-dimensional triangular Hopf algebras over \(k\). As a consequence, for example, it is proved that in any finite-dimensional triangular Hopf algebra \(H\), \(u^2 = 1\) and hence \(S^4 = \text{id}\).

Results for special dimensions

Another conjecture of Kaplansky is the following:

KAPLANSKY’S 10TH CONJECTURE. For each integer \(n > 0\) there are only finitely many isomorphism classes of \(n\)-dimensional Hopf algebras.

A positive answer to this conjecture was given in the semisimple case, any characteristic:

THEOREM 5.1.9 [216]. Let \(k\) be an algebraically closed field. Then for each integer \(n > 0\) there are only finitely many isomorphism classes of \(n\)-dimensional semisimple and cosemisimple Hopf algebras.

Using three independent methods the conjecture was shown to be false in the non-semisimple case, [8,17,96] (though all these Hopf algebras are twists of each other, [149]).
In what follows we list results for special dimensions over an algebraically closed field of characteristic zero. Let \( p \neq q \) be prime numbers, then:

1. If \( \dim H = pq \) then \( H \) is either \( kG \) or \((kG)^*\) for some group \( G \), [77,98].
2. If \( \dim H = p^2 \) then \( H = kG \) for some finite group \( G \), [147].
3. Let \( p \neq 2 \) and \( \dim H = p^3 \). Then there exist exactly \( p + 8 \) isomorphism classes. Seven of these are of the form \( kG \) or \((kG)^*\) and the other \( p + 1 \) are all non-commutative, non-cocommutative and self dual, [146].
4. If \( \dim H = p^n \), then \( H \) is solvable, in particular \( H \) is of Frobenius type, [173].

It follows that the first possible dimension for a semisimple Hopf algebra to be neither commutative nor cocommutative is 8. An example of such a Hopf algebra was constructed already in 1966, [115].

Using the above results semisimple Hopf algebras of odd dimension less then 60 over an algebraically closed field of characteristic zero are all classified. Even-dimensional Hopf algebras are less known. For specific dimensions see the surveys mentioned in the beginning of this section.

It was proved in [85].

**Theorem 5.1.10.** Let \( H \) be a Hopf algebra over an algebraically closed field of characteristic 0 whose dimension is \( pq \), where \( p, q \) are prime and \( p < q < 2p + 2 \). Then \( H \) is semisimple, hence is either \( kG \) or \((kG)^*\) for some group \( G \).

**The exponent and the Schur indicator**

Motivated by other aspects of group theory two other notions were generalized to Hopf algebras: the exponent and the Schur indicator.

The classical notion of the exponent of a group is generalized in [79], motivated by [117, 118], in which the exponent of Hopf algebras whose antipode is involutive is studied. In fact, for such Hopf algebras the notion of exponent has existed for over 30 years.

**Definition 5.1.11** [79]. The exponent \( \exp(H) \) of \( H \) is the smallest positive integer \( n \) such that \( m_n \circ (\text{id} \otimes S^{-2} \otimes \cdots \otimes S^{-2n+2}) \circ \Delta_n = \varepsilon \cdot 1 \), where \( m_n, \Delta_n \) are the iterated product and coproduct.

If \( H \) is involutive (for example, \( H \) is semisimple and cosemisimple), then \( \exp(H) \) equals the smallest positive integer \( n \) so that \( m_n \circ \Delta_n = \varepsilon \cdot 1 \).

In [79] it is shown that \( \exp(H) \) equals the order of the Drinfeld element \( u \) of the quantum double \( D(H) \), and the order of \( R^+ R \), where \( R \) is the universal \( R \)-matrix of \( D(H) \). This was motivated by a theorem in conformal field theory, [237].

In [118] it was conjectured that if \( H \) is semisimple and cosemisimple then \( \exp(H) \) is always finite and divides \( \dim(H) \).

**Theorem 5.1.12** [79]. For a semisimple and cosemisimple Hopf algebra \( H \), \( \exp(H) \) is finite and divides \( \dim(H)^3 \).
In [119] it is proved that if 2 divides \( \dim(H) \) then 2 divides \( \exp(H) \). Whether this is true for any odd prime \( p \) is still an open question.

For non-semisimple finite-dimensional Hopf algebras the exponent is usually infinite. However, it was proved in [82] that the order of unipotency of \( u \) is always finite, since all the eigenvalues of \( u \) are roots of unity. This order of unipotency of \( u \) was termed the quasi-exponent of \( H \), and it reduces to \( \exp(H) \) when \( H \) is semisimple. In [82] equivalent definitions of \( \text{qexp}(H) \), generalizing the ones in the exponent case, are given and it is proved that \( \text{qexp}(H) \) is an invariant of the tensor category \( \text{Rep}(H) \).

The theory of quasi-exponents was applied to study the group of grouplike elements of twisted quantum groups at roots of unity, [82].

Another generalization of group theory is the Schur indicator \( \nu \).

**Definition 5.1.13** [137]. Let \( H \) be a semisimple Hopf algebra over an algebraically closed field of characteristic 0, and \( t \in \int_H \) such that \( \langle \varepsilon, t \rangle = 1 \). For any irreducible character \( \chi \) define the Schur indicator by

\[
\nu(\chi) = \sum \chi(t_{12}).
\]

It is proved:

**Theorem 5.1.14** [137]. The Schur indicator satisfies the following:
1. \( \nu(\chi) \in \{0, 1, -1\} \) for all \( \chi \in \text{Irr}(H) \).
2. \( \nu(\chi) \neq 0 \) if and only if \( V_\chi \cong V_\chi^* \).

Moreover, \( \nu(\chi) = 1 \) (resp. \(-1\)) if and only if \( V_\chi \) admits a symmetric (resp. skew symmetric) non-degenerate bilinear \( H \)-invariant form.

Theorems 5.1.12 and 5.1.14 found an interesting applications in [119] where it is proved that a semisimple Hopf algebra over \( \mathbb{C} \) with a non-trivial self-dual irreducible representation or with an even-dimensional irreducible representation, must have even dimension.

**Hopf algebras with positive bases**

A finite-dimensional Hopf algebra \( H \) over \( \mathbb{C} \) is said to have a positive base if it has a linear basis with respect to which all the structure constants are positive. For example, a bicrossproduct Hopf algebra arising from a finite group \( G \) and an exact factorization \( G = G_+G_- \) of \( G \) is a Hopf algebra with a positive base (e.g., \( \mathbb{C}[G] \) and \( D(G) \)). In fact, if \( H \) is a Hopf algebra with a positive base then \( H \) is of this form, [141]. Note that in particular \( H \) is semisimple.

**5.2. Pointed Hopf algebras**

The reader is referred to [4,9,11] for more results, explanations and details.
Pointed Hopf algebras are of special interest as many important examples of Hopf algebras are such. In particular, group algebras, enveloping algebras of Lie algebras, quantized enveloping algebras and many quantum groups are all pointed.

Moreover, by Remark 1.1.11, every cocommutative coalgebra $C$ over an algebraically closed field $k$ is pointed. Group algebras and enveloping algebras of Lie algebras are clearly cocommutative, but more is true. They serve as the building blocks in one of the first fundamental theorems about cocommutative Hopf algebras. This theorem was proved independently by Cartier, [36] and Kostant, unpublished (see [221, Preface]).

**Theorem 5.2.1.** A cocommutative Hopf algebra over an algebraically closed field of characteristic 0 is a smash product of the group algebra $kG$ and the enveloping algebra $U(g)$, where $G$ is the group of grouplike elements of $H$ and $g$ is the Lie algebra of primitive elements of $H$.

The converse is obviously true. Furthermore, the converse can be generalized to any Hopf algebra $H$ as follows: If $H$ is generated as an algebra by $G(H)$ and the skew-primitive elements of $H$ (see Definition 1.1.15) then $H$ is pointed. In view of Theorem 5.2.1 and the above generalized converse, it was conjectured, [9], that all finite-dimensional pointed Hopf algebras over an algebraically closed field of characteristic 0 are generated as algebras by grouplike and skew-primitive elements. The conjecture is false in the infinite-dimensional case, [11, Example 3.6].

Finite-dimensional pointed Hopf algebras were characterized in the following cases:

1. For a prime number $p > 2$. The only pointed Hopf algebras of dimension $p^2$ are the Taft algebras. This was already shown in [177].

2. For $p = 2$ there is exactly one isomorphism class of dimension $2^n$, [177,31]. These pointed Hopf algebras are generalizations of Sweedler's four-dimensional Hopf algebra $H_4$ and were investigated by, e.g., [17,185,186].

3. Hopf algebras of dimension $p^3$, $p^4$, $p^5$ and $pq^2$, $q$ another prime, are fully characterized, [8,6,30,217,10,100].

4. Partial results are known for pointed Hopf algebras with some special properties.

An essential tool in the study of a pointed coalgebra $C$ is its coradical filtration $C_n$. This structure, given in the fundamental theorem of Taft and Wilson, allows induction starting from $C_0 = kG(C)$. The following version is somewhat stronger than the original one:

**Theorem 5.2.2** [223]. Let $C$ be a pointed coalgebra. Then for any $c \in C_n$, $n \geq 1$, we have

$$c = \sum_{g, h \in G(C)} c_{g, h}, \quad \text{where } \Delta(c_{g, h}) = c_{g, h} \otimes g + h \otimes c_{g, h} + w$$

for some $w \in C_{n-1} \otimes C_{n-1}$.

If $H$ is a pointed Hopf algebra with coradical filtration $H_n$ then $H_0 = kG(H)$ and

$$H_1 = kG(H) \oplus \left( \bigoplus_{\sigma, \tau \in G(H)} P_{\sigma, \tau}(H) \right).$$
Moreover, the coradical filtration is a Hopf algebra filtration, that is $H_i H_j \subset H_{i+j}$ and $S(H_i) \subset H_i$ for all $i, j \geq 0$.

The coradical filtration is the starting point for the lifting method which is a powerful tool in the structure theory of pointed Hopf algebras. This method was used in a series of works of Andruskiewitsch and Schneider (for more details see the references at the beginning of this section). We give a brief overview of this method and the main results.

Let $H$ be a pointed Hopf algebra, let $\{H_n \mid n \geq 0\}$ denote the coradical filtration of $H$ and set $H_{-1} = k$. Let

$$\text{gr} \ H = \bigoplus_{n \geq 0} \text{gr} \ H(n),$$

where $\text{gr} \ H(n) = H_n/H_{n-1}$ for all $n \geq 0$. Since the coradical of $H$ is a Hopf subalgebra it follows by [166, 5.2.8] that $\text{gr} \ H$ is a graded Hopf algebra. Now, there is a Hopf algebra projection $\pi : \text{gr} \ H \to \text{gr} \ H(0) = kG(H)$ and a Hopf algebra injection $i : \text{gr} \ H(0) \to \text{gr} \ H$. By Theorem 4.4.6 this implies that we have a biproduct

$$\text{gr} \ H \cong R \# kG(H),$$

where $R = \{x \in \text{gr} \ H \mid (\text{id} \otimes \pi) \Delta(x) = x \otimes 1\}$ is a Hopf algebra in the category $\mathcal{YD}_{kG(H)}$.

The structure of $R$ is the key to understanding the structure of the original Hopf algebra $H$. The vector space $V$ of all primitive elements of $R$ is also an object in $\mathcal{YD}_{kG(H)}$ and thus has a braiding

$$c : V \otimes V \to V \otimes V.$$

This braiding is called the infinitesimal braiding of $H$.

The subalgebra of $R$ generated by $V$ turns out to be $B(V)$, the so-called Nichols algebra of $V$ (see Example 4.4.3).

Given a group $G$ and a vector space $(V, c) \in \mathcal{YD}_{kG}$ the first problem is to study the structure of the Nichols algebras $B(V)$ and to determine when $B(V)$ is finite-dimensional. The second problem is to determine all pointed Hopf algebras $H$ such that $\text{gr} \ H \cong B(V) \# kG$ (the lifting problem).

The best results for this method were achieved in the case when $G$ is Abelian and the braiding is of Cartan type. That is, for a certain basis $\{v_1, \ldots, v_n\}$ of $V$ we have

$$c(v_i \otimes v_j) = q_{ij}(v_j \otimes v_i),$$

where $q_{ij}q_{ji} = q^{d_{aij}}$, $q \neq 0$ and $(a_{ij})$ is a generalized symmetrizable Cartan matrix with positive integers $\{d_1, \ldots, d_n\}$ so that $d_i a_{ij} = d_j a_{ji}$.

The Cartan type of the pointed Hopf algebra is invariant under twisting.

Pointed Hopf algebras $H$ (finite- or infinite-dimensional) such that $G(H)$ is Abelian and the braiding on $V$ is of Cartan type (plus some additional requirements) were fully characterized by Andruskiewitsch and Schneider. They are generalizations of quantized Lie algebras.
The order of $S^2$ for pointed Hopf algebras

By Theorem 2.4.1, the order of $S^2$ divides $2 \dim(H)$, when $H$ is a finite-dimensional Hopf algebra. Using the coradical filtration the following was proved:

**Theorem 5.2.3** [197]. Let $H$ be a finite-dimensional pointed Hopf algebra over $\mathbb{C}$. Then $|S^2|$ divides $\dim(H)/|G(H)|$.

**Theorem 5.2.4** [82]. Let $H$ be a finite-dimensional pointed Hopf algebra over $\mathbb{C}$. Then $|S^2|$ divides $\exp(G(H))$ (and hence $\dim(H)$).

This theorem was used to prove

**Theorem 5.2.5** [82]. Let $H$ be a finite-dimensional pointed Hopf algebra over $\mathbb{C}$. Then $q\exp(H) = \exp(G(H))$.

Hopf algebras of rooted trees

The $\mathbb{Z}$-graded Hopf algebra $A$ of rooted trees was introduced in [102], in connection with numerical algorithms for ordinary differential equations; this Hopf algebra is cocommutative but not commutative.

On the other hand, Kreimer, [125], has discovered the interesting fact that the process of renormalization in quantum field theory may be described by means of Hopf algebras related to operads of rooted trees. The Hopf algebra $L$ of decorated rooted trees described by A. Connes and D. Kreimer, [53], arises from the combinatorics of perturbative renormalization, and is related to cyclic cohomology and non-commutative geometry. It is $\mathbb{Z}$-graded and commutative but not cocommutative.

In [184] the author proves that the Hopf algebras $A$ and $L$ are dual to one another. Moreover, he identifies a certain linear operator on $A$ as a dual operator to $L$.

Related structures

There are several algebraic structures generalizing the notion of a Hopf algebra, which are very important and interesting in their own right and have been studied extensively. For example, quasi-Hopf algebras, weak Hopf algebras, Hopf algebroids, infinitesimal Hopf algebras, [2], multiplier Hopf algebras, [238], and dendriform Hopf algebras, [139], are such algebraic structures. In the following we briefly discuss some of them.

(1) **Quasi-Hopf algebras.** The notion of a quasi-Hopf algebra $H$, due to V. Drinfeld, generalizes the notion of a Hopf algebra in that the associativity constraint $(U \otimes V) \otimes W \simeq U \otimes (V \otimes W)$ in the tensor category $\text{Rep}(H)$ can be non-trivial, [28,29]. More precisely, a quasi-Hopf algebra is a unital associative algebra with comultiplication $\Delta$, counit $\varepsilon$, and antipode $S$ satisfying some axioms, where the main difference is that $\Delta$ is only coassociative up to conjugation by an invertible element in $H \otimes H \otimes H$, [73]. See also [215].
Hopf algebras

The importance of these algebras lies in the fact that their representation category is tensor (usually with a non-standard associativity). For example, they produce solutions to the Knizhnik–Zamolodchikov equation in quantum field theory.

(2) Weak Hopf algebras. A weak Hopf algebra or a quantum groupoid is a unital associative algebra with comultiplication $\Delta$, counit $\varepsilon$ and antipode $S$, satisfying some axioms. The main difference between Hopf algebras and weak Hopf algebras is that in the latter $\Delta$ need not map the identity in $H$ to the identity in $H \otimes H$. This relaxation of the axioms of Hopf algebras is very significant. For example, while not every finite (fusion) category is equivalent to Rep($H$) for some finite-dimensional (semisimple) Hopf algebra $H$, it is known that it is equivalent to Rep($H$) for some finite-dimensional (semisimple) weak Hopf algebra. Thus the theory of weak Hopf algebras is very useful in the study of finite (fusion) categories. See [27,86,87,182].

(3) Multiplier Hopf algebras. A multiplier Hopf algebra is an algebra $A$ with or without identity and a homomorphism $\Delta$ from $A$ to the multiplier algebra of $A \otimes A$ satisfying certain axioms (such as a form of coassociativity). If $A$ has an identity then $A$ is a usual Hopf algebra.

Here, as for Hopf algebras, the motivating example arises from groups. Consider the algebra $A$ of all complex valued functions on a group $G$ and define $\Delta$, $\varepsilon$, $S$ as for Fun($G$), $G$ a finite group (see the introduction). If $G$ is infinite then $\Delta(f)$ does not necessarily belong to $A \otimes A$ and thus $A$ is not a Hopf algebra. However, if $A$ is the algebra of (continuous) functions with compact support on a discrete group $G$ then $\Delta(f)(g \otimes 1), (1 \otimes f)\Delta(g) \in A \otimes A$ for all $f, g \in A$. This fact implies that $A$ is a multiplier Hopf algebra.

Many results for Hopf algebras can be generalized for multiplier Hopf algebras by using similar methods (see, e.g., [238,70,69]).

References

Hopf algebras

[104] T. Hayashi, Quantum deformations of classical groups, Publ. RIMS Kyoto Univ. 28 (1992), 57–81.
Hopf algebras


This page intentionally left blank
Difference Algebra

Alexander B. Levin

Department of Mathematics, The Catholic University of America, Washington, DC 20064, USA
E-mail: Levin@cua.edu

Contents
1. Introduction .......................................................... 243
2. Basic concepts of difference algebra ................................. 244
  2.1. Difference and inversive difference rings ...................... 244
  2.2. Rings of difference and inversive difference polynomials. Algebraic difference equations ........ 250
  2.3. Autoreduced sets of difference and inversive difference polynomials. Characteristic sets .......... 254
  2.4. Perfect difference ideals. Ritt difference rings ............... 259
  2.5. Varieties of difference polynomials ........................... 263
3. Difference modules .................................................. 268
  3.1. Ring of difference operators. Difference modules ............ 268
  3.2. Type and dimension of difference vector spaces ............... 271
  3.3. Reducible difference modules. σ*-dimension polynomials and their invariants .................. 271
  3.4. Reduction in a free difference vector space. Characteristic sets and multivariable dimension poly-
      nomials .............................................................. 276
4. Difference field extensions ......................................... 279
  4.1. Transcendental dependence. Difference transcendental bases and difference transcendental degree ...... 279
  4.2. Dimension polynomials of difference and inversive difference field extensions ...................... 281
  4.3. Limit degree. Finitely generated difference and inversive difference field extensions ................ 287
  4.4. Difference kernels. Realizations ................................ 291
  4.5. Ordinary difference polynomials. Existence theorem ........ 296
  4.6. Compatibility of difference field extensions. Specializations ............................................. 304
  4.7. Isomorphisms of difference fields. Monadicity .................. 310
  4.8. Difference valuation rings and extensions of difference specializations .............................. 312
5. Difference Galois theory .............................................. 314
  5.1. Algebraic difference field extensions. Galois correspondence .............................................. 314
  5.2. Picard-Vessiot theory of linear homogeneous difference equations ....................................... 315
  5.3. Picard-Vessiot rings and the Galois theory of difference equations ....................................... 323
References ................................................................. 330

HANDBOOK OF ALGEBRA, VOL. 4
Edited by M. Hazewinkel
© 2006 Elsevier B.V. All rights reserved
1. Introduction

Difference algebra as a separate area of mathematics was born in the 1930s when J.F. Ritt (1893–1951) developed the algebraic approach to the study of systems of difference equations over function fields. In a series of papers published during the decade from 1929 to 1939, Ritt worked out the foundations of both differential and difference algebra, the theories of abstract algebraic structures with operators that reflect the algebraic properties of derivatives and shifts of arguments of analytic functions, respectively. One can say that differential and difference algebra grew out of the study of algebraic differential and difference equations with coefficients from function fields in much the same way as the classical algebraic geometry arose from the study of polynomial equations with numerical coefficients.

Ritt’s research in differential algebra was continued and extended by H. Raudenbush, H. Levi, A. Seidenberg, A. Rosenfeld, P. Cassidy, J. Johnson, W. Keigher, S. Morrison, W. Sit and many other mathematicians, but the most important role in this area was played by E. Kolchin who recast the whole subject in the style of modern algebraic geometry with the additional presence of derivation operators. In particular, E. Kolchin developed the contemporary theory of differential fields and created differential Galois theory where finite-dimensional algebraic groups played the same role as finite groups play in the theory of algebraic equations. Kolchin’s monograph, [86], is the most deep and complete book on the subject, it contains a lot of ideas that determined the main directions of research in differential algebra for the last thirty years.

The rate of development of difference algebra after Ritt’s pioneering work and works by F. Herzog, H. Raudenbush and W. Strodt published in the 1930s (see [68,133,134,136,137,140], and [141]) was much slower than the rate of expansion of its differential counterpart. The situation began to change in the 1950s due to R.M. Cohn whose works, [19–32], not only raised difference algebra to a level comparable with the level of development of differential algebra, but also clarified why many ideas of differential algebra cannot be realized in the difference case and a number of methods and results of difference algebra cannot have differential analogs. R.M. Cohn’s book, [29], up to now remains the only fundamental monograph on difference algebra. Since the 60s various problems of difference algebra were studied by A. Babbitt, [2], I. Balaba, [3–5], I. Bentsen, [7], R.M. Cohn, [33–37], P. Evanovich, [49,50], C. Franke, [60–62], B. Greenspan, [63], P. Hendrics, [65,66], M. Kondrateva, [87–91], B. Lando, [96,97], A. Levin, [87,88] and [99–115], A. Mikhalev, [87,88,110–115] and [117–119], E. Pankratev, [87–91,117–119] and [124–127], C. Praagman, [130,131], and some other mathematicians. Difference Galois theory originated in the 60s and 70s in works by C. Franke, [56–59], A. Bialynicki-Birula, [8], H.F. Kreimer, [93–95], R. Infante, [70–75], and E. Pankratev, [124,125], was further actively developed in the last ten years by M. van der Put, M. Singer, and P.A. Hendrix, [64–67,142]. The current state of the theory is fully reflected in the recent monograph by M. van der Put and M. Singer, [142].

Since the 70s difference algebra has been enriched by new methods and ideas from the dimension theory of differential rings (see [41–43,78–82,84,85], and [139]), the theory of Gröbner bases which originated in [13] and computer algebra. A number of deep results were obtained in the model theory of difference fields developed by E. Hrushovski and
Z. Chatzidakis, [15,16] and [69] (see also [46,128] and [129]). Nowadays, difference algebra appears as a rich theory with its own methods and with applications to the study of system of equations in finite differences, functional equations, differential equations with delay, algebraic structures with operators, group and semigroup rings. A number of interesting applications of difference algebra in the theory of discrete-time non-linear systems can be found in the works by M. Fliess, [51–55], E. Aranda-Bricaire, U. Kotta and C. Moog, [1], and some other authors.

In what follows, \( \mathbb{Z}, \mathbb{N}, \mathbb{Q}, \mathbb{R}, \text{ and } \mathbb{C} \) denote the sets of integers, non-negative integers, rational numbers, real numbers, and complex numbers respectively. \( \mathbb{Q}[t] \) will denote the set of all polynomials in one variable \( t \) with rational coefficients. By a ring we always mean an associative ring with a unity. Every ring homomorphism is unitary (maps unity onto unity), every subring of a ring contains the unity of the ring. An ideal \( I \) of a ring \( R \) is said to be proper if \( I \neq R \). Unless otherwise indicated, by a module over a ring \( A \) we mean an \( A \)-module. Every module over a ring is unitary and every algebra over a commutative ring is also unitary.

2. Basic concepts of difference algebra

2.1. Difference and inversive difference rings

A difference ring is a commutative ring \( R \) together with a finite set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \) of mutually commuting injective endomorphisms of \( R \) into itself. The set \( \sigma \) is called the basic set of the difference ring \( R \), and the endomorphisms \( \alpha_1, \ldots, \alpha_n \) are called translations. In other words, a difference ring \( R \) with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \) is a commutative ring possessing \( n \) additional unitary operations \( \alpha_i : a \mapsto \alpha_i(a) \) such that

\[
\begin{align*}
\alpha_i(a) &= 0 \quad \text{if and only if} \quad a = 0, \\
\alpha_i(a + b) &= \alpha_i(a) + \alpha_i(b), \\
\alpha_i(ab) &= \alpha_i(a)\alpha_i(b), \\
\alpha_i(1) &= 1, \quad \text{and} \\
\alpha_i(\alpha_j(a)) &= \alpha_j(\alpha_i(a))
\end{align*}
\]

for any \( a \in R, 1 \leq i, j \leq n \). (Formally speaking, a difference ring is an \( (n + 1) \)-tuple \( (R, \alpha_1, \ldots, \alpha_n) \) where \( R \) is a ring and \( \alpha_1, \ldots, \alpha_n \) are mutually commuting injective endomorphisms of \( R \). Unless the notation is inconvenient or ambiguous, we always write \( R \) for \( (R, \alpha_1, \ldots, \alpha_n) \).

If \( \alpha_1, \ldots, \alpha_n \) are automorphisms of \( R \), we say that \( R \) is an inversive difference ring with the basic set \( \sigma \).

In what follows, a difference ring \( R \) with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \) will be also called a \( \sigma \)-ring. If \( n = 1 \), \( R \) is said to be an ordinary difference (\( \sigma \)-) ring, if \( \text{Card} \sigma > 1 \), the difference ring \( R \) is called partial.
If \( R \) is an inversive difference ring with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \), then the set \( \{\alpha_1, \ldots, \alpha_n, \alpha_1^{-1}, \ldots, \alpha_n^{-1}\} \) is denoted by \( \sigma^* \) and \( R \) is also called a \( \sigma^* \)-ring.

If a difference ring with a basic set \( \sigma \) is a field, it is called a difference (or \( \sigma \)-) field. An inversive difference field with a basic set \( \sigma \) is also called a \( \sigma^* \)-field.

Let \( R \) be a difference ring with basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \) and \( R_0 \) a subring of \( R \) such that \( \alpha(R_0) \subseteq R_0 \) for any \( \alpha \in \sigma \). Then \( R_0 \) is called a difference (or \( \sigma \)-) subring of \( R \). If \( R \) is a difference ring \( R_0 \) is said to be a difference (or \( \sigma \)-) overring of \( R_0 \). In this case the restriction of an endomorphism \( \alpha_i \) on \( R_0 \) is denoted by same symbol \( \alpha_i \). If the \( \sigma \)-ring \( R \) is inversive and \( R_0 \) a \( \sigma \)-subring of \( R \) such that \( \alpha^{-1}(R_0) \subseteq R_0 \) for any \( \alpha \in \sigma \), then \( R_0 \) is said to be a \( \sigma^* \)-subring of \( R \). If \( R \) is a difference \( \sigma \)-field and \( R_0 \) a subfield of \( R \) such that \( \alpha(a) \in R_0 \) for any \( a \in R_0, \alpha \in \sigma \), then \( R_0 \) is said to be a difference (or \( \sigma \)-) subfield of \( R \); \( R \) is reflexive if it is a \( \sigma^* \)-field extension.

If \( \alpha(J) \subseteq R_0 \) for any \( \alpha \in \sigma \), \( R_0 \) is said to be a difference (or \( \sigma \)-) overfield of \( R_0 \). Then \( R_0 \) is called a \( \sigma^* \)-subfield of \( R \) while \( R_0 \) is called a \( \sigma^* \)-field extension or a \( \sigma \)-overfield of \( R_0 \). We also say that we have a \( \sigma \)-field extension \( R/R_0 \). If \( R_0 \subseteq R_1 \subseteq R \) is a chain of \( \sigma \)-field extensions, we say that \( R_1/R_0 \) is a \( \sigma \)- (respectively, \( \sigma^* \))-field subextension of \( R/R_0 \).

If \( R \) is a difference ring with a basic set \( \sigma \) and \( J \) is an ideal of the ring \( R \) such that \( \alpha(J) \subseteq J \) for any \( \alpha \in \sigma \), then \( J \) is called a difference (or \( \sigma \)-) ideal of \( R \), a prime (maximal) ideal \( P \) of \( R \) is closed with respect to \( \sigma \) (that is, \( \alpha(P) \subseteq P \) for any \( \alpha \in \sigma \), \( R \) is called an \( \sigma \)-ideal). A \( \sigma \)-ideal \( J \) of a \( \sigma \)-ring \( R \) is called reflexive (or a \( \sigma^* \)-ideal) if for any translation \( \alpha \), the inclusion \( \alpha(a) \in J \) (\( a \in R \)) implies \( a \in J \). Clearly, if \( R \) is an inversive \( \sigma \)-ring, then a \( \sigma \)-ideal \( J \) of \( R \) is reflexive if and only if it is closed under all automorphisms from the set \( \sigma^* \). A prime (maximal) reflexive \( \sigma \)-ideal of a \( \sigma \)-ring \( R \) is also called a prime (respectively, maximal) \( \sigma^* \)-ideal of \( R \).

**Examples 2.1.1.**

1. Any commutative ring can be treated as a difference (or inversive difference) ring with a basic set \( \sigma \) consisting of one or several identity automorphisms.
2. Let \( z_0 \in \mathbb{C} \) and let \( U \) be a region of the complex plane such that \( z + z_0 \in U \) whenever \( z \in U \) (e.g., \( U = \{z \in \mathbb{C} \mid (\text{Re} z)(\text{Re} z_0) \geq 0\} \)). Furthermore, let \( M_U \) denote the field of all functions of one complex variable meromorphic in \( U \). Then \( M_U \) can be treated as an ordinary difference field whose basic set consists of one translation \( \alpha \) such that \( \alpha(f(z)) = f(z + z_0) \) for any function \( f(z) \in M_U \).
3. Let \( z_0 \in \mathbb{C} \) and let \( V \) be a region of the complex plane such that \( z_0 + z \in V \) whenever \( z \in V \) (e.g., \( |z_0| \leq 1 \) and \( V = \{z \in \mathbb{C} \mid |z| \leq r\} \) for some positive real number \( r \)). Then the field of all functions of one complex variable meromorphic in the region \( V \) can be considered as an ordinary difference field with a translation \( \beta \) such that \( \beta(f(z)) = f(z_0 + z) \) for any function \( f(z) \in M_V \). Clearly, \( M_V \) is inversive if and only if \( \frac{z_0}{z} \in V \) for any \( z \in V \).
4. Let \( A \) be a ring of functions of \( n \) real variables continuous on \( \mathbb{R}^n \). Let us fix some real numbers \( h_1, \ldots, h_n \) and consider the mutually commuting injective endomorphisms
\(\alpha_1, \ldots, \alpha_n\) of \(A\) given by \((\alpha_i f)(x_1, \ldots, x_n) = f(x_1, \ldots, x_i-1, x_i + h_i, x_{i+1}, \ldots, x_n)\) \((i = 1, \ldots, n)\). Then \(A\) can be treated as a difference ring with basic set \(\sigma = \{\alpha_1, \ldots, \alpha_n\}\). This difference ring is denoted by \(A_0(h_1, \ldots, h_n)\).

Similarly, one can introduce the difference structure on the ring \(C^p(\mathbb{R}^n)\) of all functions of \(n\) real variables that are continuous on \(\mathbb{R}^n\) together with all their partial derivatives up to the order \(p\) \((p \in \mathbb{N} \text{ or } p = +\infty)\). It is easy to see that \(C^p(\mathbb{R}^n)\) can be considered as a difference ring with the basic set \(\sigma = \{\alpha_1, \ldots, \alpha_n\}\) described above. This difference ring is denoted by \(A_p(h_1, \ldots, h_n)\). It is clear that \(A_p(h_1, \ldots, h_n)\) is a \(\sigma\)-subring of the \(\sigma\)-ring \(A_0(h_1, \ldots, h_n)\) whenever \(p > q\). The difference rings \(A_p(h_1, \ldots, h_n)\) often arise in connection with equations in finite differences when the \(i\)-th partial finite difference \(\Delta_i f(x_1, \ldots, x_n) = f(x_1, \ldots, x_i-1, x_i + h_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_n)\) of a function \(f(x_1, \ldots, x_n) \in C^p(\mathbb{R}^n)\) is written as \(\Delta_i f = (\alpha_i - 1)f\).

A number of interesting examples of difference rings can be found in [32, Chapter 1], [88, Section 3.3], and [142, Chapter 1].

Let \(R\) be a difference ring with a basic set \(\sigma\). An element \(c \in R\) is said to be a constant if \(\sigma(a) = a\) for any \(a \in \sigma\). Clearly, the set of constants of the ring \(R\) is a \(\sigma\)-subring of \(R\) (it is a \(\sigma^*\)-subring of \(R\), if the difference ring \(R\) is inversive). This subring is called the ring of constants of \(R\); it is denoted by \(C_R\).

If \(R\) is a difference ring with a basic set \(\sigma = \{\alpha_1, \ldots, \alpha_n\}\), then \(T_\sigma\) will denote the free commutative semigroup with identity generated by \(\alpha_1, \ldots, \alpha_n\). Elements of \(T_\sigma\) will be written in the multiplicative form \(\alpha_1^{k_1} \cdots \alpha_n^{k_n} \ (k_1, \ldots, k_n \in \mathbb{N})\) and considered as endomorphisms of \(R\). If the \(\sigma\)-ring \(R\) is inversive, then \(\Gamma_\sigma\) will denote the free commutative group generated by the set \(\sigma\). It is clear that elements of the group \(\Gamma_\sigma\) (written in the multiplicative form \(\alpha_1^{i_1} \cdots \alpha_n^{i_n}\) where \(i_1, \ldots, i_n \in \mathbb{Z}\)) act on \(R\) as automorphisms and \(T_\sigma\) is a subsemigroup of \(\Gamma_\sigma\).

For any \(a \in R\) and for any \(\tau \in T_\sigma\), the element \(\tau(a)\) is called a transform of \(a\). If the \(\sigma\)-ring \(R\) is inversive, then an element \(y(a) (a \in R, \gamma \in \Gamma_\sigma)\) is also called a transform of \(a\).

If \(J\) is a \(\sigma\)-ideal of a \(\sigma\)-ring \(R\), then \(J^* = \{a \in R \mid \tau(a) \in J\text{ for some }\tau \in T_\sigma\}\) is a reflexive \(\sigma\)-ideal of \(R\) contained in any reflexive \(\sigma\)-ideal of \(R\) containing \(J\). The ideal \(J^*\) is called a reflexive closure of the \(\sigma\)-ideal \(J\).

A difference \((\sigma-)\) ring \(R\) is called simple if the only \(\sigma\)-ideals of \(R\) are \((0)\) and \(R\). In this case the ring of constants \(C_R\) is a field.

Let \(R\) be a difference ring with a basic set \(\sigma\) and \(S \subseteq R\). Then the intersection of all \(\sigma\)-ideals of \(R\) containing \(S\) is denoted by \([S]\). Clearly, \([S]\) is the smallest \(\sigma\)-ideal of \(R\) containing \(S\); as an ideal, it is generated by the set \(T_\sigma S = \{\tau(a) \mid \tau \in T_\sigma, a \in S\}\). If \(J = [S]\), we say that the \(\sigma\)-ideal \(J\) is generated by the set \(S\) called a set of \(\sigma\)-generators of \(J\). If \(S\) is finite, \(S = \{a_1, \ldots, a_k\}\), we write \([a_1, \ldots, a_k]\) and say that \(J\) is a finitely generated \(\sigma\)-ideal of the \(\sigma\)-ring \(R\). (In this case elements \(a_1, \ldots, a_k\) are said to be \(\sigma\)-generators of \(J\).)

If \(R\) is an inversive difference \((\sigma-)\) ring and \(S \subseteq R\), then the inverse closure of the \(\sigma\)-ideal \([S]\) is denoted by \([S]^*\). It is easy to see that \([S]^*\) is the smallest \(\sigma^*\)-ideal of \(R\) containing \(S\); as an ideal, it is generated by the set \(\Gamma_\sigma S = \{\gamma(a) \mid \gamma \in \Gamma_\sigma, a \in S\}\). If \(S\) is finite, \(S = \{a_1, \ldots, a_k\}\), we write \([a_1, \ldots, a_k]^*\) for \(I = [S]^*\) and say that \(I\) is a finitely generated \(\sigma^*\)-ideal of \(R\). (In this case, the elements \(a_1, \ldots, a_k\) are said to be \(\sigma^*\)-generators of \(I\).)
Let $R$ be a difference ring with a basic set $\sigma$, $R_0$ a $\sigma$-subring of $R$ and $B \subseteq R$. The intersection of all $\sigma$-subrings of $R$ containing $R_0$ and $B$ is called the $\sigma$-subring of $R$ generated by the set $B$ over $R_0$, it is denoted by $R_0\langle B \rangle$. (As a ring, $R_0\langle B \rangle$ coincides with the ring $R_0[\{\tau(b) \mid b \in B, \tau \in T_\sigma]\}$ obtained by adjoining the set $\{\tau(b) \mid b \in B, \tau \in T_\sigma\}$ to the ring $R_0$). The set $B$ is said to be the set of $\sigma$-generators of the $\sigma$-ring $R_0\langle B \rangle$ over $R_0$. If this set is finite, $B = \{b_1, \ldots, b_k\}$, we say that $R' = R_0\langle B \rangle$ is a finitely generated difference (or $\sigma$-) ring extension (or overring) of $R_0$ and write $R' = R_0\langle b_1, \ldots, b_k \rangle$. If $R$ is a $\sigma$-field, $R_0$ a $\sigma$-subfield of $R$ and $B \subseteq R$, then the intersection of all $\sigma$-subfields of $R$ containing $R_0$ and $B$ is denoted by $R_0\langle B \rangle$ (or $R_0\langle b_1, \ldots, b_k \rangle$ if $B = \{b_1, \ldots, b_k\}$ is a finite set). This is the smallest $\sigma$-subfield of $R$ containing $R_0$ and $B$; it coincides with the field $R_0\langle \tau(b) \mid b \in B, \tau \in T_\sigma \rangle$. The set $B$ is called the set of $\sigma$-generators of the $\sigma$-field $R_0\langle B \rangle$ over $R_0$.

Let $R$ be an inversive difference ring with a basic set $\sigma$, $R_0$ a $\sigma^*$-subring of $R$ and $B \subseteq R$. Then the intersection of all $\sigma^*$-subrings of $R$ containing $R_0$ and $B$ is the smallest $\sigma^*$-subring of $R$ containing $R_0$ and $B$. This ring coincides with the ring $R_0[\{\gamma(b) \mid b \in B, \gamma \in \Gamma_\sigma\}]$; it is denoted by $R_0\langle B \rangle^*$. The set $B$ is said to be a set of $\sigma^*$-generators of $R_0\langle B \rangle^*$ over $R_0$. If $B = \{b_1, \ldots, b_k\}$ is a finite set, we say that $B = R_0\langle b_1, \ldots, b_k \rangle$ is a finitely generated inversive difference (or $\sigma^*$-) ring extension (or overring) of $R$ and write $S = R_0\langle b_1, \ldots, b_k \rangle^*$.

If $R$ is a $\sigma^*$-field, $R_0$ a $\sigma^*$-subfield of $R$ and $B \subseteq R$, then the intersection of all $\sigma^*$-subfields of $R$ containing $R_0$ and $B$ is denoted by $R_0\langle B \rangle^*$. This is the smallest $\sigma^*$-subfield of $R$ containing $R_0$ and $B$; it coincides with the field $R_0[\{\gamma(b) \mid b \in B, \gamma \in \Gamma_\sigma\}]$. The set $B$ is called the set of $\sigma^*$-generators of the $\sigma^*$-field extension $R_0\langle B \rangle^*$ over $R_0$. If $B$ is finite, $B = \{b_1, \ldots, b_k\}$, we write $R_0\langle b_1, \ldots, b_k \rangle^*$ for $R_0\langle B \rangle^*$.

In what follows we shall often consider two or more difference rings $R_1, \ldots, R_p$ with the same basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$. Formally speaking, it means that for every $i = 1, \ldots, p$, there is some fixed mapping $\nu_i$ from the set $\sigma$ into the set of all injective endomorphisms of the ring $R_i$ such that any two endomorphisms $\nu_i(\alpha_j)$ and $\nu_i(\alpha_k)$ of $R_i$ commute ($1 \leq j, k \leq n$). We shall identify elements $\alpha_j$ with their images $\nu_i(\alpha_j)$ and say that elements of the set $\sigma$ act as mutually commuting injective endomorphisms of the ring $R_i$ ($i = 1, \ldots, p$).

Let $R_1$ and $R_2$ be different rings with the same basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$. A ring homomorphism $\phi: R_1 \to R_2$ is called a difference (or $\sigma$-) homomorphism if $\phi(\alpha(a)) = \alpha(\phi(a))$ for any $\alpha \in \sigma, a \in R_1$. Clearly, if $\phi: R_1 \to R_2$ is a $\sigma$-homomorphism of inversive difference rings, then $\phi(\alpha^{-1}(a)) = \alpha^{-1}(\phi(a))$ for any $\alpha \in \sigma, a \in R_1$. If a $\sigma$-homomorphism is an isomorphism (endomorphism, automorphism, etc.), it is called a difference (or $\sigma$-) isomorphism (respectively, difference (or $\sigma$-) endomorphism, difference (or $\sigma$-) automorphism, etc.). If $R_1$ and $R_2$ are two $\sigma$-overrings of the same $\sigma$-ring $R_0$ and $\phi: R_1 \to R_2$ is a $\sigma$-homomorphism such that $\phi(a) = a$ for any $a \in R_0$, we say that $\phi$ is a difference (or $\sigma$-) homomorphism over $R_0$ or that $\phi$ leaves the ring $R_0$ fixed. It is easy to see that the kernel of any $\sigma$-homomorphism of $\sigma$-rings $\phi: R \to R'$ is an inversive $\sigma$-ideal of $R$. Conversely, let $g$ be a surjective homomorphism of a $\sigma$-ring $R$ onto a ring $S$ such that $\text{Ker} \ g$ is a $\sigma^*$-ideal of $R$. Then there is a unique structure of a $\sigma$-ring on $S$ such that $g$ is a $\sigma$-homomorphism. In particular, if $I$ is a $\sigma^*$-ideal of a $\sigma$-ring $R$, then the factor ring $R/I$ has a unique structure of a $\sigma$-ring such that the canonical surjection $R \to R/I$ is a $\sigma$-homomorphism. In this case $R/I$ is said to be the difference (or $\sigma$-) factor ring of $R$ by the $\sigma^*$-ideal $I$. 
Since a radical of a difference ideal is a difference ideal, every maximal \( \sigma \)-ideal \( I \) of a difference (\( \sigma \))-ring \( R \) is radical and inversive. In this case \( R/I \) is a reduced \( \sigma \)-ring (that is, a \( \sigma \)-ring without non-zero nilpotent elements).

**Example 2.1.2.** Let \( A \) be the set of all sequences \( a = (a_1, a_2, \ldots) \) of elements of an algebraically closed field \( C \). Consider an equivalence relation on \( A \) such that \( a = (a_1, a_2, \ldots) \) is equivalent to \( b = (b_1, b_2, \ldots) \) if and only if \( a_n = b_n \) for all sufficiently large \( n \in \mathbb{N} \) (that is, there exists \( n_0 \in \mathbb{N} \) such that \( a_n = b_n \) for all \( n > n_0 \)). Clearly, the corresponding set \( S \) of equivalence classes is a ring with respect to coordinatewise addition and multiplication of class representatives. This ring can be treated as an ordinary difference ring with respect to the mapping \( \alpha \) sending an equivalence class with a representative \( (a_1, a_2, a_3, \ldots) \) to the equivalent class with the representative \( (a_2, a_3, \ldots) \). It is easy to see that this mapping is well-defined and it is an automorphism of the ring \( S \). The field \( C \) can be naturally identified with the ring of constants of the difference ring \( S \).

Let \( C(z) \) be the field of rational functions in one complex variable \( z \). Then \( C(z) \) can be considered as an ordinary difference field with respect to the automorphism \( \beta \) such that \( \beta(z) = z + 1 \) and \( \beta(a) = a \) for any \( a \in C \). In this case, the mapping \( \phi : C(z) \to S \) that sends a function \( f(z) \) to the equivalence class of the element \( (f(0), f(1), \ldots) \) is an injective difference ring homomorphism.

Let \( R \) be a difference ring with a basic set \( \sigma = \{ \alpha_1, \ldots, \alpha_n \} \). A \( \sigma \)-overring \( U \) of \( R \) is called an inversive closure of \( R \), if the elements of \( \sigma \) act as mutually commuting automorphisms of the ring \( U \) (they are denoted by the same symbols \( \alpha_1, \ldots, \alpha_n \)) and for any \( a \in U \), there exists an automorphism \( \tau \in T_\sigma \) of the ring \( U \) such that \( \tau(a) \in R \).

The ordinary version of the following statement can be found in [32, Chapter 2, Theorem II]; the corresponding theorem for partial difference rings was proved in [7, Theorem 3.1]. (Actually, many results in difference algebra were first proved for ordinary difference rings and then generalized to difference rings with several translations. In such cases we refer to both corresponding publications.)

**Proposition 2.1.3 ([32, Chapter 2, Theorem II], [7, Theorem 3.1]).**

(i) Every difference ring has an inversive closure.

(ii) If \( U_1 \) and \( U_2 \) are two inversive closures of a difference ring \( R \), then there exists a difference \( R \)-isomorphism of \( U_1 \) onto \( U_2 \).

(iii) If \( R \) is a difference ring with a basic set \( \sigma \) and \( U \) an inversive \( \sigma \)-ring containing \( R \) as a \( \sigma \)-subring, then \( U \) contains an inversive closure of \( R \).

(iv) If a difference ring \( R \) is an integral domain (a field), then its inversive closure is also an integral domain (respectively, a field).

(v) Let \( R_1 \) and \( R_2 \) be two difference rings with the same basic set \( \sigma \), \( R_1^* \) and \( R_2^* \) their inversive closures, and \( \phi : R_1 \to R_2 \) a \( \sigma \)-homomorphism. Then \( \phi \) has a unique extension to a \( \sigma \)-homomorphism \( R_1^* \to R_2^* \).

The inversive closure of an ordinary difference ring \( R \) with a basic set \( \sigma = \{ \alpha \} \) can be constructed as follows. Let \( R' = \alpha(R) \) and let \( R'' \) be an isomorphic copy of \( R \) such that \( R \cap R'' = \emptyset \). Let \( \beta : R \to R'' \) be the corresponding isomorphism and \( R''' = \beta(R') = \beta\alpha(R) \). If one replaces all elements of \( R''' \) by the corresponding elements of \( R \), then
\( R'' \) will be transformed into an overring \( R_1 = R \). The mapping \((\rho \beta)^{-1}\), where \( \rho \) denotes the replacement mapping \( R'' \to R_1 \), is an injective endomorphism of \( R_1 \) that extends \( \alpha \). This endomorphism will be also denoted by \( \alpha \) and \( R_1 \) will be treated as a \( \sigma \)-overring of \( R \). Now, let us consider the sequence of \( \sigma \)-rings \( R = R_0, R_1, R_2, \ldots \) where every \( R_n \) is a \( \sigma \)-overring of \( R_{n-1} \) obtained by the forgoing procedure, that is, \( R_n = (R_{n-1})^\ast \) for \( n = 1, 2, \ldots \). Let us set \( R^\ast = \bigcup_{n \in \mathbb{N}} R_n \) and define the extension of \( \alpha \) to \( R^\ast \) as follows (we denote this extension by the same letter \( \alpha \)). If \( a \in R^\ast \), then \( a \in R_n \) for some \( n \) and the extension of \( \alpha \) to \( R_n \) determines an element \( \alpha(a) \in R_n \) which we consider as the image of \( a \) under the mapping \( \alpha : R^\ast \to R^\ast \). It is easy to see that the obtained mapping \( \alpha : R^\ast \to R^\ast \) is well-defined (the image of an element \( a \in R^\ast \) does not depend on the choice of \( R_n \) such that \( a \in R_n \)) and \( R^\ast \) is an inversive closure of \( R \).

Let \( R \) be a partial difference ring with a basic set \( \sigma = \{ \alpha_1, \ldots, \alpha_n \} \). Considering \( R \) as an ordinary difference ring whose basic set consists of the endomorphism \( \tau = \prod_{i=1}^n \alpha_i \), we can construct the inversive closure \( R^\ast \) of this ring. Now we can extend each \( \alpha_i (1 \leq i \leq n) \) to \( R^\ast \) as follows. For any \( a \in R^\ast \), let \( r(a) \) denote the smallest non-negative integer such that \( \tau^r(a)(a) \in R \) (we denote the extension of \( \tau \) to \( R^\ast \) by the same letter \( \tau \)). Setting \( \alpha_i(a) = \tau^{-r(a)}(a)\alpha_i \tau^{r(a)}(a) \) for any \( a \in R (1 \leq i \leq n) \), we obtain well-defined extensions of the endomorphisms \( \alpha_1, \ldots, \alpha_n \) to \( R^\ast \) that make \( R^\ast \) an inversive closure of the \( \sigma \)-ring \( R \).

If \( H \) is an inversive difference field with a basic set \( \sigma \) and \( G \) a \( \sigma \)-subfield of \( H \), then the set \( \{ a \in H \mid \tau(a) \in G \text{ for some } \tau \in T_n \} \) is a \( \sigma^\ast \)-subfield of \( H \) denoted by \( G^\ast_H \) (or \( G^\ast \) if one considers subfields of a fixed \( \sigma^\ast \)-field \( H \)). This field is said to be the inversive closure of \( G \) in \( H \). Clearly, \( G^\ast_H \) is the intersection of all \( \sigma^\ast \)-subfields of \( H \) containing \( G \).

**Proposition 2.1.4** [7, Section 5]. Let \( H \) be a \( \sigma^\ast \)-field and let \( * \) be the operation that assigns to each \( \sigma \)-subfield \( F \subseteq H \) its inversive closure in \( H \). Let \( F \) and \( G \) be two \( \sigma \)-subfields of \( H \) and \( \langle F, G \rangle \) denote the \( \sigma \)-field \( F(G) = G(F) \) (the “\( \sigma \)-compositum” of \( F \) and \( G \)). Then

(i) \( F^{**} = F^\ast \).

(ii) \( \langle F, G \rangle^{**} = \langle F^\ast, G^\ast \rangle \).

(iii) Every \( \sigma \)-isomorphism of \( F \) onto \( G \) has a unique extension to a \( \sigma \)-isomorphism of \( F^\ast \) onto \( G^\ast \).

(iv) If \( K \) is a \( \sigma \)-subfield of \( H \) contained in \( F \cap G \) and \( F \) and \( G \) are free (linearly disjoint, quasi-linearly disjoint) over \( K \), then \( F^\ast \) and \( G^\ast \) are free (linearly disjoint, quasi-linearly disjoint) over \( K^\ast \).

(The corresponding definitions can be found in [144, Vol. I, Chapter II].)

Let \( R \) be a difference ring with a basic set \( \sigma \). A subset \( S \) of the ring \( R \) is said to be a \( \sigma \)-subset of \( R \) if \( \sigma(s) \in S \) for any \( s \in S \), \( \alpha \in \sigma \). If the \( \sigma \)-ring \( R \) is inversive and \( S \) is a \( \sigma \)-subset of \( R \) such that \( \alpha^{-1}(s) \in S \) for any \( s \in S \), \( \alpha \in \sigma \), then \( S \) is said to be a \( \sigma^\ast \)-subset of the ring \( R \). By a multiplicative \( \sigma \)-subset of a \( \sigma \)-ring \( R \) we mean a \( \sigma \)-subset \( S \) of \( R \) such that \( 1 \in S \), \( 0 \notin S \), and \( st \in S \) whenever \( s \in S \) and \( t \in R \). A multiplicative \( \sigma^\ast \)-subset of an inversive \( \sigma \)-ring \( R \) is a multiplicative \( \sigma \)-subset of \( R \) such that \( \alpha^{-1}(s) \in S \) for any \( s \in S \), \( \alpha \in \sigma \).

The following statement is a natural generalization of [32, Chapter 2, Theorem III].
PROPOSITION 2.1.5. Let \( S \) be a multiplicative \( \sigma \)-subset of a \( \sigma \)-ring \( R \) and let \( S^{-1} R \) be the ring of fractions of \( R \) with denominators in \( S \). Then \( S^{-1} R \) has the unique structure of \( \sigma \)-ring such that the natural mapping \( \nu : R \to S^{-1} R \) \( (a \mapsto \frac{a}{1}) \) is a \( \sigma \)-homomorphism. If the \( \sigma \)-ring \( R \) is inversive and \( S \) is a multiplicative \( \sigma^* \)-subset of \( R \), then \( S^{-1} R \) is a \( \sigma^* \)-overring of \( R \).

If \( S \) is a multiplicative \( \sigma \)-subset of a \( \sigma \)-ring \( R \), then the ring \( S^{-1} R \) is said to be a \( \sigma \)-ring of fractions of \( R \) with denominators in \( S \). If the \( \sigma \)-ring \( R \) is inversive and \( S \) is a multiplicative \( \sigma^* \)-subset of \( R \), then \( S^{-1} R \) is called the \( \sigma^* \)-ring of fractions of \( R \) with denominators in \( S \).

PROPOSITION 2.1.6. Let \( R \) and \( R' \) be difference rings with the same basic set \( \sigma \) and let \( \phi : R \to R' \) be a \( \sigma \)-homomorphism such that \( \phi(s) \) is a unit of \( R' \) for any \( s \in S \). Then \( \phi \) factors uniquely through the canonical mapping \( \nu : R \to S^{-1} R \): there exists a unique \( \sigma \)-homomorphism \( \psi : S^{-1} R \to R' \) such that \( \psi \circ \nu = \phi \). (This \( \sigma \)-homomorphism is given by \( \psi(\frac{a}{1}) = \phi(a)\phi(s)^{-1}.\))

The last proposition shows that if a difference ring \( R \) with a basic set \( \sigma \) is an integral domain, then its quotient field \( Q(R) \) can be naturally considered as a \( \sigma \)-overring of \( R \). (We identify an element \( a \in R \) with its canonical image \( \frac{a}{1} \) in \( Q(R) \).) In this case \( Q(R) \) is said to be the quotient difference (or \( \sigma \)-) field of \( R \). Clearly, if the \( \sigma \)-ring \( R \) is inversive, then its quotient \( \sigma \)-field \( Q(R) \) is also inversive. Furthermore, if a \( \sigma \)-field \( K \) contains an integral domain \( R \) as its \( \sigma \)-subring, then \( K \) contains the quotient \( \sigma \)-field \( Q(R) \). Also, if the \( \sigma \)-field \( K \) is inversive and \( R \) is a \( \sigma^* \)-subring of \( K \), then \( Q(R) \) is a \( \sigma^* \)-subfield of \( K \).

2.2. Rings of difference and inversive difference polynomials. Algebraic difference equations

Let \( R \) be a difference ring with a basic set \( \sigma = \{a_1, \ldots, a_n\} \), \( T_\sigma \) the free commutative semi-group generated by \( \sigma \), and \( U = \{u_\lambda \mid \lambda \in \Lambda\} \) a family of elements from some \( \sigma \)-overring of \( R \). We say that the family \( U \) is transformally (or \( \sigma \)-algebraically) dependent over \( R \), if the family \( T_\sigma(U) = \{\tau(u_\lambda) \mid \tau \in T_\sigma, \lambda \in \Lambda\} \) is algebraically dependent over \( R \) (that is, there exist elements \( v_1, \ldots, v_k \in T_\sigma(U) \) and a non-zero polynomial \( f(X_1, \ldots, X_k) \) with coefficients from \( R \) such that \( f(v_1, \ldots, v_k) = 0 \)). Otherwise, the family \( U \) is said to be transformally (or \( \sigma \)-algebraically) independent over \( R \) or a family of difference (or \( \sigma \)-) indeterminates over \( R \). In the last case, the \( \sigma \)-ring \( \mathbb{R}((u_\lambda)_{\lambda \in \Lambda})_\sigma \) is called the algebra of difference (or \( \sigma \)-) polynomials in the difference (or \( \sigma \)-) indeterminates \( \{(u_\lambda)_{\lambda \in \Lambda}\} \) over \( R \). If a family consisting of one element \( u \) is \( \sigma \)-algebraically dependent over \( R \), the element \( u \) is said to be transformally algebraic (or \( \sigma \)-algebraic) over the \( \sigma \)-ring \( R \). If the set \( \{\tau(u) \mid \tau \in T\} \) is algebraically independent over \( R \), we say that \( u \) is transformally (or \( \sigma \)-) transcendental over the ring \( R \).

Let \( R \) be a \( \sigma \)-field, \( L \) a \( \sigma \)-overfield of \( R \), and \( S \subseteq L \). We say that the set \( S \) is \( \sigma \)-algebraic over \( R \) if every element \( a \in S \) is \( \sigma \)-algebraic over \( R \). If every element of \( L \) is \( \sigma \)-algebraic over \( R \), we say that \( L \) is \( \sigma \)-algebraic field extension of the \( \sigma \)-field \( R \).
PROPOSITION 2.2.1 ([32, Chapter 2, Theorem 1], [88, Proposition 3.3.7]). Let \( R \) be a difference ring with a basic set \( \sigma \) and \( I \) an arbitrary set. Then there exists an algebra of \( \sigma \)-polynomials over \( R \) in a family of \( \sigma \)-indeterminates with indices from the set \( I \). If \( S \) and \( S' \) are two such algebras, then there exists a \( \sigma \)-isomorphism \( S \rightarrow S' \) that leaves the ring \( R \) fixed. If \( R \) is an integral domain, then any algebra of \( \sigma \)-polynomials over \( R \) is an integral domain.

The algebra of \( \sigma \)-polynomials over the \( \sigma \)-ring \( R \) can be constructed as follows. Let \( T = T_\sigma \) and let \( S \) be the polynomial \( R \)-algebra in the set of indeterminates \( \{ y_{i,t} \}_{i \in I, t \in T} \) with indices from the set \( I \times T \). For any \( f \in S \) and \( \alpha \in \sigma \), let \( \alpha(f) \) denote the polynomial from \( S \) obtained by replacing every indeterminate \( y_{i,t} \) that appears in \( f \) by \( y_{i,\alpha(t)} \) and every coefficient \( a \in R \) by \( \alpha(a) \). We obtain an injective endomorphism \( S \rightarrow S \) that extends the original endomorphism \( \alpha \) of \( R \) to the ring \( S \) (this extension is denoted by the same letter \( \alpha \)). Setting \( y_i = y_{i,1} \) (where \( 1 \) denotes the identity of the semigroup \( T \)) we obtain a \( \sigma \)-algebraically independent over \( R \) set \( \{ y_i \mid i \in I \} \) such that \( S = R[ \{ y_i \}_{i \in I} ] \). Thus, \( S \) is an algebra of \( \sigma \)-polynomials over \( R \) in a family of \( \sigma \)-indeterminates \( \{ y_i \mid i \in I \} \).

Let \( R \) be an inversive difference ring with a basic set \( \sigma \), \( \Gamma = \Gamma_\sigma \), \( I \) a set, and \( S^* \) a polynomial ring in the set of indeterminates \( \{ y_{i,y} \}_{i \in I, y \in \Gamma} \) with indices from the set \( I \times \Gamma \). If we extend the automorphisms \( \beta \in \sigma^* \) to \( S^* \) setting \( \beta(y_{i,y}) = y_{i,\beta(y)} \) for any \( y_{i,y} \) and denote \( y_{i,1} \) by \( y_i \), then \( S^* \) becomes an inversive difference overring of \( R \) generated (as a \( \sigma^* \)-overring) by the family \( \{ (y_i)_{i \in I} \} \). Obviously, this family is \( \sigma^* \)-algebraically independent over \( R \), that is, the set \( \{ y_i \mid i \in I \} \) is algebraically independent over \( R \).

(Note that a set is \( \sigma^* \)-algebraically dependent (independent) over an inversive \( \sigma \)-ring if and only if this set is \( \sigma \)-algebraically dependent (respectively, independent) over this ring.) The ring \( S^* = R[ \{ y_i \}_{i \in I} ]^* \) is called the algebra of inversive difference (or \( \sigma^* \))-polynomials over \( R \) in the set of \( \sigma^* \)-indeterminates \( \{ (y_i)_{i \in I} \} \). It is easy to see that \( S^* \) is the inversive closure of the ring of \( \sigma \)-polynomials \( R[ \{ y_i \}_{i \in I} ] \) over \( R \). Furthermore, if a family \( \{ (u_i)_{i \in I} \} \) from some \( \sigma^* \)-overring of \( R \) is \( \sigma^* \)-algebraically independent over \( R \), then the inversive difference ring \( R[ \{ u_i \}_{i \in I} ]^* \) is naturally \( \sigma \)-isomorphic to \( S^* \). Any such overring \( R[ \{ u_i \}_{i \in I} ]^* \) is said to be an algebra of inversive difference (or \( \sigma^* \))-polynomials over \( R \) in the set of \( \sigma^* \)-indeterminates \( \{ (u_i)_{i \in I} \} \). We obtain the following analog of Proposition 2.2.1.

PROPOSITION 2.2.1* [88, Proposition 3.4.4]. Let \( R \) be an inversive difference ring with a basic set \( \sigma \) and \( I \) an arbitrary set. Then there exists an algebra of \( \sigma^* \)-polynomials over \( R \) in a family of \( \sigma^* \)-indeterminates with indices from the set \( I \). If \( S \) and \( S' \) are two such algebras, then there exists a \( \sigma^* \)-isomorphism \( S \rightarrow S' \) that leaves the ring \( R \) fixed. If \( R \) is an integral domain, then any algebra of \( \sigma^* \)-polynomials over \( R \) is an integral domain.

Let \( R \) be a \( \sigma \)-ring, \( R[ \{ (y_i)_{i \in I} \} ] \) an algebra of difference polynomials in a family of \( \sigma \)-indeterminates \( \{ (y_i)_{i \in I} \} \), and \( \{ (\eta_i)_{i \in I} \} \) a set of elements from some \( \sigma \)-overring of \( R \). Since the set \( \{ \tau(y_i) \mid i \in I, \tau \in T_\sigma \} \) is algebraically independent over \( R \), there exists a unique ring homomorphism \( \phi_\eta : R[ \{ \tau(y_i)_{i \in I, \tau \in T_\sigma} \} ] \rightarrow R[ \{ \tau(\eta_i)_{i \in I, \tau \in T_\sigma} \} ] \) that maps every \( \tau(y_i) \) onto \( \tau(\eta_i) \) and leaves \( R \) fixed. Clearly, \( \phi_\eta \) is a surjective \( \sigma \)-homomorphism of \( R[ \{ (y_i)_{i \in I} \} ] \) onto \( R[ \{ (\eta_i)_{i \in I} \} ] \); it is called the substitution of \( (\eta_i)_{i \in I} \) for \( (y_i)_{i \in I} \). Similarly, if \( R \) is an inversive
\(\sigma\)-ring, \(R[(y_i)_{i\in I}]^*\) an algebra of \(\sigma^*\)-polynomials over \(R\) and \((\eta_i)_{i\in I}\) a family of elements from a \(\sigma^*\)-overring of \(R\), one can define a surjective \(\sigma\)-homomorphism \(R[(y_i)_{i\in I}]^* \rightarrow R[(\eta_i)_{i\in I}]^*\) that maps every \(y_i\) onto \(\eta_i\) and leaves the ring \(R\) fixed. This homomorphism is also called the substitution of \((\eta_i)_{i\in I}\) for \((y_i)_{i\in I}\). (It will always be clear whether we talk about substitutions for difference or inversive difference polynomials.) If \(g\) is a \(\sigma\)- or \(\sigma^*\)-polynomial, then its image under a substitution of \((\eta_i)_{i\in I}\) for \((y_i)_{i\in I}\) is denoted by \(g((\eta_i)_{i\in I})\). The kernel of a substitution \(\phi\) is an inversive difference ideal of the \(\sigma\)-ring \(R[(y_i)_{i\in I}]\) (or the \(\sigma^*\)-ring \(R[(y_i)_{i\in I}]^*\)); it is called the defining difference (or \(\sigma\)-) ideal of the family \((\eta_i)_{i\in I}\) over \(R\). If \(R\) is a \(\sigma\)- (or \(\sigma^*\)-) field and \((\eta_i)_{i\in I}\) is a family of elements from some its \(\sigma\) (respectively, \(\sigma^*\)) overfield \(S\), then \(R[(\eta_i)_{i\in I}]\) (respectively, \(R[(\eta_i)_{i\in I}]^*\)) is an integral domain (it is contained in the field \(S\)). It follows that the defining \(\sigma\)-ideal \(P\) of the family \((\eta_i)_{i\in I}\) over \(R\) is a prime inversive difference ideal of the ring \(R[(\eta_i)_{i\in I}]\) (respectively, of the ring of \(\sigma^*\)-polynomials \(R[(y_i)_{i\in I}]^*\)). Therefore, the difference field \(R[(\eta_i)_{i\in I}]\) can be treated as the quotient \(\sigma\)-field of the \(\sigma\)-ring \(R[(y_i)_{i\in I}]^*/P\). (In the case of inversive difference rings, the \(\sigma^*\)-field \(R[(\eta_i)_{i\in I}]^*/P\) can be considered as a quotient \(\sigma\)-field of the \(\sigma^*\)-ring \(R[(y_i)_{i\in I}]^*/P\).)

Let \(F\) be a difference field with a basic set \(\sigma\) and \(s\) a positive integer. By an \(s\)-tuple over \(F\) we mean an \(s\)-dimensional vector \(a = (a_1, \ldots, a_s)\) whose coordinates belong to some \(\sigma\)-overfield of \(F\). If the \(\sigma\)-field \(F\) is inversive, the coordinates of an \(s\)-tuple over \(F\) are supposed to lie in some \(\sigma^*\)-overfield of \(F\). If each \(a_i\) \((1 \leq i \leq s)\) is \(\sigma\)-algebraic over the \(\sigma\)-field \(F\), we say that the \(s\)-tuple \(a\) is \(\sigma\)-algebraic over \(F\).

**DEFINITION 2.2.2.** Let \(F\) be a difference (inversive difference) field with a basic set \(\sigma\) and let \(R\) be the algebra of \(\sigma\) (respectively, \(\sigma^*\))-polynomials in finitely many \(\sigma\) (respectively, \(\sigma^*\)) indeterminates \(y_1, \ldots, y_s\) over \(F\). Furthermore, let \(\Phi = \{f_j \mid j \in J\}\) be a set of \(\sigma\) (respectively, \(\sigma^*\))-polynomials from \(R\). An \(s\)-tuple \(\eta = (\eta_1, \ldots, \eta_s)\) over \(F\) is said to be a solution of the set \(\Phi\) or a solution of the system of difference algebraic equations \(f_j(\eta_1, \ldots, \eta_s) = 0\) \((j \in J)\) if \(\Phi\) is contained in the kernel of the substitution of \((\eta_1, \ldots, \eta_s)\) for \((y_1, \ldots, y_s)\). In this case we also say that \(\eta\) annuls \(\Phi\). (If \(\Phi\) is a subset of a ring of inversive difference polynomials, the system is said to be a system of algebraic \(\sigma^*\)-equations.)

As we have seen, if one fixes an \(s\)-tuple \(\eta = (\eta_1, \ldots, \eta_s)\) over a \(\sigma\)-field \(F\), then all \(\sigma\)-polynomials of the ring \(F[y_1, \ldots, y_s]\), for which \(\eta\) is a solution, form a prime inversive difference ideal. It is called the defining \(\sigma\)-ideal of \(\eta\). If \(\eta\) is an \(s\)-tuple over a \(\sigma^*\)-field \(F\), then all \(\sigma^*\)-polynomials \(g\) of the ring \(F[y_1, \ldots, y_s]^*\) such that \(g(\eta_1, \ldots, \eta_s) = 0\) form a prime \(\sigma^*\)-ideal of \(F[y_1, \ldots, y_s]^*\). This ideal is called the defining \(\sigma^*\)-ideal of \(\eta\) over \(F\).

Let \(\Phi\) be a subset of the algebra of \(\sigma\)-polynomials \(F[y_1, \ldots, y_s]\) over a \(\sigma\)-field \(F\). An \(s\)-tuple \(\eta = (\eta_1, \ldots, \eta_s)\) over \(F\) is called a generic zero of \(\Phi\) if for any \(\sigma\)-polynomial \(A \in F[y_1, \ldots, y_s]\), the inclusion \(A \in \Phi\) holds if and only if \(A(\eta_1, \ldots, \eta_s) = 0\). If the \(\sigma\)-field \(F\) is inversive, then the notion of a generic zero of a subset of \(F[y_1, \ldots, y_s]^*\) is defined similarly.

Two \(s\)-tuples \(\eta = (\eta_1, \ldots, \eta_s)\) and \(\zeta = (\zeta_1, \ldots, \zeta_s)\) over a \(\sigma\)- (or \(\sigma^*\)-) field \(F\) are called equivalent over \(F\) if there is a \(\sigma\)-isomorphism \(F[\eta_1, \ldots, \eta_s] \rightarrow F[\zeta_1, \ldots, \zeta_s]\) (respec
is easy to check that the \( \sigma \)-algebra of set -tuples are solutions in the sense of classical algebraic geometry.\)

**PROPOSITION 2.2.3** ([32, Chapter 2, Theorem VII], [88, Proposition 3.3.20]). Let \( S \) denote the algebra of \( \sigma \)-polynomials \( F\{y_1, \ldots, y_s\} \) or the algebra of \( \sigma^* \)-polynomials \( F\{y_1, \ldots, y_s\}^* \) over a difference (respectively, inversive difference) field \( F \) with a basic set \( \sigma \).

(i) A set \( \Phi \subseteq S \) has a generic zero if and only if \( \Phi \) is a prime \( \sigma^* \)-ideal of \( S \). If \( (\eta_1, \ldots, \eta_k) \) is a generic zero of \( \Phi \), then \( F\{\eta_1, \ldots, \eta_k\} \) (or \( F\{\eta_1, \ldots, \eta_s\}^* \) if we consider the algebra of \( \sigma^* \)-polynomials over a \( \sigma^* \)-field \( F \)) is \( \sigma \)-isomorphic to the quotient \( \sigma \)-field of \( S/\Phi \).

(ii) Any \( s \)-tuple over \( F \) is a generic zero of some prime \( \sigma^* \)-ideal of \( S \).

(iii) If two \( s \)-tuples over \( F \) are generic zeros of the same prime \( \sigma^* \)-ideal of \( S \), then these \( s \)-tuples are equivalent.

**EXAMPLE 2.2.4** (see [32, Chapter 2, Example 4]). Let us consider \( \mathbb{C} \) as an ordinary difference field whose basic set \( \sigma \) consists of the identity automorphism \( \alpha \). Let \( \mathbb{C}\{y\} \) be the algebra of \( \sigma \)-polynomials in one \( \sigma \)-indeterminate \( y \) over \( \mathbb{C} \) and let \( \langle k \rangle y \) denote the \( k \)-th transform \( \alpha^k y \) \((k = 1, 2, \ldots)\). Furthermore, let \( M \) be the field of functions of one complex variable \( z \) meromorphic on the whole complex plane. Then \( M \) can be viewed as a \( \sigma \)-overfield of \( \mathbb{C} \) if one extends \( \sigma \) by setting \( \alpha f(z) = f(z + 1) \) for any function \( f \in M \). It is easy to check that the \( \sigma \)-polynomial \( A = (\langle 1 \rangle y - y)^2 - 2(\langle 1 \rangle y + y) + 1 \) is irreducible in \( \mathbb{C}\{y\} \) (when this ring is treated as a polynomial ring in the denumerable set of indeterminates \( y, \langle 1 \rangle y, \langle 2 \rangle y, \ldots \)). Furthermore, if \( c(z) \) is a periodic function from \( M \) with period 1, then \( \xi = (z + c(z))^2 \) and \( \eta = (c(z)e^{i\pi z} + \frac{1}{2})^2 \) are solutions of \( A \). (\( \xi \) is a solution of the system of the \( \sigma \)-polynomials \( A \) and \( A' = \langle 2 \rangle y - 2 \langle 1 \rangle y + y - 2 \), while \( \eta \) is the solution of the system of \( A \) and \( A'' = \langle 2 \rangle y - y \).) The fact that an irreducible \( \sigma \)-polynomial in one \( \sigma \)-indeterminate may have two distinct sets of solutions, each of which depends on an arbitrary periodic function, does not have an analog in the theory of differential polynomials.

Let \( F \) be a difference field with a basic set \( \sigma \), \( R = F\{y_1, \ldots, y_s\} \) the algebra of \( \sigma \)-polynomials in a set of \( s \) \( \sigma \)-indeterminates \( y_1, \ldots, y_s \) over \( F \), and \( \Phi \subseteq R\{y_1, \ldots, y_s\} \). Let \( \bar{a} = \{a_{i, \tau} \mid i = 1, \ldots, s, \tau \in T_\sigma\} \) be a family of elements from some \( \sigma \)-overfield of \( F \). The family \( \bar{a} \) (indexed by the set \( \{1, \ldots, s\} \times T_\sigma \)) is said to be an algebraic solution of the set of \( \sigma \)-polynomials \( \Phi \) if \( \bar{a} \) is a solution of \( \Phi \) when this set is treated as a set of polynomials in the polynomial ring \( F\{y_{i, \tau} \mid i = 1, \ldots, s, \tau \in T_\sigma\} \). (This polynomial ring in the denumerable family of indeterminates \( \{y_{i, \tau} \mid i = 1, \ldots, s, \tau \in T_\sigma\} \) coincides with \( R \) \((y_{i, \tau} \) stands for \( \tau(y_i)\)), but it is not considered as a difference ring, so the solutions of its subsets are solutions in the sense of classical algebraic geometry.)

It is easy to see that every solution \( a = (a_1, \ldots, a_s) \) of a set \( \Phi \subseteq F\{y_1, \ldots, y_s\} \) produces its algebraic solution \( \bar{a} = \{\tau(a_i) \mid i = 1, \ldots, s, \tau \in T_\sigma\} \). On the other hand, not every algebraic solution can be obtained from a solution in this way.

**EXAMPLE 2.2.5** (see [32, Chapter 2, Example 6]). Let \( \mathbb{C} \) be the field of complex numbers considered as an ordinary difference field whose basic set consists of the complex conjuga-
tion (that is, \( \sigma = \{ a \} \) where \( a(a + bi) = a - bi \) for any complex number \( a + bi \)). Let \( C(y) \) be the ring of \( \sigma \)-polynomials in one \( \sigma \)-indeterminate \( y \). If \( A = y^2 + 1 \in C(y) \), then the
1-tuples \( (i) \) and \((-i)\) are solutions of the \( \sigma \)-polynomial \( A \) that produce algebraic solutions \((i, -i, i, -i, \ldots)\) and 
\((-i, -i, i, -i, \ldots)\) of \( A \). At the same time, the sequence \((-i, i, i, i, \ldots)\) is an algebraic solution of \( A \) that is not a solution of this \( \sigma \)-polynomial.

If \( \Phi \) is a subset of an algebra of \( \sigma^* \)-polynomials \( F[y_1, \ldots, y_s]^\sigma \) over an inversive difference field \( F \) with a basic set \( \sigma \), then an algebraic solution of \( \Phi \) is defined as a family \( a^\sigma = \{ a_i, y \mid i = 1, \ldots, s, y \in \Gamma_\sigma \} \) that annihilates every polynomial from \( \Phi \) when \( \Phi \) is treated as a subset of the polynomial ring \( F[y(y_i) \mid i = 1, \ldots, s, y \in \Gamma_\sigma] \). As in the case of \( \sigma \)-polynomials, every solution \( a = \{ a_1, \ldots, a_s \} \) of a set of \( \sigma^* \)-polynomials generates the algebraic solution \( a^\sigma = \{ y(a_i) \mid i = 1, \ldots, s, y \in \Gamma_\sigma \} \) of this set, but not all algebraic solutions can be obtained in this way.

2.3. Autoreduced sets of difference and inversive difference polynomials.

Characteristic sets

Let \( F \) be a difference field with a basic set \( \sigma = \{ a_1, \ldots, a_n \} \), \( T = T_\sigma \), and \( R = F[y_1, \ldots, y_s] \) the algebra of difference polynomials in \( \sigma \)-indeterminates \( y_1, \ldots, y_s \) over \( F \). Then \( R \) can be viewed as a polynomial ring in the set of indeterminates \( TY = \{ \tau y_i \mid \tau \in T, 1 \leq i \leq s \} \) over \( F \) (here and below we often write \( \tau y_i \) instead of \( \tau(y_i) \)). Elements of this set are called terms. If \( \tau = \alpha_1^{k_1} \cdots \alpha_n^{k_n} \in T \) \((k_1, \ldots, k_n \in \mathbb{N})\), then the number \( \text{ord} \tau = \sum_{v=1}^n k_v \) is called the order of \( \tau \). The order \( \text{ord} \tau \) of a term \( \tau = \tau y_i \in TY \) is defined as the order of \( \tau \). As usual, if \( \tau, \tau' \in T \), we say that \( \tau' \) divides \( \tau \) (and write \( \tau' | \tau \)) if \( \tau = \tau' \tau'' \) for some \( \tau'' \in T \). If \( \tau = \tau y_i \) and \( \nu = \tau' y_j \) are two terms from \( TY \), we say that \( \nu \) divides \( \tau \) (and write \( \nu | \tau \)) if \( i = j \) and \( \tau | \tau' \).

By a ranking of the family of indeterminates \( \{ y_1, \ldots, y_s \} \) we mean a well-ordering of the set of terms from \( TY \) that satisfies the following conditions. (We denote the order on \( TY \) by the usual symbol \( \leq \) and write \( u < v \) if \( u \leq v \) and \( u \neq v \).)

(i) \( u \leq \tau u \) for any \( u \in TY, \tau \in T \).

(ii) If \( u, v \in TY \) and \( u \leq v \), then \( tu \leq \tau v \) for any \( \tau \in T \).

A ranking of the family \( \{ y_1, \ldots, y_s \} \) is also referred to as a ranking of the set of terms \( TY \). It is said to be orderly if the inequality \( \text{ord} u \leq \text{ord} v \) \( (u, v \in TY) \) implies that \( u < v \). An important example of an orderly ranking is the standard ranking defined as follows: \( u = \alpha_1^{k_1} \cdots \alpha_n^{k_n} y_i \leq v = \alpha_1^{l_1} \cdots \alpha_n^{l_n} y_j \in TY \) if and only if \( (\sum_{v=1}^n k_v, i, k_1, \ldots, k_n) \) is less than or equal to \( (\sum_{v=1}^n l_v, j, l_1, \ldots, l_n) \) with respect to the lexicographic order on \( \mathbb{N}^{n+2} \). In what follows, we assume that an orderly ranking of \( TY \) has been fixed.

Let \( A \in F[y_1, \ldots, y_s] \). The greatest (with respect to the given ranking) element of \( TY \) that appears in the \( \sigma \)-polynomial \( A \) is called the leader of \( A \); it is denoted by \( u_A \). If \( A \) is written as a polynomial in \( u_A \), \( A = \sum_{i=0}^d l_i u_A^i \) \((d = \text{deg}_{u_A} A \) and the \( \sigma \)-polynomials \( l_0, \ldots, l_d \) do not contain \( u_A \)), then \( I_d \) is called the initial of the \( \sigma \)-polynomial \( A \); it is denoted by \( I_A \).

Let \( A \) and \( B \) be two \( \sigma \)-polynomials from \( F[y_1, \ldots, y_s] \). We say that \( A \) has lower rank than \( B \) (or simply \( A \) is less than \( B \)) and write \( A < B \), if either \( A \in F, B \notin F \) or \( u_A < u_B \).
or \( u_A = u_B = u \), \( \deg_u A < \deg_u B \). If neither \( A < B \) nor \( B < A \), we say that \( A \) and \( B \) have the same rank and write \( rk A = rk B \). The \( \sigma \)-polynomial \( A \) is said to be reduced with respect to \( B \) if \( A \) does not contain any power of a transform \( \tau u_B \) \((\tau \in T_\sigma)\) whose exponent is greater than or equal to \( \deg u_B \). If \( \Sigma \) is any subset of \( F[y_1, \ldots, y_s] \setminus F \), then a \( \sigma \)-polynomial \( A \in F[y_1, \ldots, y_s] \) is said to be reduced with respect to \( \Sigma \) if \( A \) is reduced with respect to every element of \( \Sigma \).

A set \( \Sigma \subseteq F[y_1, \ldots, y_s] \) is called an autoreduced set if either \( \Sigma = \emptyset \) or \( \Sigma \cap F = \emptyset \) and every element of \( \Sigma \) is reduced with respect to all other elements of \( \Sigma \). It is easy to see that distinct elements of an autoreduced set have distinct leaders. It follows from [86, Chapter 0, Lemma 15(a)] (see also [88, Lemma 2.2.1]) that every autoreduced set is finite.

The ordinary version of the following reduction theorem was proved in [136, Section 5]. In [36] R.M. Cohn generalized the result to the case of partial difference polynomial rings (actually, to the rings of partial difference-differential polynomials).

**THEOREM 2.3.1.** Let \( A = \{A_1, \ldots, A_p\} \) be an autoreduced set in a ring of \( \sigma \)-polynomials \( F[y_1, \ldots, y_s] \) over a difference field \( F \) with basic set \( \sigma \). Let \( I(A) = \{B \in F[y_1, \ldots, y_s] \mid \text{either } B = 1 \text{ or } B \text{ is a product of finitely many } \sigma \text{-} polynomials of the form } \tau(I_{A_i}) \text{ } (\tau \in T_\sigma, \text{ } i = 1, \ldots, p)\). Then for any \( C \in F[y_1, \ldots, y_s] \), there exist \( \sigma \)-polynomials \( J \in I(A) \) and \( C_0 \in F[y_1, \ldots, y_s] \) such that \( C_0 \) is reduced with respect to \( A \) and \( JC \equiv C_0 \text{ (mod } [A]\text{)} \) (i.e., \( JC - C_0 \in [A]\)).

With the notation of the theorem, the \( \sigma \)-polynomial \( C_0 \) is called the remainder of the \( \sigma \)-polynomial \( C \) with respect to \( A \). We also say that \( C \) reduces to \( C_0 \) modulo \( A \). (If \( A = \{A\} \), we say that \( C_0 \) is a remainder of \( C \) with respect to the \( \sigma \)-polynomial \( A \).)

The reduction process, that is, a transition from a given \( \sigma \)-polynomial \( C \) to a \( \sigma \)-polynomial \( C_0 \) satisfying the conditions of the theorem, can be performed in many ways. Let us describe one of them.

If \( C \) is reduced with respect to \( A \), we can take \( C_0 = C \) and \( J = 1 \). If \( C \) is not reduced with respect to \( A \), then \( C \) contains a power \( (\tau u_{A_j})^k \) of some term \( \tau u_{A_j} \) \((\tau \in T_\sigma, \text{ } 1 \leq i \leq p)\) whose exponent is greater than or equal to \( \deg u_{A_j} \). Such a term \( \tau u_{A_j} \) of the highest possible rank is called the \( A \)-leader of \( C \) and denoted by \( v_{A,C} \). Obviously, \( C \) can be written as \( C = Dv_{A,C}^d + E \) where \( D \) does not contain \( v_{A,C} \) and \( \deg v_{A,C} Q < d \). Let \( v_{A,C} = \tau u_{A_j} \) \((\tau \in T_\sigma, \text{ } 1 \leq j \leq p)\). Then \( v_{A,C} \) is the leader of \( \tau A_j \), \( I\tau A_j = I A_j \), and \( \deg v_{A,C} (\tau A_j) = d_j \) where \( d_j = \deg u_{A_j} A_j \). Consider the \( \sigma \)-polynomial \( C' = (\tau I_{A_j})C - v_{A,C}^d (\tau A_j)D \).

Clearly, \( v_{A,C'} \leq v_{A,C} \) and in the case of equality, \( \deg v_{A,C'} C' < d \). Furthermore, \( C' \equiv C \text{ (mod } [A]\text{)} \). Applying the same procedure to \( C' \) instead of \( C \) and continuing this process, we obtain a \( \sigma \)-polynomial \( C' \) such that \( C' \equiv C \text{ (mod } [A]\text{)} \) and \( v_{A,C'} \leq v_{A,C} \). Repeating the foregoing procedure we obtain a \( \sigma \)-polynomial \( C_0 \) that satisfies the conditions of the last theorem.

In what follows, the elements of an autoreduced set are always written in the order of increasing rank. (Thus, if \( A = \{A_1, \ldots, A_p\} \) an autoreduced set in \( F[y_1, \ldots, y_s] \), we assume that \( A_1 < \cdots < A_p \).)
**Definition 2.3.2.** Let \( A = \{A_1, \ldots, A_p\} \) and \( B = \{B_1, \ldots, B_q\} \) be two autoreduced sets in the algebra of difference polynomials \( F\{y_1, \ldots, y_s\} \). We say that \( A \) has lower rank than \( B \) and write \( rk_A < rk_B \) if one of the following conditions holds:

(i) there exists \( k \in \mathbb{N} \), \( 1 \leq k \leq \min\{p, q\} \), such that \( rk A_i = rk B_i \) for \( i = 1, \ldots, k - 1 \) and \( A_k < B_k \);

(ii) \( p > q \) and \( rk A_i = rk B_i \) for \( i = 1, \ldots, q \).

The proof of the following result is similar to the proof of the corresponding statement about autoreduced sets of differential polynomials (see [86, Chapter 1, Proposition 3]).

**Proposition 2.3.3.** In every non-empty set of autoreduced subsets of \( F\{y_1, \ldots, y_s\} \) there exists an autoreduced set of lowest rank.

If \( J \) is a non-empty subset (in particular, an ideal) of the ring \( F\{y_1, \ldots, y_s\} \), then the family of all autoreduced subsets of \( J \) is not empty (if \( 0 \neq A \in J \), then \( A = \{A\} \) is an autoreduced set). It follows from the last proposition that \( J \) contains an autoreduced subset of lowest rank. Such a subset is called a characteristic set of \( J \). The following proposition describes some properties of characteristic sets of difference polynomials.

**Proposition 2.3.4.** Let \( F \) be a difference field with a basic set \( \sigma \), \( J \) a difference ideal of the algebra of \( \sigma \)-polynomials \( F\{y_1, \ldots, y_s\} \), and \( \Sigma \) a characteristic set of \( J \). Then:

(i) The \( \sigma \)-ideal \( J \) does not contain non-zero difference polynomials reduced with respect to \( \Sigma \). In particular, if \( A \in \Sigma \), then \( I_A \not\in J \).

(ii) Let \( I = \prod_{A \in \Sigma} I_A \). If the ideal \( J \) is prime, then \( J = [\Sigma]; \Lambda(\Sigma) \) where \( \Lambda(\Sigma) \) is the free commutative multiplicative semigroup generated by the set \( \{\tau(I) \mid \tau \in T_\sigma\} \).

Let \( F \) be an inversive difference field with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \) and \( \Gamma \) the free commutative group generated by \( \sigma \). Let \( \mathbb{Z}_- \) denote the set of all non-negative integers and let \( \mathbb{Z}_1, \mathbb{Z}_2, \ldots, \mathbb{Z}_{2^n} \) be all distinct Cartesian products of \( n \) factors each of which is either \( \mathbb{N} \) or \( \mathbb{Z}_- \) (we assume that \( \mathbb{Z}_1 = \mathbb{N}^n \)). These sets are called orthants of \( \mathbb{Z}^n \). For any \( j = 1, \ldots, 2^n \), we set \( \Gamma_j = \{\gamma = \alpha_1^{k_1} \ldots \alpha_n^{k_n} \in \Gamma \mid (k_1, \ldots, k_n) \in \mathbb{Z}_j\} \). Furthermore, if \( \gamma = \alpha_1^{k_1} \ldots \alpha_n^{k_n} \in \Gamma \), then the number \( ord \gamma = \sum_{i=1}^{n} |k_i| \) will be called the order of \( \gamma \).

Let \( F\{y_1, \ldots, y_s\}^* \) be the algebra of \( \sigma^* \)-polynomials in \( \sigma^* \)-indeterminates \( y_1, \ldots, y_s \) over \( F \) and let \( Y \) denote the set \( \{y y_i \mid \gamma \in \Gamma, 1 \leq i \leq s\} \) whose elements are called terms (here and below we often write \( \gamma y_i \) for \( y(\gamma y_i) \)). By the order of a term \( u = \gamma y_j \) we mean the order of the element \( \gamma \in \Gamma \). Setting \( Y_j = \{y y_i \mid \gamma \in \Gamma_j, 1 \leq i \leq s\} \) \( (j = 1, \ldots, 2^n) \) we obtain a representation of the set of terms as a union \( Y = \bigcup_{j=1}^{2^n} Y_j \).

**Definition 2.3.5.** A term \( v \in Y \) is called a transform of a term \( u \in Y \) if and only if \( u \) and \( v \) belong to the same \( Y_j \) \( (1 \leq j \leq 2^n) \) and \( v = \gamma u \) for some \( \gamma \in \Gamma_j \). If \( \gamma \neq 1 \), \( v \) is said to be a proper transform of \( u \).

**Definition 2.3.6.** A well-ordering of the set of terms \( Y \) is called a ranking of the family of \( \sigma^* \)-indeterminates \( y_1, \ldots, y_s \) (or a ranking of the set \( Y \)) if it satisfies the following
distinct elements of an autoreduced set have distinct leaders and every autoreduced set is
will be always written in the order of increasing rank. If
is reduced with respect to
or
where
and for any two terms
,\ v \in Y_j \ (1 \leq j \leq 2^n), \ u \leq v \ and \ \gamma \in \Gamma_j, \ then \ \gamma u \leq \gamma v.

A ranking of the \( \sigma^*_\)-indeterminates \( y_1, \ldots, y_s \) is called orderly if for any \( j = 1, \ldots, 2^n \) and for any two terms \( u, v \in Y_j \), the inequality \( ord u < ord v \) implies that \( u < v \) (as usual, \( u < v \) means \( v \leq w \) and \( v \neq w \)). As an example of an orderly ranking of the \( \sigma^*_\)-indeterminates \( y_1, \ldots, y_s \) one can consider the standard ranking defined as follows:

\[
u = a_1^{\alpha_1} \cdots a_n^{\alpha_n} y_1 \leq v = a_1^{\alpha_1} \cdots a_n^{\alpha_n} y_j \quad \text{if and only if} \quad (\sum_{i=1}^n k_i, i, k_1, \ldots, k_n) \text{ is less than or equal to } (\sum_{i=1}^n |k_i|, j, l_1, \ldots, l_n) \quad \text{with respect to the lexicographic order on } \mathbb{Z}^{n+2}.
\]

In what follows, we assume that an orderly ranking \( \leq \) of the set of \( \sigma^*_\)-indeterminates \( y_1, \ldots, y_s \) has been fixed. If \( A \in F\{y_1, \ldots, y_s\}^* \), then the greatest (with respect to the ranking \( \leq \) term of \( Y \) that appears in \( A \) is called the leader of \( A \); it is denoted by \( u_A \). If
\[
d = \deg u_A,
\]
then the \( \sigma^*_\)-polynomial \( A \) can be written as \( A = I_d u^d + I_{d-1} u^{d-1} + \cdots + I_0 \) where \( I_k \ (0 \leq k \leq d) \) do not contain \( u \). The \( \sigma^*_\)-polynomial \( I_d \) is called the initial of \( A \); it is denoted by \( I_A \).

The ranking of the set of \( \sigma^*_\)-indeterminates \( y_1, \ldots, y_s \) generates the following relation on \( F\{y_1, \ldots, y_s\}^* \). If \( A \) and \( B \) are two \( \sigma^*_\)-polynomials, then \( A \) is said to have rank less than \( B \) (we write \( A \prec B \)) if either \( A \in F, B \notin F \) or \( A, B \in F\{y_1, \ldots, y_s\}^* \setminus F \) and \( u_A < u_B \) or \( u_A = u_B = u, \ deg u_A < deg u_B \). If \( u_A = u_B = u \) and \( \deg u_A = \deg u_B \), we say that \( A \) and \( B \) are of the same rank and write \( rk A = rk B \).

Let \( A, B \in F\{y_1, \ldots, y_s\}^* \). The \( \sigma^*_\)-polynomial \( A \) is said to be reduced with respect to \( B \) if \( A \) does not contain any power of a transform \( \gamma u_B (\gamma \in \Gamma_n) \) whose exponent is greater than \( \deg u_B \). If \( \Sigma \subseteq F\{y_1, \ldots, y_s\}^* \setminus F \), then a \( \sigma^*_\)-polynomial \( A \in F\{y_1, \ldots, y_s\}^* \), is said to be reduced with respect to \( \Sigma \) if \( A \) is reduced with respect to every element of the set \( \Sigma \).

A set \( \Sigma \subseteq F\{y_1, \ldots, y_s\}^* \) is said to be autoreduced if either it is empty or \( \Sigma \cap F = \emptyset \) and every element of \( \Sigma \) is reduced with respect to all others. As in the case of \( \sigma \)-polynomials, distinct elements of an autoreduced set have distinct leaders and every autoreduced set is finite. The following statement is an analog of Theorem 2.3.1.

**Theorem 2.3.7** [88, Theorem 3.4.27]. Let \( A = \{A_1, \ldots, A_r\} \) be an autoreduced subset of \( F\{y_1, \ldots, y_s\}^* \) and let \( D \in F\{y_1, \ldots, y_s\}^* \). Furthermore, let \( I(\mathcal{A}) = \{B \in F\{y_1, \ldots, y_s\} \mid \) either \( B = 1 \) or \( B \) is a product of finitely many polynomials of the form \( \gamma(I_{A_i}) (\gamma \in \Gamma_n, i = 1, \ldots, r) \). Then there exist \( \sigma \)-polynomials \( J \in I(\mathcal{A}) \) and \( D_0 \in F\{y_1, \ldots, y_s\} \) such that \( D_0 \) is reduced with respect to \( A \) and \( JD \equiv D_0 \mod [\mathcal{A}] \).

The transition from a \( \sigma^*_\)-polynomial \( D \) to a \( \sigma^*_\)-polynomial \( D_0 \) satisfying the conditions of the theorem can be performed in the same way as in the case of \( \sigma \)-polynomials (see the description of the corresponding reduction process after Theorem 2.3.1). We say that \( D \) reduces to \( D_0 \) modulo \( \mathcal{A} \).

As in the case of \( \sigma \)-polynomials, the elements of an autoreduced set in \( F\{y_1, \ldots, y_s\}^* \) will be always written in the order of increasing rank. If \( A = \{A_1, \ldots, A_r\} \) and \( B = \{B_1, \ldots, B_s\} \) are two autoreduced sets of \( \sigma^*_\)-polynomials, we say that \( A \) has lower rank
than \( B \) and write \( \text{rk} \, A < \text{rk} \, B \) if either there exists \( k \in \mathbb{N}, \, 1 \leq k \leq \min\{r, s\} \), such that 
\( \text{rk} \, A_i = \text{rk} \, B_i \) for \( i = 1, \ldots, k-1 \) and \( A_k < B_k \), or \( r > s \) and \( \text{rk} \, A_i = \text{rk} \, B_i \) for \( i = 1, \ldots, s \).

Repeating the proof of [86, Chapter 1, Proposition 3.1], one obtains that every family of autoreduced subsets of \( F[y_1, \ldots, y_s]^* \) contains an autoreduced set of lowest rank. In particular, if \( \emptyset \neq J \subseteq F[y_1, \ldots, y_s]^* \), then the set \( J \) contains an autoreduced set of lowest rank called a characteristic set of \( J \). The following statement is the version of Proposition 2.3.4 for inversive difference polynomials (see [88, Proposition 3.4.32]).

**Proposition 2.3.8.** Let \( F \) be a difference field with a basic set \( \sigma \), \( J \) a \( \sigma^* \)-ideal of the algebra of \( \sigma^* \)-polynomials \( F[y_1, \ldots, y_s]^* \), and \( \Sigma \) a characteristic set of \( J \). Then:

(i) The ideal \( J \) does not contain non-zero \( \sigma^* \)-polynomials reduced with respect to \( \Sigma \). In particular, if \( A \in \Sigma \), then \( I_A \notin J \).

(ii) If \( J \) is a prime \( \sigma^* \)-ideal, then \( J = [\Sigma]; \, \gamma(\Sigma) \) where \( \gamma(\Sigma) \) denote the set of all finite products of elements of the form \( \gamma(I_A) \) (\( \gamma \in \Gamma_\sigma, \, A \in \Sigma \)).

Let \( F \) be a difference field with a basic set \( \sigma \) and \( F[y_1, \ldots, y_s] \) an algebra of \( \sigma \)-polynomials in \( \sigma \)-indeterminates \( y_1, \ldots, y_s \) over \( F \). A \( \sigma \)-ideal \( I \) of \( F[y_1, \ldots, y_s] \) is called linear if it is generated (as a \( \sigma \)-ideal) by linear \( \sigma \)-polynomials (that is, \( \sigma \)-polynomials of the form \( \sum_{i=1}^{m} a_i \tau_i y_{k_i} \) where \( a_i \in F, \, \tau_i \in T_\sigma, \, 1 \leq k_i \leq s \) for \( i = 1, \ldots, m \)). If the \( \sigma \)-field \( F \) is inversive, then a \( \sigma^* \)-ideal of an algebra of \( \sigma^* \)-polynomials \( F[y_1, \ldots, y_s]^* \) is called linear if it is generated (as a \( \sigma^* \)-ideal) by linear \( \sigma^* \)-polynomials, i.e., \( \sigma^* \)-polynomials of the form \( \sum_{i=1}^{m} a_i \gamma_i y_{k_i} \) (\( a_i \in F, \, \gamma_i \in \Gamma_\sigma, \, 1 \leq k_i \leq s \) for \( i = 1, \ldots, m \)). As in the case of linear differential polynomials (see [88, Proposition 3.2.28]), one can show that if \( I \) is a proper linear \( \sigma \)-ideal of \( F[y_1, \ldots, y_s] \) or a proper linear \( \sigma^* \)-ideal of \( F[y_1, \ldots, y_s]^* \) then the ideal \( I \) is prime.

**Definition 2.3.9.** Let \( F \) be a difference field with a basic set \( \sigma \) and \( A \) an autoreduced set in \( F[y_1, \ldots, y_s] \) that consists of linear \( \sigma \)-polynomials (respectively, let \( F \) be a \( \sigma^* \)-field and \( A \) an autoreduced set in \( F[y_1, \ldots, y_s]^* \) that consists of linear \( \sigma^* \)-polynomials). The set \( A \) is called coherent if the following two conditions hold:

(i) If \( A \in A \) and \( \tau \in T_\sigma \) (respectively, \( \gamma \in \Gamma_\sigma \)), then \( \tau A \) (respectively, \( \gamma A \)) reduces to zero modulo \( A \).

(ii) If \( A, B \in A \) and \( v = \tau_1 u_A = \tau_2 u_B \) is a common transform of the leaders \( u_A \) and \( u_B \) (\( \tau_1, \tau_2 \in T_\sigma \) or \( \tau_1, \tau_2 \in \Gamma_\sigma \) if we consider the case of \( \sigma^* \)-polynomials), then the \( \sigma \)-polynomial \( (\tau_2 I_B)(\tau_1 A) - (\tau_1 I_A)(\tau_2 B) \) reduces to zero modulo \( A \).

The following result is proved in [88, Theorem 6.5.3] for autoreduced sets of inverse difference polynomials. The proof for the case of difference polynomials is similar.

**Theorem 2.3.10.** Let \( F \) be a difference field with a basic set \( \sigma \) and \( I \) a linear \( \sigma \)-ideal of the algebra of \( \sigma \)-polynomials \( F[y_1, \ldots, y_s] \) (respectively, let \( F \) be a \( \sigma^* \)-field and \( I \) a linear \( \sigma^* \)-ideal of \( F[y_1, \ldots, y_s]^* \)). Then any characteristic set of \( I \) is a coherent autoreduced set of linear \( \sigma \)- (respectively, \( \sigma^* \)-) polynomials.

Conversely, if \( A \subseteq F[y_1, \ldots, y_s] \) (respectively, \( A \subseteq F[y_1, \ldots, y_s]^* \)) is any coherent autoreduced set consisting of linear \( \sigma \)- (respectively, \( \sigma^* \)-) polynomials, then \( A \) is a characteristic set of the linear \( \sigma \)-ideal \([A] \) (respectively, of the linear \( \sigma^* \)-ideal \([A]^* \)).
COROLLARY 2.3.11. Let $F$ be an inversive difference field with a basic set $\sigma$ and let $\preceq$ be a preorder on $F\{y_1, \ldots, y_s\}^*$ such that for any two $\sigma^*$-polynomials $A_1$ and $A_2$, $A_1 \preceq A_2$ if and only if $u_{A_2}$ is a transform of $u_{A_1}$. Let $\Gamma$ be a linear $\sigma^*$-polynomial from $F\{y_1, \ldots, y_s\}^* \setminus F$ and $\Gamma_\sigma A = \{\gamma A \mid \gamma \in \Gamma_\sigma\}$. Then the set of all minimal (with respect to $\preceq$) elements of $\Gamma_\sigma A$ is a characteristic set of the $\sigma^*$-ideal $[A]^*$.

Theorem 2.3.10 implies the following method of constructing a characteristic set of a proper linear $\sigma^*$-ideal $I$ in $F\{y_1, \ldots, y_s\}^*$ (a similar method can be used for building a characteristic set of a $\sigma$-ideal in $F\{y_1, \ldots, y_s\}$). Suppose that $I = [A_1, \ldots, A_p]^*$ where $A_1, \ldots, A_p$ are linear $\sigma^*$-polynomials and $A_1 < \cdots < A_p$. It follows from Theorem 2.3.10 that one should find a coherent autoreduced set $\Phi \subseteq F\{y_1, \ldots, y_s\}^*$ such that $[\Phi]^* = I$. Such a set can be obtained from the set $A = \{A_1, \ldots, A_p\}$ via the following two-step procedure.

Step 1. Constructing an autoreduced set $\Sigma \subseteq I$ such that $[\Sigma]^* = I$.

If $A$ is autoreduced, set $\Sigma = A$. If $A$ is not autoreduced, choose the smallest $i$ ($1 \leq i \leq p$) such that some $\sigma^*$-polynomial $A_j$, $1 \leq i < j \leq p$, is not reduced with respect to $A_i$. Replace $A_j$ by its remainder with respect to $A_i$ (obtained by the procedure described after Theorem 2.3.1) and arrange the $\sigma^*$-polynomials of the new set $A_1$ in ascending order. Then apply the same procedure to the set $A_1$ and so on. After each iteration the number of $\sigma^*$-polynomials in the set does not increase, one of them is replaced by a $\sigma^*$-polynomial of lower or equal rank, and the others do not change. Therefore, the process terminates after a finite number of steps and then we have the desired autoreduced set $\Sigma$.

Step 2. Constructing a coherent autoreduced set $\Phi \subseteq I$.

Let $\Sigma_0 = \Sigma$ be an autoreduced subset of $I$ such that $[\Sigma]^* = I$. If $\Sigma$ is not coherent, we build a new autoreduced set $\Sigma_1 \subseteq I$ by adding to $\Sigma_0$ new $\sigma^*$-polynomials of the following types.

(a) $\sigma^*$-polynomials $(\gamma_1 I_{B_1})\gamma_2 B_2 - (\gamma_2 I_{B_2})\gamma_1 B_1$ constructed for every pair $B_1, B_2 \in \Sigma$ such that the leaders $u_{B_1}$ and $u_{B_2}$ have a common transform $v = \gamma_1 u_{B_1} = \gamma_2 u_{B_2}$ and $(\gamma_1 I_{B_1})\gamma_2 B_2 - (\gamma_2 I_{B_2})\gamma_1 B_1$ is not reducible to zero modulo $\Sigma_0$.

(b) $\sigma^*$-polynomials of the form $\gamma A$ ($\gamma \in \Gamma_\sigma, A \in \Sigma_0$) that are not reducible to zero modulo $\Sigma_0$.

It is clear that $rk \Sigma_1 < rk \Sigma_0$. Applying the same procedure to $\Sigma_1$ and continuing in the same way, we obtain autoreduced subsets $\Sigma_0, \Sigma_1, \ldots$ of $I$ such that $rk \Sigma_{i+1} < rk \Sigma_i$ for $i = 0, 1, \ldots$. Obviously, the process terminates after finitely many steps, and so we obtain an autoreduced set $\Phi \subseteq I$ such that $\Phi = \Sigma_k = \Sigma_{k+1} = \cdots$ for some $k \in \mathbb{N}$. It is easy to see that $\Phi$ is coherent, so it is a characteristic set of the ideal $I$.

2.4. Perfect difference ideals. Ritt difference rings

DEFINITION 2.4.1. Let $R$ be a difference ring with a basic set $\sigma$. A $\sigma$-ideal $J$ of the ring $R$ is called perfect if for any $a \in R$, $\tau_1, \ldots, \tau_r \in \Gamma_\sigma$ and $k_1, \ldots, k_r \in \mathbb{N}$, the inclusion $\tau_1(a)^k_1 \cdots \tau_r(a)^k_r \in J$ implies $a \in J$.

It is easy to see that every perfect ideal is reflexive and every reflexive prime ideal is perfect. Furthermore, if the $\sigma$-ring $R$ is inversive, then a $\sigma$-ideal $J$ of $R$ is perfect if and only
if any inclusion $γ_1(a)^{k_1} \cdots γ_r(a)^{k_r} ∈ J$ $(a ∈ R, γ_1, \ldots, γ_r ∈ Γ_σ, k_1, \ldots, k_r ∈ \mathbb{N})$ implies $a ∈ J$.

If $B$ is a subset of a difference ring $R$ with a basic set $σ$, then the intersection of all perfect $σ$-ideals of $R$ containing $B$ is the smallest perfect ideal containing $B$. It is denoted by $\{B\}$ and called the perfect closure of the set $B$. The ideal $\{B\}$ can be obtained from the set $B$ via the following procedure introduced in [137] and called shuffling. For any set $M ⊆ R$, let $M′$ denote the set of all $a ∈ R$ such that $τ_1(a)^{k_1} \cdots τ_r(a)^{k_r} ∈ M$ for some $τ_1, \ldots, τ_r ∈ T_σ$ and $k_1, \ldots, k_r ∈ \mathbb{N}$ $(r ≥ 1)$. Starting with $B_0 = B$, set $B_1 = [B_0]^′$, $B_2 = [B_1]^′$, and so on. Then $\{B\} = \bigcup_{i=0}^{∞} B_i$.

**Proposition 2.4.2 [32, Chapter 3, Section 2].** Let $A$ and $B$ be two subsets of a difference ring $R$. Then

(i) $A_k B_k ⊆ (AB)_k$ for any $k ∈ \mathbb{N}$. (By the product $UV$ of two sets $U, V ⊆ R$ we mean the set $UV = \{uv | u ∈ U, v ∈ V\}$.)

(ii) $\{A\}\{B\} ⊆ \{AB\}$.

(iii) $(AB)_k ⊆ A_k ∩ B_k$ for any $k ∈ \mathbb{N}, k ≥ 1$.

(iv) $A_k ∩ B_k ⊆ (AB)_{k+1}$ for any $k ∈ \mathbb{N}$.

(v) $\{A\} ∩ \{B\} = \{AB\}$.

Let $J$ be a subset of a difference ring $R$. Then a finite subset $A$ of $J$ is called a basis of $J$ if $\{A\} = \{J\}$. If $\{J\} = A_m$ for some $m ∈ \mathbb{N}$, $A$ is said to be an $m$-basis of $J$. A difference ring in which every subset has a basis is called a Ritt difference ring.

**Proposition 2.4.3 [32, Chapter 3, Theorem I].** A difference ring $R$ is a Ritt difference ring if and only if every perfect difference ideal of $R$ has a basis. If every perfect difference ideal of $R$ has an $m$-basis, then every set in $R$ has an $m$-basis. (In this and similar statements the number $m$ is not fixed but depends on the set.)

**Proposition 2.4.4 [32, Chapter 3, Theorem II].** A difference ring $R$ is a Ritt difference ring if and only if it satisfies the ascending chain condition for perfect difference ideals.

In [35] R.M. Cohn introduced and studied conservative systems of ideals of a commutative ring, that is, sets of ideals closed with respect to unions of linearly ordered subsets and intersections. The set of all perfect difference ideals of a difference ring $R$ is an example of such a system. If $R$ is a Ritt difference ring, then its perfect difference ideals form a Noetherian perfect conservative system, that is, a conservative system where every ideal coincides with its radical and every ascending chain of ideals is finite. A number of results that describe general properties of conservative systems can be also found in [86, Chapter 0, Section 7] and [88, Section 1.4].

The following statement is a version of the J. Ritt and H. Raudenbush theorem, [137], for partial difference rings.

**Theorem 2.4.5 [88, Theorem 3.3.42].** Let $R$ be a Ritt difference ring with a basic set $σ$ and let $S = R[η_1, \ldots, η_s]$ be a $σ$-overring of $R$ generated by a finite family of elements $[η_1, \ldots, η_s]$. Then $S$ is a Ritt $σ$-ring. Moreover, if every set in $R$ has an $m$-basis, then every
set in $S$ has an $m$-basis. In particular, an algebra of difference polynomials $R\{y_1, \ldots, y_s\}$ in a finite set of difference indeterminates $y_1, \ldots, y_s$ is a Ritt difference ring.

If $R$ is an inversive Ritt $\sigma$-ring and $S^* = R\{\eta_1, \ldots, \eta_s\}^*$ a finitely generated $\sigma^*$-overring of $R$, then $S^*$ is a Ritt $\sigma^*$-ring. If every set in $R$ has an $m$-basis, then every set in $S^*$ has an $m$-basis. In particular, an algebra of $\sigma^*$-polynomials in a finite set of $\sigma^*$-indeterminates over $R$ is a Ritt $\sigma^*$-ring.

It is not known whether the existence of bases of perfect difference ideals implies the existence of $m$-bases. Another open problem is to find out whether there is a positive integer $k$ such that every set in an algebra of difference polynomials over a difference field has a $k$-basis. In [32, Chapter 3, Section 13] R.M. Cohn showed that if such an integer $k$ exists, it exceeds 1. More precisely, R.M. Cohn proved that if $Q$ is considered as an ordinary difference field with the identity automorphism and $S = Q[u, v]$ is an algebra of difference polynomials in two difference indeterminates $u, v$ over $Q$, then the perfect difference ideal $\{uv\}$ of $S$ has no 1-basis.

The following statement strengthens the result of Theorem 2.4.5 for a wide class of rings of difference polynomials over ordinary difference fields.

**THEOREM 2.4.6 [21].** Let $F$ be an ordinary difference field containing an element $t$ which is distinct from all its non-trivial transforms. Let $S = F\{y_1, \ldots, y_s\}$ be a ring of difference polynomials in difference indeterminates $y_1, \ldots, y_s$ over $F$. Then every perfect difference ideal of $S$ has a basis consisting of $n + 1$ difference polynomials.

Let $R$ be a difference ring with a basic set $\sigma$ and $J$ a proper perfect $\sigma$-ideal of $R$. For every $x \in R \setminus J$, let $\mathcal{P}_x$ denote the set of all perfect $\sigma$-ideals $I$ of $R$ such that $J \subseteq I$ and $x \notin I$. By the Zorn lemma, the set $\mathcal{P}_x$ contains a maximal (relative to inclusion) perfect ideal $P_x$ ($\mathcal{P}_x \neq \emptyset$, since $J \in \mathcal{P}_x$). It follows from Proposition 2.4.2(ii) that the ideal $P_x$ is prime. Since $J = \bigcap_{x \in R \setminus J} P_x$, we obtain that every proper perfect $\sigma$-ideal of a difference ring can be represented as an intersection of prime reflexive $\sigma$-ideals. The following statement, the first version of which appeared in [137], specifies this result for Ritt difference rings. (Recall that a representation of a radical ideal $J$ of a commutative ring $R$ as a finite intersection of prime ideals, $J = \bigcap_{i=1}^r P_i$, is called *irredundant* if $P_i \not\subseteq P_j$ for $i \neq j$. Prime ideals $P_i$ from such a representation are called the *essential prime divisors of $J$*.)

**PROPOSITION 2.4.7.** Let $R$ be a difference ring with a basic set $\sigma$ and $J$ a proper perfect $\sigma$-ideal of $R$.

(i) There exists an irredundant representation of $J$ as an intersection of prime $\sigma$-ideals, $J = P_1 \cap \cdots \cap P_r$.

(ii) The $\sigma$-ideals $P_1, \ldots, P_r$ are reflexive and uniquely determined by the ideal $J$.

Let $R$ be a difference ring with a basic set $\sigma$ and $I, J$ two $\sigma$-ideals of $R$. The ideals $I$ and $J$ are called *separated* if $[I, J] = R$ and *strongly separated* if $[I, J] = R$. The $\sigma$-ideals $I_1, \ldots, I_r$ of $R$ are called *strongly separated in pairs* if any two of the ideals $I_i, I_j$ ($1 \leq i < j \leq r$) are (strongly) separated. It is easy to show that if the ideals $I_1, \ldots, I_r$ are strongly separated in pairs, then $I_{i+1} \cap \cdots \cap I_r = \prod_{i=1}^r I_i$. 

In [32, Chapter 3, Sections 11, 15] R.M. Cohn gives two examples that illustrate the relationships between some characteristics of difference ideals.

Let \( S = \mathbb{Q}[y] \) be the algebra of \( \sigma \)-polynomials in one \( \sigma \)-indeterminate \( y \) over \( \mathbb{Q} \) (treated as an ordinary difference field whose basic set consists of the identity isomorphism \( \alpha \)). The first example presents two separated \( \sigma \)-ideals that are not strongly separated. Let \( A = 1 + y\alpha(y), \, B = y + \alpha(y) \in S \). Then \( \{A, \, B\} = S \), but \( \{A, \, B\} \) is a proper ideal of the ring \( S \). (Moreover, even \([\{A, \, B\}]\) is a proper ideal of \( S \).)

The second example gives the irreducible representation of the perfect ideal \( \{y^2 - 1\} \) of \( S \) as an intersection of prime reflexive ideals: \( \{y^2 - 1\} = \{y - 1\} \cap \{y + 1\} \). At the same time, the \( \sigma \)-ideal \( \{y^2 - 1\} \) cannot be represented as an intersection (or product) of two \( \sigma \)-ideals whose perfect closures are \( \{y - 1\} \) and \( \{y + 1\} \). Moreover, \( \{y^2 - 1\} \) cannot be represented as an intersection (or product) of any two proper \( \sigma \)-ideals (such an ideal is called indecomposable).

We conclude this section with the description of two more types of difference ideals introduced and studied in [134] (see also [32, Chapter 3]). The corresponding results are formulated for partial difference rings.

Let \( R \) be a difference ring with a basic set \( \sigma \). A \( \sigma \)-ideal \( I \) of \( R \) is called \textit{complete} if for every element \( a \in \{I\} \), there exist \( \tau \in T_\sigma, \, k \in \mathbb{N} \) such that \( \tau(a)^k \in I \). It is easy to check that a \( \sigma \)-ideal is complete if and only if the presence in \( I \) of a product of powers of transforms of an element implies the presence in \( I \) of a power of a transform of the element. One can also define a complete \( \sigma \)-ideal as a \( \sigma \)-ideal \( I \) such that \( \{I\} \) is the reflexive closure of \( \sqrt{I} \) (here and below \( \sqrt{I} \) denotes the radical of \( I \)).

**Proposition 2.4.8** [134, Theorem II]. Let \( I \) be a complete difference ideal in a difference ring \( R \) with a basic set \( \sigma \). Suppose that \( \{I\} \) is the intersection of \( s \) perfect ideals \( J_1, \ldots, J_s \), which are strongly separated in pairs. Then there exist uniquely determined complete \( \sigma \)-ideals \( I_1, \ldots, I_s \) such that \( I = I_1 \cap \cdots \cap I_s \) and \( \{I_k\} = J_k \) (\( 1 \leq k \leq s \)). In this representation, the ideals \( I_1, \ldots, I_s \) are strongly separated in pairs. Furthermore, if the ideal \( I \) is reflexive, then so are \( I_1, \ldots, I_s \).

As we have seen, the ideal \( \{y^2 - 1\} \) of the ring of difference polynomials \( \mathbb{Q}[y] \) (\( \mathbb{Q} \) is treated as an ordinary difference ring with the identity translation) is indecomposable and its essential prime divisors \( \{y - 1\} \) and \( \{y + 1\} \) are strongly separated. In this case Proposition 2.4.8 shows that the difference ideal \( \{y^2 - 1\} \) is not complete.

Propositions 2.4.7 and 2.4.8 lead to the following decomposition theorem for complete difference ideals.

**Proposition 2.4.9** [32, Chapter 3, Theorem IX]. Let \( I \) be a proper complete difference ideal in a Ritt difference ring \( R \) with a basic set \( \sigma \). Then there exist proper complete \( \sigma \)-ideals \( I_1, \ldots, I_s \) of \( R \) such that

(i) \( I = I_1 \cap \cdots \cap I_s \),

(ii) \( I_1, \ldots, I_s \) are strongly separated in pairs (hence, \( I = I_1I_2 \cdots I_s \)).

(iii) no \( I_k \) (\( 1 \leq k \leq s \)) is the intersection of two strongly separated proper \( \sigma \)-ideals.

The ideals \( I_1, \ldots, I_s \) are uniquely determined. If \( I \) is reflexive, so are the \( I_k \) (\( 1 \leq k \leq s \)).
The ideals $I_1, \ldots, I_s$ whose existence is established by Proposition 2.4.9 are called the \textit{essential strongly separated divisors} of the $\sigma$-ideal $I$. The following result is a consequence of Proposition 2.4.9.

**Proposition 2.4.10.** Let $R$ be a Ritt difference ring with a basic set $\sigma$ and let $J$ be a proper perfect $\sigma$-ideal of $R$. Then

(i) the essential strongly separated divisors of $J$ are perfect $\sigma$-ideals;
(ii) the ideal $J$ can be represented as $J = J_1 \cap \cdots \cap J_s$ where $J_1, \ldots, J_s$ are proper perfect $\sigma$-ideals separated in pairs and no $J_k (1 \leq k \leq s)$ can be represented as an intersection of two separated proper perfect $\sigma$-ideals. The ideals $J_1, \ldots, J_s$ are uniquely determined by the ideal $J$. (They are called the essential separated divisors of $J$.)

A difference ideal $I$ of a difference ($\sigma$-) ring $R$ is called \textit{mixed} if the inclusion $ab \in I$ $(a,b \in R)$ implies that $a \alpha(b) \in I$ for every $\alpha \in \sigma$. It is easy to see that every perfect ideal is mixed and every mixed ideal is complete.

Let $Q[y]$ be the ring of difference polynomials in one difference indeterminate $y$ over $Q$ (treated as an ordinary difference ring with the identity translation $\alpha$). In [32, Chapter 3, Section 21] R.M. Cohn showed that the difference ideal $[y\alpha(y)]$ is not complete, while $[y^2]$ is complete, but not mixed. He also proved that if the ideal $I$ in the hypothesis of Proposition 2.4.9 is mixed, then so are the $I_1, \ldots, I_s$. Thus, the essential strongly separated divisors of a mixed difference ideal in a Ritt difference ring are mixed difference ideals.

### 2.5. Varieties of difference polynomials

Let $F$ be a difference field with a basic set $\sigma$, $F\{y_1, \ldots, y_s\}$ an algebra of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$ and $E$ a family of $\sigma$-overfields of $F$. Furthermore, let $\Phi \subseteq F\{y_1, \ldots, y_s\}$ and let $M_E(\Phi)$ denote the set of all $s$-tuples $a = (a_1, \ldots, a_s)$ with coordinates from some field $F_a \in E$ which are solutions of the set $\Phi$ (that is, $f(a_1, \ldots, a_s) = 0$ for any $f \in \Phi$). Then $M_E(\Phi)$ is said to be the $E$-variety defined by the set $\Phi$ (it is also called the $E$-variety of the set $\Phi$ over $F\{y_1, \ldots, y_s\}$ or over $F$). The $\sigma$-field $F$ is said to be the \textit{ground $\sigma$-field} of the $E$-variety.

Now, let $\mathcal{M}$ be a set of $s$-tuples such that the coordinates of every point $a \in \mathcal{M}$ belong to some field $F_a \in E$. If there exists a set $\Phi \subseteq F\{y_1, \ldots, y_s\}$ such that $\mathcal{M} = M_E(\Phi)$, then $\mathcal{M}$ is said to be an $E$-variety over $F\{y_1, \ldots, y_s\}$ (or over $F$).

**Definition 2.5.1.** Let $F$ be a difference field with a basic set $\sigma$, $G = F(x_1, x_2, \ldots)$ the field of rational fractions in a denumerable set of indeterminates $x_1, x_2, \ldots$ over $F$, and $\overline{G}$ the algebraic closure of $G$. Then the family $U(F)$ of all $\sigma$-overfields of $F$ which are defined on subfields of $\overline{G}$ is called the universal system of $\sigma$-overfields of $F$. If $\Phi$ is a subset of the algebra of $\sigma$-polynomials $F\{y_1, \ldots, y_s\}$ and $U = U(F)$, then the $U$-variety $M_U(\Phi)$ (also denoted by $M(\Phi)$) is called the variety defined by the set $\Phi$ over $F$ (or over $F\{y_1, \ldots, y_s\}$).
A set of $s$-tuples $M$ over the field $F$ is said to be a variety over $F\{y_1,\ldots,y_s\}$ (or a variety over $F$) if there exists a set $\Phi \subseteq F\{y_1,\ldots,y_s\}$ such that $M = \mathcal{M}_{\mathcal{U}(F)}(\Phi)$.

In what follows, we assume that a difference field $F$ with a basic set $\sigma$, an algebra of $\sigma$-polynomials $F\{y_1,\ldots,y_s\}$, and a family $E$ of $\sigma$-overfields of $F$ are fixed. $E$-varieties and varieties over $F\{y_1,\ldots,y_s\}$ will be called simply $E$-varieties and varieties, respectively.

**Proposition 2.5.2** [*32, Chapter 4, Section 3*]. Let $\eta = (\eta_1,\ldots,\eta_s)$ be an $s$-tuple over the $\sigma$-field $F$. Then there exists and $s$-tuple $\xi = (\xi_1,\ldots,\xi_s)$ over $F$ such that $\xi$ is equivalent to $\eta$ and all $\xi_i$ ($1 \leq i \leq s$) belong to some field from the universal system $\mathcal{U}(F)$.

If $A_1$ and $A_2$ are two $E$-varieties and $A_1 \subseteq A_2$ ($A_1 \not\subseteq A_2$), then $A_1$ is said to be a $E$-subvariety (respectively, a proper $E$-subvariety) of $A_2$. $\mathcal{U}(F)$-subvarieties of a variety $A$ are called subvarieties of $A$. An $E$-variety (variety) $A$ is called reducible if it can be represented as a union of two its proper $E$-subvarieties (subvarieties). If such a representation does not exist, the $E$-variety (variety) $A$ is said to be irreducible.

Let an $E$-variety (variety) $A$ be represented as a union of its $E$-subvarieties (subvarieties): $A = A_1 \cup \cdots \cup A_k$. This representation is called irredundant if $A_i \not\subseteq A_j$ for $i \neq j$ ($1 \leq i, j \leq k$).

The following proposition summarizes basic properties of $E$-varieties. As before, we assume that a family $E$ of $\sigma$-overfields of $F$ is fixed. Furthermore, if $A$ is a set of $s$-tuples $a$ with coordinates from a $\sigma$-field $F_a \in E$ (we say that $A$ is a set of $s$-tuples from $E$ over $F$), then $\Phi_E(A)$ denotes the perfect $\sigma$-ideal $\{ f \in F\{y_1,\ldots,y_s\} | f(a_1,\ldots,a_s) = 0 \}$ for any $a = (a_1,\ldots,a_s) \in A$ of the ring $F\{y_1,\ldots,y_s\}$.

**Proposition 2.5.3.**
(i) If $\Phi_1 \subseteq \Phi_2 \subseteq F\{y_1,\ldots,y_s\}$, then $\mathcal{M}_E(\Phi_2) \subseteq \mathcal{M}_E(\Phi_1)$.
(ii) If $A_1$ and $A_2$ are two sets of $s$-tuples from $E$ over $F$ and $A_1 \subseteq A_2$, then $\Phi_E(A_2) \subseteq \Phi_E(A_1)$.
(iii) If $A$ is an $E$-variety, then $A = \mathcal{M}_E(\Phi_E(A))$.
(iv) If $J_1,\ldots,J_k$ are $\sigma$-ideals of the ring $F\{y_1,\ldots,y_s\}$ and $J = J_1 \cap \cdots \cap J_k$, then $\mathcal{M}_E(J) = \mathcal{M}_E(J_1) \cup \cdots \cup \mathcal{M}_E(J_k)$.
(v) If $A_1,\ldots,A_k$ are $E$-varieties over $F$ and $A = A_1 \cup \cdots \cup A_k$, then $A$ is an $E$-variety over $F$ and $\Phi_E(A) = \Phi_E(A_1) \cap \cdots \cap \Phi_E(A_k)$.
(vi) The intersection of any family of $E$-varieties is an $E$-variety.
(vii) An $E$-variety $A$ is irreducible if and only if $\Phi_E(A)$ is a prime reflexive ideal of $F\{y_1,\ldots,y_s\}$.
(ix) Every $E$-variety $A$ has a unique irredundant representation as a union of irreducible $E$-varieties, $A = A_1 \cup \cdots \cup A_k$. (The $E$-varieties $A_1,\ldots,A_k$ are called irreducible $E$-components of $A$.) Furthermore, $A_i \not\subseteq \bigcup_{j \neq i} A_j$ for $i = 1,\ldots,k$.
(x) If $A_1,\ldots,A_k$ are the irreducible $E$-components of an $E$-variety $A$, then the prime $\sigma^+$-ideals $\Phi_E(A_1),\ldots,\Phi_E(A_k)$ are the essential prime divisors of the perfect $\sigma$-ideal $\Phi_E(A)$.

The last proposition implies that $A \mapsto \Phi_E(A)$ is an injective mapping of the set of all $E$-varieties over $F\{y_1,\ldots,y_s\}$ into a set of all perfect $\sigma$-ideals of the ring $F\{y_1,\ldots,y_s\}$. If $E$
is the universal system of $\sigma$-overfields of $F$, then this mapping is bijective. More precisely, we have the following statement about varieties (see [32, Chapter 4, Section 5]).

**Proposition 2.5.4.**

(i) If $J$ is a perfect $\sigma$-ideal of the ring $F[y_1, \ldots, y_s]$, then $\Phi(M(J)) = J$.
(ii) $M(J) = \emptyset$ if and only if $J = F[y_1, \ldots, y_s]$.
(iii) The mappings $\mathcal{A} \mapsto \Phi(\mathcal{A})$ and $P \mapsto \mathcal{M}(P)$ are two mutually inverse mappings that establish one-to-one correspondence between the set of all varieties over $F$ and the set of all perfect $\sigma$-ideals of the ring $F[y_1, \ldots, y_s]$.
(iv) The correspondence $\mathcal{A} \mapsto \Phi(\mathcal{A})$ maps irreducible components of an arbitrary variety $\mathcal{B}$ onto essential prime divisors of the perfect $\sigma$-ideal $\Phi(\mathcal{B})$ in $F[y_1, \ldots, y_s]$. In particular there is a one-to-one correspondence between irreducible varieties over $F$ and prime $\sigma$-ideals of the $\sigma$-ring $F[y_1, \ldots, y_s]$.

If $\mathcal{A}$ is an irreducible variety over $F[y_1, \ldots, y_s]$, then a generic zero of the corresponding prime ideal $\Phi(\mathcal{A})$ is called a generic zero of the variety $\mathcal{A}$.

The following result is a version of the Hilbert’s Nullstellensatz for difference fields.

**Theorem 2.5.5.** Let $F$ be a difference field with a basic set $\sigma$ and $F[y_1, \ldots, y_s]$ an algebra of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$. Let $f \in F[y_1, \ldots, y_s]$, $\Phi \subseteq F[y_1, \ldots, y_s]$, and $\mathcal{M}(\Phi)$ the variety defined by the set $\Phi$ over $F$. Then the following conditions are equivalent.

(i) Every $s$-tuple from $\mathcal{M}(\Phi)$ is a solution of the $\sigma$-polynomial $f$.
(ii) Every $s$-tuple $\eta = (\eta_1, \ldots, \eta_s) \in \mathcal{M}(\Phi)$ which is $\sigma$-algebraic over $F$ is a solution of $f$.
(iii) $f \in \{\Phi\}$.

Two varieties $\mathcal{A}_1$ and $\mathcal{A}_2$ over the ring of difference polynomials $F[y_1, \ldots, y_s]$ are called separated if $\mathcal{A}_1 \cap \mathcal{A}_2 = \emptyset$.

If a variety $\mathcal{A}$ is represented as a union of pairwise separated varieties $\mathcal{A}_1, \ldots, \mathcal{A}_k$ and no $\mathcal{A}_i$ is the union of two non-empty separated varieties, then $\mathcal{A}_1, \ldots, \mathcal{A}_k$ are said to be *essential separated components* of $\mathcal{A}$. (All varieties are considered over the same ring $F[y_1, \ldots, y_s]$.)

The following statement is due to R.M. Cohn (see [32, Chapter 4, Section 7]).

**Proposition 2.5.6.** Let $F$ be a difference field with a basic set $\sigma$ and $F[y_1, \ldots, y_s]$ an algebra of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$.

(i) Two varieties $\mathcal{A}_1$ and $\mathcal{A}_2$ over $F[y_1, \ldots, y_s]$ are separated if and only if the perfect $\sigma$-ideals $\Phi(\mathcal{A}_1)$ and $\Phi(\mathcal{A}_2)$ are separated.
(ii) Two perfect ideals $J_1$ and $J_2$ of the ring $F[y_1, \ldots, y_s]$ are separated if and only if the varieties $\mathcal{M}(J_1)$ and $\mathcal{M}(J_2)$ are separated.
(iii) Every non-empty variety $\mathcal{A}$ over $F[y_1, \ldots, y_s]$ can be represented as a union of a uniquely determined family of its essential separated components. Each of these components is a union of some irreducible components of $\mathcal{A}$.
(iv) If $A_1, \ldots, A_k$ are essential separated components of a variety $A$, then $\Phi(A_1), \ldots, \Phi(A_k)$ are essential separated divisors of the perfect $\sigma$-ideal $\Phi(A)$ in $F[y_1, \ldots, y_s]$.

Now, let $F$ be an inverse difference field with a basic set $\sigma$, $F[y_1, \ldots, y_s]$ an algebra of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$, and $E$ a set of $\sigma$-overfields of $F$. If $\Phi \subseteq F[y_1, \ldots, y_s]^*$, then the set $M_E(\Phi)$ consisting of all $s$-tuples $a$ which have coordinates in some field $F_a \in E$ and annul every $\sigma$-polynomial from $\Phi$ is called an $E$-variety over $F[y_1, \ldots, y_s]^*$ determined by the set $\Phi$.

Let $A$ be a set of $s$-tuples over $F$ such that all coordinates of each $s$-tuple $a \in A$ belong to some $\sigma$-field $F_a \in E$. Then $A$ is said to be an $E$-variety over $F[y_1, \ldots, y_s]^*$ if there exists a set $\Phi \subseteq F[y_1, \ldots, y_s]^*$ such that $A = M_E(\Phi)$.

Let $L = F(x_1, x_2, \ldots)$ be the field of rational fractions in a denumerable set of indeterminates $x_1, x_2, \ldots$ over the $\sigma$-field $F$ and let $\bar{L}$ be the algebraic closure of $L$. Then the family $U^*(F)$ consisting of all $\sigma$-overfields of $F$ defined on subfields of $\bar{L}$ is called the universal system of $\sigma$-overfields of $F$. As in the case of non-inverse difference fields, one can prove that if $\eta = (\eta_1, \ldots, \eta_s)$ is any $s$-tuple over the $\sigma$-field $F$, then there exists an $s$-tuple $\zeta = (\zeta_1, \ldots, \zeta_s)$ such that $\zeta$ is equivalent to $\eta$ and all coordinates of the point $\zeta$ lie in some $\sigma$-field $F_\zeta \in U^*(F)$ (see [88, Proposition 3.4.34]). A $U^*(F)$-variety over $F[y_1, \ldots, y_s]^*$ is called a variety over this ring of $\sigma$-polynomials.

The concepts of $E$-subvariety, proper $E$-subvariety, subvariety, and proper subvariety over $F[y_1, \ldots, y_s]^*$, as well as the notions of reducible and irreducible $E$-varieties and varieties, are precisely the same as in the case of $s$-tuples over an algebra of (non-inverse) difference polynomials. If an $E$-variety (variety) $A$ over $F[y_1, \ldots, y_s]^*$ is represented as a union of its $E$-subvarieties (subvarieties), $A = A_1 \cup \cdots \cup A_k$, and $A_i \nsubseteq A_j$ for $i \neq j$ ($1 \leq i, j \leq k$), then this representation is called irredundant.

All properties of $E$-varieties and varieties over an algebra of difference polynomials listed in Propositions 2.5.3, 2.5.4, 2.5.6 and Theorem 2.5.5 are also valid for $E$-varieties and varieties over $F[y_1, \ldots, y_s]^*$. The formulations of the corresponding statements are practically the same. One should just replace the ring $F[y_1, \ldots, y_s]$ by $F[y_1, \ldots, y_s]^*$ and treat $M_E(\Phi)$ ($\Phi \subseteq F[y_1, \ldots, y_s]^*$) and $\Phi_E(A)$ ($A$ is a set of $s$-tuples $a$ over $F$ whose coordinates belong to some $\sigma$-field $F_a \in E$) as the set $\{a = (a_1, \ldots, a_s) \mid a_1, \ldots, a_s$ belong to some field $F_a \in E$ and $f(a_1, \ldots, a_s) = 0$ for all $f \in \Phi$ and the perfect $\sigma$-ideal $\{f \in F[y_1, \ldots, y_s]^* \mid f(a_1, \ldots, a_s) = 0$ for any $a = (a_1, \ldots, a_s) \in A\}$ of the ring $F[y_1, \ldots, y_s]^*$, respectively. (If $E = U^*(F)$, then $\Phi_E(A)$ and $M_E(\Phi)$ are denoted by $M(\Phi)$ and $\Phi(A)$, respectively.) By a generic zero of an irreducible variety $A$ over $F[y_1, \ldots, y_s]^*$ we mean a generic zero of the corresponding perfect $\sigma$-ideal $\Phi(A)$ of the ring $F[y_1, \ldots, y_s]^*$. The concept of separated varieties over $F[y_1, \ldots, y_s]^*$ is introduced in the same way as in the case of varieties over an algebra of difference polynomials.

A family $E$ of difference overfields of a difference ($\sigma$-) field $F$ is called a complete system of $\sigma$-overfields of $F$ if distinct perfect $\sigma$-ideals of any ring of $\sigma$-polynomials over $F$ have distinct $E$-varieties. Clearly, the universal system of $\sigma$-overfields of $F$ is complete.

The proof of the following result can be found in [32, Chapter 8].
PROPOSITION 2.5.7. Let $F$ be an ordinary difference field with a basic set $\sigma = \{\alpha\}$.

(i) There exists a complete system $C$ of $\sigma$-overfields of $F$ where each $\sigma$-field is $\sigma$-algebraic over $F$. If $F$ is algebraically closed, $C$ may be chosen to consist of one $\sigma$-field.

(ii) Let the $\sigma$-field $F$ be aperiodic (that is, there is no $n \in \mathbb{N}$ such that $\alpha^n(a) = a$ for all $a \in F$) and $\text{Char} F = 0$. Let $\mathcal{E}$ be a family of difference overfields of $F$ and let $F\{y\}$ be a ring of $\sigma$-polynomials in one $\sigma$-indeterminate $y$ over $F$. Then $\mathcal{E}$ is a complete system if and only if given any prime $\sigma^*$-ideal $P$ of $F\{y\}$ and any $\sigma$-polynomial $A \in F\{y\} \setminus P$, $\mathcal{M}_E(P)$ contains a solution not annulling $A$.

We conclude this section with a brief discussion of a realization of an abstract variety of difference polynomials as a set of complex-valued functions. The following results are due to R.M. Cohn, [33], who provided such a realization by means of an existence theorem yielding solutions of difference equations as complex-valued functions defined, except for isolated singularities, on the non-negative real axis.

A complex-valued function $f(x)$ is said to be permitted function if it is defined for all real values $x \geq 0$ except at a set $S(f)$ which has no limit points, is analytic in each of the intervals into which the non-negative real axis is divided by omission of the points of $S(f)$, and is either identically 0 or is 0 at only finitely many points in any finite interval. A permitted difference ring is an ordinary difference ring $R$ whose elements are permitted functions and whose basic set consists of the translation $\alpha$ such that $(\alpha f)(x) = f(x + 1)$ for any $f(x) \in R$. (More precisely, the elements of $R$ are equivalence classes of permitted functions, with $f(x)$ equivalent to $g(x)$ if they coincide except possibly on $S(f) \cup S(g)$.) It follows from the definition of permitted functions that $\alpha$ is an isomorphism of $R$ into itself.) A permitted difference ring which is a field is called a permitted difference field.

Let $K_0$ be the ordinary difference field of rational functions $f : \mathbb{R} \to \mathbb{C}$ with complex coefficients whose basic set consists of the translation $\alpha : f(x) \mapsto f(x + 1)$ ($f(x) \in K_0$). Then $K_0$ may be regarded as a permitted difference field by restricting the domain of its members to $x \geq 0$. In what follows, $\mathcal{K}$ denotes the set of permitted difference overfields of $K_0$ and $\mathcal{K}^* = \{K \in \mathcal{K} \mid$ there exists an infinite set of functions analytic throughout $[0, 1]$ which is algebraically independent over $K$ (regarded as a field of functions over $[0, 1])\}$.

The following existence theorem was proved in [33] where one can also find a discussion of the existence of continuous solutions of reflexive prime ideals in $K_0\{y\}$ and $K\{y\}$ ($K \in \mathcal{K}^*$).

THEOREM 2.5.8. With the above notation, let $K \in \mathcal{K}^*$ and let $J$ be a proper reflexive prime ideal of the ring of difference polynomials $K\{y\}$ in one difference indeterminate $y$. Then

(i) $\mathcal{K}$ is a complete system of difference overfields of $K$.

(ii) The ideal $J$ has a generic zero in one of the members of $\mathcal{K}$. 

3. Difference modules

3.1. Ring of difference operators. Difference modules

Let \( R \) be a difference ring with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \) and \( T \) the free commutative semigroup generated by the elements \( \alpha_1, \ldots, \alpha_n \). Furthermore, for any \( r \in \mathbb{N} \), let \( T(r) = \{\tau \in T \mid \text{ord} \tau \leq r\} \) (as before, by the order of an element \( \tau = \alpha_1^{k_1} \cdots \alpha_n^{k_n} \in T \) we mean the number \( \text{ord} \tau = \sum_{i=1}^{n} k_i \)).

**Definition 3.1.1.** An expression of the form \( \sum_{\tau \in T} a_{\tau} \tau \), where \( a_{\tau} \in R \) for any \( \tau \in T \) and only finitely many elements \( a_{\tau} \) are different from 0, is called a difference (or \( \sigma \)-) operator over the difference ring \( R \). Two \( \sigma \)-operators \( \sum_{\tau \in T} a_{\tau} \tau \) and \( \sum_{\tau \in T} b_{\tau} \tau \) are considered to be equal if and only if \( a_{\tau} = b_{\tau} \) for any \( \tau \in T \).

Let \( \mathcal{D} \) denote the set of all \( \sigma \)-operators over the \( \sigma \)-ring \( R \). This set can be equipped with a ring structure if we set \( \sum_{\tau \in T} a_{\tau} \tau + \sum_{\tau \in T} b_{\tau} \tau = \sum_{\tau \in T} (a_{\tau} + b_{\tau}) \tau \), \( a \sum_{\tau \in T} a_{\tau} \tau = \sum_{\tau \in T} (aa_{\tau}) \tau \), \( (\sum_{\tau \in T} a_{\tau} \tau) \tau_1 = \sum_{\tau \in T} a_{\tau} (\tau \tau_1) \), \( \tau_1 a = \tau_1 (a \tau) \tau_1 \) for any \( \sum_{\tau \in T} a_{\tau} \tau \), \( \sum_{\tau \in T} b_{\tau} \tau \in \mathcal{D} \), \( a \in R \), \( \tau_1 \in T \), and extend the multiplication by distributivity. The ring obtained in this way is called the ring of difference (or \( \sigma \)-) operators over \( R \).

The order of a \( \sigma \)-operator \( A = \sum_{\tau \in T} a_{\tau} \tau \in \mathcal{D} \) is defined as the number \( \text{ord} A = \max\{\text{ord} \tau \mid a_{\tau} \neq 0\} \). If for any \( q \in \mathbb{N} \) we define \( \mathcal{D}^{(q)} = \{\sum_{\tau \in T} a_{\tau} \tau \in \mathcal{D} \mid \text{ord} \tau = q\} \) for every \( \tau \in T \) such that \( a_{\tau} \neq 0 \) and set \( \mathcal{D}^{(q)} = 0 \) for any \( q \in \mathbb{Z}, q < 0 \), then the ring \( \mathcal{D} \) can be considered as a graded ring (with positive grading): \( \mathcal{D} = \bigoplus_{q \in \mathbb{Z}} \mathcal{D}^{(q)} \). It can be also treated as a filtered ring with the ascending filtration \( (\mathcal{D}_r)_{r \in \mathbb{Z}} \) such that \( \mathcal{D}_r = 0 \) for any \( r < 0 \) and \( \mathcal{D}_r = \{ A \in \mathcal{D} \mid \text{ord} A \leq r\} \) for any \( r \in \mathbb{N} \). Below, while considering \( \mathcal{D} \) as a graded or filtered ring, we always mean the grading with the homogeneous components \( \mathcal{D}^{(q)} \) \((q \in \mathbb{Z})\) or the filtration \( (\mathcal{D}_r)_{r \in \mathbb{Z}} \), respectively.

**Definition 3.1.2.** Let \( R \) be a difference ring with a basic set \( \sigma \) and \( \mathcal{D} \) the ring of \( \sigma \)-operators over \( R \). Then a left \( \mathcal{D} \)-module is called a difference \( R \)-module or a \( \sigma \)-\( R \)-module. In other words, an \( R \)-module \( M \) is a difference (or \( \sigma \)-) \( R \)-module, if the elements of \( \sigma \) act on \( M \) in such a way that \( \alpha(x + y) = \alpha(x) + \alpha(y), \alpha(\beta x) = \beta(\alpha x), \) and \( \alpha(\alpha x) = \alpha(a)\alpha(x) \) for any \( x, y \in M, \alpha, \beta \in \sigma, a \in R \).

If \( R \) is a difference (\( \sigma \)-) field, then a \( \sigma \)-\( R \)-module \( M \) is also called a difference vector space over \( R \) or a vector \( \sigma \)-\( R \)-space.

We say that a difference \( R \)-module \( M \) is finitely generated, if it is finitely generated as a left \( \mathcal{D} \)-module. By a graded difference (\( \sigma \)-) \( R \)-module we always mean a graded left module over the ring of \( \sigma \)-operators \( \mathcal{D} = \bigoplus_{q \in \mathbb{Z}} \mathcal{D}^{(q)} \). If \( M = \bigoplus_{q \in \mathbb{Z}} M^{(q)} \) is a graded \( \sigma \)-\( R \)-module and \( M^{(q)} = 0 \) for all \( q < 0 \), we say that \( M \) is positively graded and write \( M = \bigoplus_{q \in \mathbb{N}} M^{(q)} \).

Let \( R \) be a difference ring with a basic set \( \sigma \) and \( \mathcal{D} \) the ring of \( \sigma \)-operators over \( R \) equipped with the ascending filtration \( (\mathcal{D}_r)_{r \in \mathbb{Z}} \). In what follows, by a filtered \( \sigma \)-\( R \)-module we always mean a left \( \mathcal{D} \)-module equipped with an exhaustive and separated filtration. Thus, a filtration of a \( \sigma \)-\( R \)-module \( M \) is an ascending chain \( (M_r)_{r \in \mathbb{Z}} \) of \( R \)-submodules of
Let $R$ be an Artinian difference ring with a basic set $\sigma = \{a_1, \ldots, a_n\}$ and $M = \bigoplus_{q \in \mathbb{Z}} M^{(q)}$ a finitely generated positively graded $\sigma$-R-module. Then

(i) the length $l_R(M^{(q)})$ of every $R$-module $M^{(q)}$ is finite;
(ii) there exists a polynomial $\phi(t) \in \mathbb{Q}[t]$ such that $\phi(q) = l_R(M^{(q)})$ for all sufficiently large $q \in \mathbb{N}$ (i.e., there exists $q_0 \in \mathbb{N}$ such that the equality holds for all $q \geq q_0$);
(iii) $\deg \phi(t) \leq n - 1$ and the polynomial $\phi(t)$ can be written as $\phi(t) = \sum_{i=0}^{n-1} a_i t^i$ where $a_0, a_1, \ldots, a_{n-1} \in \mathbb{Z}$.

Let us consider the ring $R$ as a filtered ring with the trivial filtration $(R_r)_{r \in \mathbb{Z}}$ such that $R_r = R$ for all $r \geq 0$ and $R_r = 0$ for any $r < 0$. Let $P$ be an $R$-module and let $(P_r)_{r \in \mathbb{Z}}$ be a non-descending chain of $R$-submodules of $P$ such that $\bigcup_{r \in \mathbb{Z}} P_r = P$ and $P_r = 0$ for all sufficiently small $r \in \mathbb{Z}$. Then $P$ can be treated as a filtered $R$-module with the filtration $(P_r)_{r \in \mathbb{Z}}$ and the left $D$-module $D \otimes_R P$ can be considered as a filtered $\sigma$-R-module with the filtration $((D \otimes_R P)_r)_{r \in \mathbb{Z}}$ where $(D \otimes_R P)_r$ is the $R$-submodule of $D \otimes_R P$ generated by the set $\{u \otimes x \mid u \in D_i \text{ and } x \in P_{r-i}, 0 \leq i \leq r\}$. In what follows, while considering $D \otimes_R P$ as a filtered $\sigma$-R-module ($P$ is an exhaustively and separately filtered module over the $\sigma$-ring $R$ with the trivial filtration) we shall always mean the filtration $((D \otimes_R P)_r)_{r \in \mathbb{Z}}$.

Theorem 3.1.4 [88, Theorem 6.2.5]. Let $R$ be an Artinian difference ring with a basic set $\sigma = \{a_1, \ldots, a_n\}$ and let $(M_r)_{r \in \mathbb{Z}}$ be an excellent filtration of a $\sigma$-R-module $M$. Then there exists a polynomial $\psi(t) \in \mathbb{Q}[t]$ such that $\psi(r) = l_R(M_r)$ for all sufficiently large $r \in \mathbb{Z}$. Furthermore, $\deg \psi(t) \leq n$ and the polynomial $\psi(t)$ can be written as $\psi(t) = \sum_{i=0}^{n} c_i t^i$ where $c_0, c_1, \ldots, c_n \in \mathbb{Z}$.
The polynomial \( \psi(t) \) whose existence is established by theorem 3.1.4 is called the difference \((\sigma\cdot)\) dimension polynomial or characteristic polynomial of the module \( M \) associated with the excellent filtration \((M_r)_{r \in \mathbb{Z}}\).

**Example 3.1.5.** Let \( R \) be a difference field with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \) and \( D \) the ring of \( \sigma \)-operators over \( R \) treated as a filtered \( \sigma\cdot R \)-module with the excellent filtration \((D_r)_{r \in \mathbb{Z}}\). If \( r \in \mathbb{N} \), then the elements \( \alpha_1^{k_1}, \ldots, \alpha_n^{k_n}, \) where \( k_1, \ldots, k_n \in \mathbb{N} \) and \( \sum_{i=1}^{n} k_i \leq r \), form a basis of the vector \( R \)-space \( D_r \). Therefore, \( l_R(D_r) = \dim_R D_r = \text{Card}(k_1, \ldots, k_n) \in \mathbb{N}^n \mid k_1 + \cdots + k_n \leq r \rangle = \binom{r+n}{n} \) whence \( \psi_D(t) = \binom{t+n}{n} \) is the characteristic polynomial of the ring \( D \) associated with the filtration \((D_r)_{r \in \mathbb{Z}}\).

Let \( R \) be a difference ring with a basic set \( \sigma \), \( D \) the ring of \( \sigma \)-operators over \( R \), \( M \) a filtered \( \sigma\cdot R \)-module with a filtration \((M_r)_{r \in \mathbb{Z}}\), and \( R[x] \) the ring of polynomials in one indeterminate \( x \) over \( R \). Let \( D' \) denote the subring \( \sum_{r \in \mathbb{N}} D_r \otimes_R R[x] \) of the ring \( D \otimes_R R[x] \) and \( M' \) denote the left \( D' \)-module \( \sum_{r \in \mathbb{N}} M_r \otimes_R R[x] \). The proof of the following three results can be found in [88, Section 6.2].

**Lemma 3.1.6.** With the above notation, let all components of the filtration \((M_r)_{r \in \mathbb{Z}}\) be finitely generated \( R \)-modules. Then the filtration \((M_r)_{r \in \mathbb{Z}}\) is excellent if and only if \( M' \) is a finitely generated \( D' \)-module.

**Lemma 3.1.7.** Let \( R \) be a Noetherian inversive difference ring with a basic set \( \sigma \). Then the ring of \( \sigma \)-operators \( D \) and the ring \( D' \) are left Noetherian.

**Theorem 3.1.8.** Let \( R \) be a Noetherian inversive difference ring with a basic set \( \sigma \). Let \( \rho : N \to M \) be an injective homomorphism of filtered \( \sigma\cdot R \)-modules and let the filtration of \( M \) be excellent. Then the filtration of \( N \) is also excellent.

Let \( R \) be a difference field with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \), \( D \) the ring of \( \sigma \)-operators over \( R \), and \( M \) a finitely generated \( \sigma\cdot R \)-module with generators \( x_1, \ldots, x_m \) (i.e., \( M = \sum_{i=1}^{m} D x_i \)). Then the vector \( R \)-spaces \( M_r = \sum_{i=1}^{m} D_r x_i \) \( (r \in \mathbb{Z}) \) form an excellent filtration of \( M \). It is easy to see that if \((M'_r)_{r \in \mathbb{Z}}\) is another excellent filtration of \( M \), then there exist \( k \in \mathbb{Z}, p \in \mathbb{N} \) such that \( M_r \subseteq M'_r + p \) and \( M'_r \subseteq M_r + p \) for all \( r \geq k \). Thus, if \( \psi(t) \) and \( \psi_1(t) \) are the characteristic polynomials of the \( \sigma\cdot R \)-module \( M \) associated with the excellent filtrations \((M_r)_{r \in \mathbb{Z}}\) and \((M'_r)_{r \in \mathbb{Z}}\), respectively, then \( \psi(r) \leq \psi_1(r + p) \) and \( \psi_1(r) \leq \psi(r + p) \) for all sufficiently large \( r \in \mathbb{Z} \). It follows that \( \deg \psi(t) = \deg \psi_1(t) \) and the leading coefficients of the polynomials \( \psi(t) \) and \( \psi_1(t) \) are equal. Since the degree of a characteristic polynomial of \( M \) does not exceed \( n \), \( \Delta^n \psi(t) = \Delta^n \psi_1(t) \in \mathbb{Z} \). (The \( n \)-th finite difference \( \Delta^n f(t) \) of a polynomial \( f(t) \) is defined as usual: \( \Delta f(t) = f(t + 1) - f(t), \Delta^k f(t) = \Delta(\Delta^{k-1} f(t)) \) for \( k = 1, 2, \ldots ) \) We arrive at the following result.

**Theorem 3.1.9.** Let \( R \) be a difference field with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \), \( M \) a finitely generated \( \sigma\cdot R \)-module, and \( \psi(t) \) the difference dimension polynomial associated with an excellent filtration of \( M \). Then the integers \( \Delta^n \psi(t), d = \deg \psi(t), \) and \( \Delta^d \psi(t) \) do not depend on the choice of the excellent filtration.
With the notation of the last theorem, the numbers $\Delta^n \psi(t), d = \deg \psi(t), and \Delta^d \psi(t)$ are called the difference (or $\sigma$-) dimension, difference (or $\sigma$-) type, and typical difference (or $\sigma$-) dimension of $M$, respectively. These characteristics of the $\sigma$-$R$-module $M$ are denoted by $\delta(M), t(M),$ and $\delta(M)$, respectively.

The following two theorems give some properties of the difference dimension (see [99] or [88, Section 6.2]).

**Theorem 3.1.10.** Let $R$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and let $0 \rightarrow N \rightarrow M \rightarrow P \rightarrow 0$ be an exact sequence of finitely generated $\sigma$-$R$-modules. Then $\delta(N) + \delta(P) = \delta(M)$.

**Theorem 3.1.11.** Let $R$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}, D$ the ring of $\sigma$-operators over $R$, and $M$ a finitely generated $\sigma$-$R$-module. Then $\delta(M)$ is equal to the maximal number of elements of $M$ linearly independent over $D$.

*Type and dimension of difference vector spaces*

Let $M$ be a module over a commutative ring $R$, $U$ a family of $R$-submodules of $M$, and $\mathcal{B}_U$ the set of all pairs $(N, N') \in U \times U$ such that $N' \subseteq N$. Furthermore, let $\overline{Z}$ denote the set $\mathbb{Z} \cup \{\infty\}$ considered as a linearly ordered set with the natural order $(a < \infty$ for any $a \in \mathbb{Z})$. As is shown in [80], there exists a unique map $\mu_U : \mathcal{B}_U \rightarrow \overline{Z}$ such that

(i) $\mu_U(N, N') \geq -1$ for every pair $(N, N') \in \mathcal{B}_U$;

(ii) for any $d \in \mathbb{N}$, the inequality $\mu_U(N, N') \geq d$ holds if and only if $N \not\subseteq N'$ and there exists an infinite chain $N = N_0 \supseteq N_1 \supseteq \cdots \supseteq N'$ such that $N_i \in U$ and $\mu_U(N_i, N_j) \geq d - 1$ for $i = 1, 2, \ldots$.

With the above notation, $\sup \{\mu_U(N, N') : (N, N') \in \mathcal{B}_U\}$ is called the *type of the $R$-module* $M$ over the family $U$; it is denoted by $\text{type}_U M$. The least upper bound of the lengths $p$ of chains $N_0 \supseteq N_1 \supseteq \cdots \supseteq N_p$ such that $N_i \in U$ $(0 \leq i \leq p)$ and $\mu_U(N_{i-1}, N_i) = \text{type}_U M$ for $i = 1, \ldots, p$ is called the *dimension of $M$ over $U$*; it is denoted by $\dim_U M$.

**Theorem 3.1.12** [102]. Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}, M$ a finitely generated $\sigma$-$K$-module, and $U$ the family of all $\sigma$-$K$-submodules of $M$.

(i) If $\delta(M) > 0$, then $\text{type}_U M = n$ and $\dim_U M = \delta(M)$.

(ii) If $\delta(M) = 0$, then $\text{type}_U M < n$.

*3.2. Inversive difference modules. $\sigma^*$-dimension polynomials and their invariants*

Let $R$ be an inversive difference ring with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and let $\Gamma$ denote the free commutative group generated by $\sigma$. As before, we set $\sigma^* = \{\alpha_1, \ldots, \alpha_n, \alpha_1^{-1}, \ldots, \alpha_n^{-1}\}$ and call $R$ a $\sigma^*$-ring. If $\gamma = \alpha_1^{k_1} \ldots \alpha_n^{k_n} \in \Gamma$, then the number $\sum_{i=1}^n |k_i|$ is called the *order* of the element $\gamma$; it is denoted by $\text{ord} \gamma$. For any $r \in \mathbb{N}$, the set $\{\gamma \in \Gamma \mid \text{ord} \gamma \leq r\}$ is denoted by $\Gamma(r)$. 
DEFINITION 3.2.1. An expression of the form $\sum_{\gamma \in \Gamma} a_{\gamma} \gamma$, where $a_{\gamma} \in R$ for any $\gamma \in \Gamma$ and only finitely many elements $a_{\gamma}$ are different from 0, is called an inversive difference (or $\sigma^*$-) operator over the difference ring $R$. Two $\sigma^*$-operators $\sum_{\gamma \in \Gamma} a_{\gamma} \gamma$ and $\sum_{\gamma \in \Gamma} b_{\gamma} \gamma$ are considered to be equal if and only if $a_{\gamma} = b_{\gamma}$ for any $\gamma \in \Gamma$.

The set of all $\sigma^*$-operators over $R$ can be naturally equipped with a ring structure if one sets $\sum_{\gamma \in \Gamma} a_{\gamma} \gamma + \sum_{\gamma \in \Gamma} b_{\gamma} \gamma = \sum_{\gamma \in \Gamma} (a_{\gamma} + b_{\gamma}) \gamma$, $a \sum_{\gamma \in \Gamma} a_{\gamma} \gamma = \sum_{\gamma \in \Gamma} (aa_{\gamma}) \gamma$, $(\sum_{\gamma \in \Gamma} a_{\gamma} \gamma) \gamma_1 = \sum_{\gamma \in \Gamma} a_{\gamma} (\gamma \gamma_1)$, $\gamma_1 a = \gamma_1 (a) \gamma_1$ for any $\sigma^*$-operators $\sum_{\gamma \in \Gamma} a_{\gamma} \gamma$, $\sum_{\gamma \in \Gamma} b_{\gamma} \gamma$ and for any $a \in R$, $\gamma_1 \in \Gamma$, and extends the multiplication by distributivity.

The ring obtained in this way is called the ring of inversive difference (or $\sigma^*$-) operators over $R$; it is denoted by $E$. Clearly, the ring of difference ($\sigma$-) operators $D$ introduced in the preceding section is a subring of $E$.

If $A = \sum_{\gamma \in \Gamma} a_{\gamma} \gamma \in E$, then the number $\text{ord} A = \max \{\text{ord} \gamma \mid a_{\gamma} \neq 0\}$ is called the order of the $\sigma^*$-operator $A$. Setting $E_r = \{A \in E \mid \text{ord} A \leq r\}$ for any $r \in \mathbb{N}$ and $E_0 = \{A \in E \mid \text{ord} A = 0\}$, we obtain an ascending filtration $(E_r)_{r \in \mathbb{Z}}$ of the ring $E$ called the standard filtration of this ring. In what follows, while considering $E$ as a filtered ring, we always mean this filtration.

THEOREM 3.2.2 [100]. If $R$ is a Noetherian inversive difference ring with a basic set $\sigma$, then the corresponding ring of $\sigma^*$-operators $E$ is left Noetherian. If $R$ is a $\sigma^*$-field, then $E$ is a left Ore ring.

DEFINITION 3.2.3. Let $R$ be an inversive difference ring with a basic set $\sigma$ and $E$ the ring of inversive difference operators over $R$. Then a left $E$-module is said to be an inversive difference $R$-module or a $\sigma^*$-$R$-module. In other words, an $R$-module $M$ is called a $\sigma^*$-$R$-module if elements of the set $\sigma^*$ act on $M$ in such a way that $\alpha(x + y) = \alpha x + \alpha y$, $\alpha(\beta x) = \beta(\alpha x)$, $\alpha(\alpha x) = \alpha (\alpha x)$, and $\alpha (\alpha^{-1} x) = x$ for any $\alpha, \beta \in \sigma^*$; $x, y \in M$; $a \in R$.

If $R$ is a $\sigma^*$-field, a $\sigma^*$-$R$-module $M$ is said to be a vector $\sigma^*$-$R$-space (or an inversive difference vector space over $R$).

It is clear that any $\sigma^*$-$R$-module can be naturally treated as a $\sigma$-$R$-module. Also, if $M$ and $N$ are two $\sigma^*$-$R$-modules, any difference ($\sigma$-) homomorphism $f : M \rightarrow N$ has the property that $f(\alpha x) = \alpha f(x)$ for any $x \in M$ and $\alpha \in \sigma^*$.

Let $R$ be an inversive difference ring with a basic set $\sigma$ and $E$ the ring of $\sigma^*$-operators over $R$. For any $\sigma^*$-$R$-module $M$, the set $C(M) = \{x \in M \mid \alpha x = x \text{ for all } \alpha \in \sigma\}$ is called the set of constants of $M$, its elements are called constants. Obviously, $C(M)$ is a subgroup of the additive group of $M$ and $C : M \mapsto C(M)$ is a functor from the category of $\sigma^*$-$R$-modules (i.e., the category of all left $E$-modules) to the category of Abelian groups.

If $M$ and $N$ are two $\sigma^*$-$R$-modules, then each of the $R$-modules $\text{Hom}_R(M, N)$ and $M \otimes_R N$ can be equipped with a structure of a $\sigma^*$-$R$-module if for any $f \in \text{Hom}_R(M, N)$, $\sum_{i=1}^k x_i \otimes y_i \in M \otimes_R N$ ($x_1, \ldots, x_k \in M; y_1, \ldots, y_k \in N$), and $\alpha \in \sigma^*$, one defines $\alpha f$ by $(\alpha f) x = \alpha f(\alpha^{-1} x)$ and sets $\alpha f(\sum_{i=1}^k x_i \otimes y_i) = \sum_{i=1}^k \alpha x_i \otimes \alpha y_i$. It is easy to check that $\alpha f \in \text{Hom}_R(M, N)$ and the actions of elements of $\sigma^*$ on $\text{Hom}_R(M, N)$ and $M \otimes_R N$ satisfy the conditions of Definition 3.2.3. In what follows, while considering $\text{Hom}_R(M, N)$
and $M \otimes_R N$ as $\sigma^*\cdot R$-modules (for some $\sigma^*\cdot R$-modules $M$ and $N$), we always mean these invasive difference structures of these modules.

**Proposition 3.2.4** [88, Section 3.4]. Let $R$ be an inversive difference ring with a basic set $\sigma$, and let $M$, $N$, and $P$ be $\sigma^*\cdot R$-modules.

(i) The natural isomorphism of $R$-modules

$$ \eta: \text{Hom}_R(P \otimes_R M, N) \to \text{Hom}_R(P, \text{Hom}_R(M, N)) $$

(defined by $[(\eta f)x](y) = f(x \otimes y)$ for any $f \in \text{Hom}_R(P \otimes_R M, N)$, $x \in P$, $y \in M$) is a $\sigma^*\cdot$-isomorphism.

(ii) $C(\text{Hom}_R(M, N)) = \text{Hom}_\mathbb{C}(M, N)$.

(iii) The functors $C$ and $\text{Hom}_\mathbb{C}(\cdot \otimes_R M, N)$ are naturally isomorphic.

(iv) The functor $C$ is left exact and for any positive integer $p$, its $p$-th right derived functor $R^pC$ is naturally isomorphic to the functor $\text{Ext}_\mathbb{C}^p(R, \cdot)$.

(v) The functors $\text{Hom}_\mathbb{C}(\cdot \otimes_R M, N)$ and $\text{Hom}_\mathbb{C}(\cdot, \text{Hom}_R(M, N))$ are naturally isomorphic and the same is true for the functors $\text{Hom}_\mathbb{C}(M \otimes_R \cdot, N)$ and $\text{Hom}_\mathbb{C}(M, \text{Hom}_R(\cdot, N))$.

(vi) For any positive integers $p$ and $q$, there exists a spectral sequence converging to $\text{Ext}_\mathbb{C}^{p+q}(M, N)$ whose second term is equal to $E^{p,q}_2 = (R^pC)(\text{Ext}_R^q(M, N))$.

**Definition 3.2.5.** Let $R$ be an inversive difference ring with a basic set $\sigma$, $\mathcal{E}$ the ring of $\sigma^*$-operators over $R$ (considered as a filtered ring with the standard filtration $(\mathcal{E}_r)_{r \in \mathbb{Z}}$), and $M$ a $\sigma^*\cdot R$-module. An ascending chain $(M_r)_{r \in \mathbb{Z}}$ of $R$-submodules of $M$ is called a filtration of $M$ if $\mathcal{E}_r M_s \subseteq M_{r+s}$ for all $r, s \in \mathbb{Z}$, $M_r = 0$ for all sufficiently small $r \in \mathbb{Z}$, and $\bigcup_{r \in \mathbb{Z}} M_r = M$. A filtration $(M_r)_{r \in \mathbb{Z}}$ of the $\sigma^*\cdot R$-module $M$ is called excellent if all $R$-modules $M_r$ ($r \in \mathbb{Z}$) are finitely generated and there exists $r_0 \in \mathbb{Z}$ such that $M_r = \mathcal{E}_{r-r_0} M_{r_0}$ for any $r \in \mathbb{Z}$, $r \geq r_0$.

**Theorem 3.2.6.** Let $R$ be an Artinian $\sigma^*$-ring with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and let $(M_r)_{r \in \mathbb{Z}}$ be an excellent filtration of a $\sigma^*\cdot R$-module $M$. Then there exists a numerical polynomial $\chi(t) \in \mathbb{Q}[t]$ such that

(i) $\chi(r) = l_R(M_r)$ for all sufficiently large $r \in \mathbb{Z}$;

(ii) $\deg \chi(t) \leq n$ and the polynomial $\chi(t)$ can be represented in the form $\chi(t) = \sum_{i=0}^n 2^i a_i \binom{t+i}{i}$ where $a_0, \ldots, a_n \in \mathbb{Z}$.

The polynomial $\chi(t)$ whose existence is established by Theorem 3.2.6 is called the $\sigma^*$-dimension polynomial or the characteristic polynomial of the module $M$ associated with the excellent filtration $(M_r)_{r \in \mathbb{Z}}$.

**Example 3.2.7.** Let $R$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $\mathcal{E}$ the ring of $\sigma^*$-operators over $R$, and $\chi \mathcal{E}(t)$ the difference dimension polynomial associated with the standard filtration $(\mathcal{E}_r)_{r \in \mathbb{Z}}$. Then $\chi \mathcal{E}(r) = l_R(\mathcal{E}_r) = \dim_R(\mathcal{E}_r) = \text{Card}(\gamma = \alpha_1^k \ldots \alpha_n^k \in \Gamma \mid \text{ord } \gamma = \sum_{i=1}^n |k_i| \leq r)$ for all sufficiently large $r \in \mathbb{Z}$.
Proposition 2.1.9] we obtain three expressions for the last number that lead to the following three forms of the polynomial $\chi_\mathcal{E}(t)$:

$$\chi_\mathcal{E}(t) = \sum_{i=0}^{n} 2^i \binom{n}{i} \frac{(t+i)}{n} = \sum_{i=0}^{n} \binom{n}{i} \left( t + i \right) = \sum_{i=0}^{n} (-1)^n 2^i \binom{n}{i} \left( t + i \right).$$

Let $F_m$ be a free left $\mathcal{E}$-module of rank $m$ ($m \geq 1$) with free generators $f_1, \ldots, f_m$. Then for every $l \in \mathbb{Z}$, one can consider the excellent filtration $((F_m^l)_r)_{r \in \mathbb{Z}}$ of the module $F_m$ such that $(F_m^l)_{r} = \sum_{i=l}^{m} \mathcal{E}_{r-i} f_i$ for every $r \in \mathbb{Z}$. We obtain a filtered $\sigma^*-R$-module that will be denoted by $F_m^l$. A finite direct sum of such filtered $\sigma^*-R$-modules is called a free filtered $\sigma^*-R$-module. The representations of $\chi_\mathcal{E}(i)$ imply the following expressions for the $\sigma^*$-dimension polynomial $\chi(t)$ of $F_m^l$:

$$\chi(t) = m \chi_\mathcal{E}(t-l) = m \sum_{i=0}^{n} 2^i \binom{n}{i} \left( t - l \right) = m \sum_{i=0}^{n} \binom{n}{i} \left( t + i - l \right) = m \sum_{i=0}^{n} (-1)^n 2^i \binom{n}{i} \left( t + i - l \right).$$

Let $R$ be an inversive difference ring with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $\mathcal{E}$ the ring of $\sigma^*$-operators over $R$, $M$ a filtered $\sigma^*-R$-module with a filtration $(M_r)_{r \in \mathbb{Z}}$, and $R[x]$ the ring of polynomials in one indeterminate $x$ over $R$. Let $\mathcal{E}'$ denote the subring $\sum_{r \in \mathbb{N}} \mathcal{E}_r \otimes_R R x^r$ of the ring $\mathcal{E} \otimes_R R[x]$ and $M'$ denote the left $\mathcal{E}'$-module $\sum_{r \in \mathbb{N}} M_r \otimes_R R x^r$. The following three results are similar to the corresponding statements for difference modules.

**Lemma 3.2.8.** With the above notation, let all components of the filtration $(M_r)_{r \in \mathbb{Z}}$ be finitely generated $R$-modules. Then the filtration $(M_r)_{r \in \mathbb{Z}}$ is excellent if and only if $M'$ is a finitely generated $\mathcal{E}'$-module.

**Lemma 3.2.9.** If $R$ is a Noetherian inversive difference ring, then the ring $\mathcal{E}'$ considered above is left Noetherian.

Let $R$ be an inversive difference ring with a basic set $\sigma$ and let $M$ and $N$ be filtered $\sigma^*-R$-modules with filtrations $(M_r)_{r \in \mathbb{Z}}$ and $(N_r)_{r \in \mathbb{Z}}$, respectively. Then a $\sigma$-homomorphism $f : M \rightarrow N$ is said to be a $\sigma$-homomorphism of filtered $\sigma^*-R$-modules if $f(M_r) \subseteq N_r$ for any $r \in \mathbb{Z}$.

**Theorem 3.2.10.** Let $R$ be a Noetherian inversive difference ring with a basic set $\sigma$ and let $\rho : N \rightarrow M$ be an injective homomorphism of filtered $\sigma^*-R$-modules. Furthermore, suppose that the filtration of the module $M$ is excellent. Then the filtration of $N$ is also excellent.

As in the case of difference modules, for any two excellent filtrations $(M_r)_{r \in \mathbb{Z}}$ and $(M'_r)_{r \in \mathbb{Z}}$ of a finitely generated $\sigma^*$-module $M$ over an inversive difference $(\sigma^*)$ field $K$, there exist $p \in \mathbb{N}$ such that $M_r \subseteq M'_{r+p}$ and $M'_r \subseteq M_{r+p}$ for all $r \in \mathbb{Z}$. This observation
leads to the following statement that gives some invariants of the $\sigma^*$-dimension polynomial of $M$.

**Theorem 3.2.11.** Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $M$ a finitely generated $\sigma^*$-$K$-module, and $\chi(t)$ the characteristic polynomial associated with an excellent filtration of $M$. Then the integers $\frac{\Delta^n \chi(t)}{\gcd^n}$, $d = \deg \chi(t)$, and $\frac{\Delta^d \chi(t)}{\gcd^d}$ do not depend on the choice of the excellent filtration of $M$.

With the notation of the last theorem, the numbers $\frac{\Delta^n \chi(t)}{\gcd^n}$, $d = \deg \chi(t)$, and $\frac{\Delta^d \chi(t)}{\gcd^d}$ are called the inversive difference (or $\sigma^*$-) dimension, inversive difference (or $\sigma^*$-) type, and typical inversive difference (or typical $\sigma^*$-) dimension of the module $M$, respectively. These characteristics of the $\sigma^*$-$K$-module $M$ are denoted by $i\delta(M)$, $it(M)$, and $ti\delta(M)$, respectively. (If we want to indicate the basic set with respect to which $K$ is considered as a difference field, we will use the notation $i\delta_n(M)$, $it_n(M)$, and $ti\delta_n(M)$, respectively.)

Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $\Gamma$ the free commutative group generated by $\sigma$, and $M$ a $\sigma^*$-$K$-module. Elements $z_1, \ldots, z_m \in M$ are said to be $\sigma^*$-linearly independent over $K$ if the set $\{yz_i \mid 1 \leq i \leq m, y \in \Gamma\}$ is linearly independent over the $\sigma^*$-field $K$. Otherwise, $z_1, \ldots, z_m$ are said to be $\sigma^*$-linearly dependent over $K$.

**Theorem 3.2.12 [100].** Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and let $0 \rightarrow N \xrightarrow{i} M \xrightarrow{j} P \rightarrow 0$ be an exact sequence of finitely generated $\sigma^*$-$K$-modules. Then $i\delta(N) + i\delta(P) = i\delta(M)$.

**Theorem 3.2.13 [100].** Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $E$ the ring of $\sigma^*$-operators over $K$, and $M$ a finitely generated $\sigma^*$-$K$-module. Then $i\delta(M)$ is equal to the maximal number of elements of $M$ that are $\sigma^*$-linearly independent over $K$.

**Theorem 3.2.14 [102].** Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $M$ a finitely generated $\sigma^*$-$K$-module, and $U$ the family of all $\sigma^*$-$K$-submodules of $M$. Then

(i) If $i\delta(M) > 0$, then $\text{type}_U M = n$ and $\dim_U M = i\delta(M)$.

(ii) If $i\delta(M) = 0$, then $\text{type}_U M < n$.

Let $R$ be an inversive difference ring with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and $\Gamma$ the free commutative group generated by $\sigma$. It is easy to see that if $\sigma_1 = \{\tau_1, \ldots, \tau_n\}$ is another system of free generators of $\Gamma$, then there exists a matrix $K = (k_{ij})_{1 \leq i, j \leq n} \in GL(n, \mathbb{Z})$ such that $\alpha_i = \tau_1^{k_{i1}} \cdots \tau_n^{k_{in}}$ ($1 \leq i \leq n$). The $\sigma^*$-ring $R$ can be also treated as a $\sigma_1^*$-ring and the corresponding ring of $\sigma_1^*$-operators coincides with the ring of $\sigma^*$-operators $\mathcal{E}$ over $R$. We say that two finite sets of automorphisms $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and $\sigma_1 = \{\tau_1, \ldots, \tau_n\}$ of the same ring $R$ are equivalent if there exists a matrix $(k_{ij})_{1 \leq i, j \leq n} \in GL(n, \mathbb{Z})$ such that $\alpha_i = \tau_1^{k_{i1}} \cdots \tau_n^{k_{in}}$ ($1 \leq i \leq n$).
Theorem 3.2.15 [88, Theorem 6.3.19]. Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $M$ a finitely generated $\sigma^+K$-module, and $d = i\sigma(M)$. Then there exists a set $\sigma' = \{\beta_1, \ldots, \beta_n\}$ of pairwise commuting automorphisms of $K$ such that

(i) The sets $\sigma$ and $\sigma'$ are equivalent.

(ii) Let $\sigma'' = \{\beta_1, \ldots, \beta_d\}$. Then $M$ is a finitely generated $\sigma''^+K$-module and $i\sigma''(M) > 0$.

There are several publications devoted to methods and algorithms of computation of characteristic polynomials of difference and inversive difference modules (see [87,89–91], [88, Chapters 6, 9], [113,119], and [127]). The corresponding techniques are based either on constructing resolutions of free filtered difference modules (whose characteristic polynomials are given in Example 3.2.7) or by applying the Gröbner basis method to modules over rings of difference and inversive difference operators.

3.3. Reduction in a free difference vector space. Characteristic sets and multivariable dimension polynomials

In this section we apply the method of characteristic sets to difference and inversive difference modules whose basic sets of translations are represented as unions of their disjoint subsets. In the case of difference and inversive difference vector spaces, we generalize the results of the two preceding sections and show the existence of characteristic polynomials in several variables associated with partitions of the basic set. We also present invariants of such polynomials.

Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and let a partition of $\sigma$ into a union of $p$ its proper disjoint subsets be fixed: $\sigma = \bigcup_{j=1}^p \sigma_j$ where $\sigma_1 = \{\alpha_1, \ldots, \alpha_{n_1}\}$, $\sigma_2 = \{\alpha_{n_1+1}, \ldots, \alpha_{n_1+n_2}\}$, $\ldots$, $\sigma_p = \{\alpha_{n_1+\ldots+n_{p-1}+1}, \ldots, \alpha_n\}$ ($p \geq 1, n_1, \ldots, n_p \in \mathbb{N}$). As before, $T$ and $D$ denote the free commutative semigroup generated by $\sigma$ and the ring of $\sigma$-operators over $K$, respectively. If $t = \alpha_1^{k_1} \ldots \alpha_n^{k_n} \in T$, then the order of the element $t$ with respect to $\sigma_i$ ($1 \leq i \leq p$) is defined as $\text{ord}_i t = \sum_{\nu = n_1 + \ldots + n_{i-1} + 1}^{n_1 + \ldots + n_i} k_\nu$ (if $i = 1$, then the lower index in the last sum is 1). The order of $i$ is still defined as $\text{ord}_i \theta = k_1 + \ldots + k_n$.

Our partition of the set $\sigma$ induces $p$ orderings $<_1, \ldots, <_p$ of the semigroup $T$ defined as follows: $t = \alpha_1^{l_1} \ldots \alpha_n^{l_n} <_i t' = \alpha_1^{l_1'} \ldots \alpha_n^{l_n'}$ if and only if the $(n+p+1)$-tuple $(\text{ord}_1 t, \text{ord}_1 t', \ldots, \text{ord}_{i-1} t, \text{ord}_{i-1} t', \ldots, \text{ord}_p t, \text{ord}_p t', k_{n_1+\ldots+n_{i-1}+1}, \ldots, k_{n_1+\ldots+n_i})$ is less than the $(n+p+1)$-tuple $(\text{ord}_1 t', \text{ord}_1 t', \ldots, \text{ord}_{i-1} t', \text{ord}_{i-1} t', \ldots, \text{ord}_p t', l_{n_1+\ldots+n_{i-1}+1}, \ldots, l_{n_1+\ldots+n_i}, l_1, \ldots, l_{n_1+\ldots+n_{i-1}})$ with respect to the lexicographic order on $\mathbb{N}^{n+p+1}$.

For any $r_1, \ldots, r_p \in \mathbb{N}$, $T(r_1, \ldots, r_p)$ will denote the set $\{t \in T | \text{ord}_1 t \leq r_1, \ldots, \text{ord}_p t \leq r_p\}$; the vector $K$-subspace of $D$ generated by this set will be denoted by $D_{r_1\ldots r_p}$. Setting $D_{r_1\ldots r_p} = 0$ for $(r_1, \ldots, r_p) \in \mathbb{Z}^p \setminus \mathbb{N}^p$, we obtain a family $\{D_{r_1\ldots r_p} | (r_1, \ldots, r_p) \in \mathbb{Z}^p\}$ the standard $p$-dimensional filtration of the ring $D$.

Definition 3.3.1. A family $\{M_{r_1\ldots r_p} | (r_1, \ldots, r_p) \in \mathbb{Z}^p\}$ of vector $K$-subspaces of a $\sigma$-$K$-module $M$ is called a $p$-dimensional filtration of $M$ if
(i) for any fixed integers $r_1, \ldots, r_{i-1}, r_i+1, \ldots, r_p$ $(1 \leq i \leq p)$, $M_{r_1 \ldots r_{i-1} r_i+1 \ldots r_p} \subseteq M_{r_1 \ldots r_{i-1} r_i+1 \ldots r_p}$ and $M_{r_1 \ldots r_p} = 0$ for all sufficiently small $r_i \in \mathbb{Z}$;
(ii) $\bigcup \{M_{r_1 \ldots r_p} \mid (r_1, \ldots, r_p) \in \mathbb{Z}^p\} = M$;
(iii) $D_{r_1 \ldots r_p} k_{1} s_p \subseteq M_{r_1 + s_1, \ldots, r_p + s_p}$ for any $(r_1, \ldots, r_p) \in \mathbb{N}^p, (s_1, \ldots, s_p) \in \mathbb{Z}^p$.

If every vector $K$-space $M_{r_1 \ldots r_p}$ is finitely generated and there exists an element $(h_1, \ldots, h_p) \in \mathbb{Z}^p$ such that $R_{r_1 \ldots r_p} M_{h_1 \ldots h_p} = M_{r_1 + h_1 \ldots r_p + h_p}$ for any $(r_1, \ldots, r_p) \in \mathbb{N}^p$, then the $p$-dimensional filtration is called excellent.

It is easy to see that if $u_1, \ldots, u_n$ is a finite system of generators of a left $D$-module $M$, then the filtration $\{\sum_{i=1}^n R_{r_1 \ldots r_p} u_i \mid (r_1, \ldots, r_p) \in \mathbb{Z}^p\}$ is excellent.

Let $F$ be a finitely generated free $\sigma$-$K$-module (that is, a free left $D$-module) and let $f_1, \ldots, f_q$ be a fixed basis of $F$ over $D$. Then elements $tf_k$ $(t \in T, 1 \leq k \leq q)$ are called terms; the set of all terms is denoted by $T_f$. The order of a term $tf_k$ and the order of this term with respect to $\sigma_i$ $(1 \leq i \leq p)$ are defined as the order of the element $t \in T$ and the order of $t$ relative to $\sigma_i$, respectively. A term $tf_i$ is said to be a multiple of a term $t'f_j$ if $i = j$ and $t'$ divides $t$ in the semigroup $T$ (that is, $t = t''t'$ for some $t'' \in T$). In this case we also say that the term $t'f_j$ divides $tf_i$ and write $t'f_j \mid tf_i$.

Below we consider $p$ orderings of the set $T_f$ that correspond to the orderings of the set $T$. These ordering are denoted by the same symbols $<_1, \ldots, <_p$ and defined as follows: if $tf_k, tf_j \in T_f$, then $tf_k <_1 tf_j$ if and only if $t <_1 t'$ in $T$ or $t = t'$ and $k < l$.

Since the set $T_f$ is a basis of the vector $K$-space $F$, every element $f \in F$ has a unique representation in the form

$$f = a_1 t_1 f_{i_1} + \cdots + a_m t_m f_{i_m},$$

(\*)

where $t_i \in T$, $a_i \in K$, $a_i \neq 0$ $(1 \leq i \leq m)$, $1 \leq i_1, \ldots, i_m \leq q$ and all terms $t_v f_{i_v}$ $(1 \leq v \leq m)$ are distinct. For any $j = 1, \ldots, p$, the greatest with respect to $<_j$ order of the term of the set $\{t_v f_{i_v} \mid 1 \leq v \leq m\}$ is called the $j$-leader of the element $f$; it is denoted by $u_f^{(j)}$. (Of course, it is possible that $u_f^{(j)} = u_f^{(l)}$ for some distinct numbers $j$ and $l$.)

In what follows, we say that an element $f \in F$ contains a term $tf_j$ if the term appears in the representation (\*) with a non-zero coefficient. The coefficient of the $j$-leader of an element $f \in F$ will be denoted by $lc_j(f)$ $(1 \leq j \leq p)$.

**Definition 3.3.2.** Let $f$ and $g$ be two elements of the free $\sigma$-$K$-module $F$ considered above. The element $f$ is said to be reduced with respect to $g$ if $f$ does not contain any multiple $t u_g^{(j)}$ $(t \in T)$ of the 1-leader $u_g^{(j)}$ such that $\text{ord}_j(t u_g^{(j)}) \leq \text{ord}_j u_g^{(j)}$ for $j = 2, \ldots, p$. An element $h \in F$ is said to be reduced with respect to a set $A \subseteq F$, if $h$ is reduced with respect to every element of $A$. A set $\Sigma \subseteq F$ is called autoreduced if every element of $\Sigma$ is reduced with respect to any other element of this set. An autoreduced set $\Sigma$ is called normal if $lc_1(g) = 1$ for every element $g \in \Sigma$.

**Theorem 3.3.3.** Let $\Sigma$ be an autoreduced set in the free $\sigma$-$K$-module $F$. Then:

(i) The set $\Sigma$ is finite, $\Sigma = \{g_1, \ldots, g_r\}$.
(ii) For any $f \in F$, there exists an element $g \in F$ such that $f - g = \sum_{i=1}^r \lambda_i g_i$ for some $\lambda_1, \ldots, \lambda_r \in D$ and $g$ is reduced with respect to $\Sigma$. 

Let $f$ and $g$ be two elements of the free $\sigma$-$K$-module $F$. We say that $f$ has lower rank than $g$ and write $rk(f) < rk(g)$ if either $u_f^{(i)} < u_g^{(i)}$ or there exists some $k$, $2 \leq k \leq p$, such that $u_f^{(v)} = u_g^{(v)}$ for $v = 1, \ldots, k - 1$ and $u_f^{(k)} < k u_g^{(k)}$. If $u_f^{(i)} = u_g^{(i)}$ for $i = 1, \ldots, p$, we say that $f$ and $g$ have the same rank and write $rk(f) = rk(g)$. In what follows, while considering autoreduced sets in $F$, we always assume that their elements are arranged in order of increasing rank.

**Definition 3.3.4.** Let $\Sigma = \{h_1, \ldots, h_r\}$ and $\Sigma' = \{h'_1, \ldots, h'_s\}$ be two autoreduced subsets of the free $\sigma$-$K$-module $F$. An autoreduced set $\Sigma$ is said to have lower rank than $\Sigma'$ if one of the following two cases holds:

1. There exists $k \in N$ such that $k \leq \min\{r, s\}$, $rk(h_i) = rk(h'_i)$ for $i = 1, \ldots, k - 1$ and $rk(h_k) < rk(h'_k)$.

2. $r > s$ and $rk(h_i) = rk(h'_i)$ for $i = 1, \ldots, s$.

If $r = s$ and $rk(h_i) = rk(h'_i)$ for $i = 1, \ldots, r$, then $\Sigma$ is said to have the same rank as $\Sigma'$.

As in the case of autoreduced sets of difference polynomials, one can show that in every non-empty set of autoreduced subsets of the free $\sigma$-$K$-module $F$ there exists an autoreduced subset of lowest rank. If $N$ is a $D$-submodule of $F$, then an autoreduced subset of $N$ of lowest rank is called a characteristic set of the module $N$.

**Theorem 3.3.5.** Let $N$ be a $D$-submodule of the free $\sigma$-$K$-module $F$ and let $\Sigma = \{g_1, \ldots, g_r\}$ be a characteristic set of $N$.

(i) An element $f \in N$ is reduced with respect to $\Sigma$ if and only if $f = 0$.

(ii) If $N$ is a cyclic $D$-submodule of $F$ generated by an element $g$, then $\{g\}$ is a characteristic set of $N$.

(iii) The set $\Sigma$ generates $N$ as a left $D$-module.

(iv) Let the characteristic set $\Sigma$ be normal and let $\Sigma_1 = \{h_1, \ldots, h_s\}$ be another normal characteristic sets of $N$. Then $r = s$ and $g_i = h_i$ for $i = 1, \ldots, r$.

**Theorem 3.3.6.** Let $K$ be a difference field with a basic set $\sigma$, $D$ the ring of $\sigma$-operators over $K$, and $M$ a finitely generated $\sigma$-$K$-module with a system of generators $\{e_1, \ldots, e_q\}$. Furthermore, let $F$ be a free $\sigma$-$K$-module with a basis $f_1, \ldots, f_q$, $\pi : F \rightarrow M$ the natural $D$-epimorphism ($\pi(f_i) = e_i$ for $i = 1, \ldots, q$), and $\Sigma = \{g_1, \ldots, g_d\}$ a characteristic set of the $D$-module $N = \ker \pi$. Finally, for any $r_1, \ldots, r_p \in N$, let $M_{r_1 \ldots r_p} = \sum_{i=1}^{p} D_{r_1 \ldots r_p} e_i$ and $U_{r_1 \ldots r_p} = \{w \in T f \mid \text{ord}_j w \leq r_j \text{ for } j = 1, \ldots, p \text{ and either } w \text{ is not a multiple of any } u_g^{(i_j)} \ (1 \leq i_j \leq d) \text{ or for any } t \in T, g_i \in \Sigma \text{ such that } w = tu_g^{(i_j)} \text{ we have } \text{ord}_j(tu_g^{(i_j)}) > r_j \text{ for some } j, 2 \leq j \leq p\}$. Then $\pi(U_{r_1 \ldots r_p})$ is a basis of the vector $K$-space $M_{r_1 \ldots r_p}$.

**Theorem 3.3.7.** Let $K$ be a difference field with a basic set $\sigma$ and let $\{M_{r_1 \ldots r_p} \mid (r_1, \ldots, r_p) \in Z^p\}$ be an excellent $p$-dimensional filtration of a $\sigma$-$K$-module $M$. Then there exists a polynomial $\phi(t_1, \ldots, t_p) \in Q[t_1, \ldots, t_p]$ such that

(i) $\phi(r_1, \ldots, r_p) = \dim_K M_{r_1 \ldots r_p}$ for all sufficiently large $(r_1, \ldots, r_p) \in Z^p$ (i.e., there exists $(r_1^{(0)}, \ldots, r_p^{(0)}) \in Z^p$ such that the equality holds for all $(r_1, \ldots, r_p)$ that exceed $(r_1^{(0)}, \ldots, r_p^{(0)}) \in Z^p$ with respect to the product order on $Z^p$);
(ii) \( \deg_{i_1} \phi \leq n_i \) for \( i = 1, \ldots, p \) (in particular, the total degree of \( \phi \) does not exceed \( n \)) and the polynomial \( \phi(t_1, \ldots, t_p) \) can be represented as

\[
\phi = \sum_{i_1=0}^{n_1} \cdots \sum_{i_p=0}^{n_p} a_{i_1 \ldots i_p} \binom{t_1 + i_1}{i_1} \cdots \binom{t_p + i_p}{i_p},
\]

where \( a_{i_1 \ldots i_p} \in \mathbb{Z} \) for all \( i_1, \ldots, i_p \).

The numerical polynomial \( \phi(t_1, \ldots, t_p) \), whose existence is established by Theorem 3.3.7, is called the difference (or \( \sigma \)-) dimension polynomial of the module \( M \) associated with the \( p \)-dimensional filtration \( \{ M_{r_1 \ldots r_p} \mid (r_1, \ldots, r_p) \in \mathbb{Z}^p \} \).

Any permutation \( (j_1, \ldots, j_p) \) of the set \( \{1, \ldots, p\} \), defines a lexicographic order \( \preceq_{j_1, \ldots, j_p} \) on \( \mathbb{N}^p \) such that \( (r_1, \ldots, r_p) \preceq_{j_1, \ldots, j_p} (s_1, \ldots, s_p) \) if and only if either \( r_{j_1} < s_{j_1} \) or there exists \( k \in \mathbb{N}, 1 \leq k \leq p - 1 \), such that \( r_{j_v} = s_{j_v} \) for \( v = 1, \ldots, k \) and \( r_{j_{k+1}} < s_{j_{k+1}} \). In what follows, we use these orders to associate with every set \( \Sigma \subseteq \mathbb{N}^p \) the set \( \Sigma' = \{ e \in \Sigma \mid e \) is a maximal element of \( \Sigma \) with respect to one of the \( p! \) lexicographic orders \( \preceq_{j_1, \ldots, j_p} \} \).

For example, if \( \Sigma = \{(3, 0, 2), (2, 1, 1), (0, 1, 4), (1, 0, 3), (1, 1, 6), (3, 1, 0), (1, 2, 0)\} \subseteq \mathbb{N}^3 \), then \( \Sigma' = \{(3, 0, 2), (3, 1, 0), (1, 1, 6), (1, 2, 0)\} \).

**Theorem 3.3.8.** Let \( K \) be a difference field with a basic set \( \sigma \), \( M \) be a finitely generated \( \sigma \)-\( K \)-module, \( \{ M_{r_1 \ldots r_p} \mid (r_1, \ldots, r_p) \in \mathbb{Z}^p \} \) an excellent \( p \)-dimensional filtration of \( M \), and

\[
\phi(t_1, \ldots, t_p) = \sum_{i_1=0}^{n_1} \cdots \sum_{i_p=0}^{n_p} a_{i_1 \ldots i_p} \binom{t_1 + i_1}{i_1} \cdots \binom{t_p + i_p}{i_p}
\]

the \( \sigma \)-dimension polynomial associated with this filtration. Let \( E = \{(i_1, \ldots, i_p) \in \mathbb{N}^p \mid 0 \leq i_k \leq n_k \ (k = 1, \ldots, p) \) and \( a_{i_1 \ldots i_p} \neq 0 \).

Then the total degree \( d \) of the polynomial \( \phi \), \( a_{n_1 \ldots n_p} \), \( p \)-tuples \( (j_1, \ldots, j_p) \in E' \), the corresponding coefficients \( a_{j_1 \ldots j_p} \), and the coefficients of the terms of total degree \( d \) do not depend on the choice of the excellent filtration.

Methods of computation of multivariable characteristic polynomials of difference and inversive difference vector spaces are similar to those for multivariable Hilbert and differential dimension polynomials, see [106,107], and [109]. In particular, these papers contain examples showing that multivariable characteristic polynomials can carry essentially more invariants than classical dimension polynomials in one variable.

4. Difference field extensions

4.1. Transformal dependence. Difference transcendental bases and difference transcendental degree

Let \( K \) be a difference field with a basic set \( \sigma \), \( T \) the free commutative semigroup generated by \( \sigma \), and \( L \) a \( \sigma \)-overfield of \( K \). We say that an element \( v \in L \) is transformally dependent or
σ-algebraically dependent on a set $A \subseteq L$ over $K$ if $v$ is σ-algebraic over the field $K(A)$. Obviously, an element $v \in L$ is σ-algebraically dependent on a set $A \subseteq L$ if and only if there exists a finite family $\{\eta_1, \ldots, \eta_s\} \subseteq A$ such that $v$ is σ-algebraic over $K(\eta_1, \ldots, \eta_s)$.

Let $K$ be an inversive σ-field, $L$ a σ*-overfield of $K$ and $A \subseteq L$. It is easy to see that an element $v \in L$ is σ*-algebraically dependent on $A$ over $K$ (that is, $v$ is σ-algebraic over the σ*-field $K(A^*)$) if and only if $v$ is σ-algebraically dependent on $A$ over $K$.

All the statements in the rest of this section can be proved in the same way as their ordinary versions (see [32, Chapter 5]).

**Proposition 4.1.1.** Let $K$ be a difference field with a basic set $\sigma$, $L$ a σ*-overfield of $K$ and $A \subseteq L$.

(i) The set $A$ is σ-algebraically dependent over $K$ if and only if there exists $v \in A$ such that $v$ is σ-algebraically dependent on $A \setminus \{v\}$ over $K$.

(ii) The set $A$ contains a maximal subset σ-algebraically independent over $K$. In other words, there exists a set $B \subseteq A$ such that $B$ is σ-algebraically independent over $K$ and any subset of $A$ properly containing $B$ is σ-algebraically dependent over $K$.

A set $B$, whose existence is established by Proposition 4.1.1(ii), is called a basis for transformal transcendence or a difference (or σ-) transcendence basis of $A$ over $K$. If $A = L$, the set $B$ is called a basis for transformal transcendence or a difference (or σ-) transcendence basis of $L$ over $K$. (We use this terminology when $L$ is a difference overfield of a σ-field $K$ or an inversive difference field extension of a σ*-field $K$.)

**Proposition 4.1.2.** Let $K$ be a difference (inversive difference) field with a basic set $\sigma$ and $L$ a σ*-overfield of $K$. Furthermore, let $B$ and $B'$ be two subsets of $L$ and $v, u_1, \ldots, u_m \in L$.

(i) If $v$ is σ-algebraically dependent on $B$ over $K$ and every element of $B$ is σ-algebraically dependent on $B'$ over $K$, then $v$ is σ-algebraically dependent on $B'$ over $K$.

(ii) If $v$ is σ-algebraically dependent on $\{u_1, \ldots, u_m\}$, but not on $\{u_1, \ldots, u_{m-1}\}$ over $K$, then $u_m$ is σ-algebraically dependent on the set $\{u_1, \ldots, u_{m-1}, v\}$ over $K$.

(iii) Suppose that $B' \subseteq B, u_1, \ldots, u_m \in B$ are σ-algebraically independent over $K$, and each $u_i$ ($1 \leq i \leq m$) is σ-algebraically dependent on $B'$ over $K$. Then there exist elements $v_1, \ldots, v_m \in B'$ such that each $v_i$ is σ-algebraically dependent over $K$ on the set $B''$ obtained from $B'$ by replacing $v_j$ by $u_j$ ($j = 1, \ldots, m$).

**Proposition 4.1.3.** Let $K$ be a difference field with a basic set $\sigma$, $L$ a σ*-overfield of $K$ and $A \subseteq L$.

(i) Suppose that $B$ is a subset of $A$ which is σ-algebraically independent over $K$. Then $B$ is a σ-transcendence basis of $A$ over $K$ if and only if every element of $A$ is σ-algebraically dependent on $B$ over $K$.

(ii) All σ-transcendence bases of $A$ over $K$ either contain the same finite number of elements or are infinite.
DEFINITION 4.1.4. Let $K$ be a difference (in particular, inversive difference) field with a basic set $\sigma$, $L$ a $\sigma$-overfield of $K$ and $A \subseteq L$. Then the $\sigma$-transcendence degree of $A$ over $K$ is the number of elements of any $\sigma$-transcendence basis of $A$ over $K$, if this number is finite, or infinity in the contrary case.

The $\sigma$-transcendence degree of $A$ over $K$ is denoted by $\sigma$-trdeg$_K A$. In particular, if $A = L$, then $\sigma$-trdeg$_K L$ denotes the $\sigma$-transcendence degree of the $\sigma$-(or $\sigma^*$-) field extension $L/K$.

PROPOSITION 4.1.5. Let $K$ be a difference field with a basic set $\sigma$ and $L$ a $\sigma$-overfield of $K$.

(i) Any family of $\sigma$-generators of $L$ over $K$ contains a $\sigma$-transcendence basis of this difference field extension. If the $\sigma$-field $K$ is inversive and $L$ a $\sigma^*$-overfield of $K$, then any system of $\sigma^*$-generators of $L$ over $K$ contains a $\sigma$-transcendence basis of $L$ over $K$.

(ii) Let $\eta_1, \ldots, \eta_m \in L$. Then $\sigma$-trdeg$_K K(\eta_1, \ldots, \eta_m) \leq m$. If $K$ is inversive and $L$ a $\sigma^*$-overfield of $K$, then $\sigma$-trdeg$_K K(\eta_1, \ldots, \eta_m)^* \leq m$.

(iii) Let $\{\eta_1, \ldots, \eta_m\}$ and $\{\zeta_1, \ldots, \zeta_s\}$ be two finite subsets of $L$ such that $K(\eta_1, \ldots, \eta_m) = K(\zeta_1, \ldots, \zeta_s)$ or $K(\eta_1, \ldots, \eta_m)^* = K(\zeta_1, \ldots, \zeta_s)^*$ if $K$ is inversive and $L$ a $\sigma^*$-overfield of $K$. If the set $\{\zeta_1, \ldots, \zeta_s\}$ is $\sigma$-algebraically independent over $K$, then $s \leq m$.

(iv) Let $S = K(y_1, \ldots, y_s)$ be the ring of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $K$. If $k \neq s$, then $S$ cannot be a ring of $\sigma$-polynomials in $k$ $\sigma$-indeterminates over $K$. Similarly, if the difference field $K$ is inversive, then a ring of $\sigma^*$-polynomials $K(y_1, \ldots, y_s)^*$ cannot be a ring of $\sigma^*$-polynomials in $k$ $\sigma^*$-indeterminates over $K$ if $k \neq s$.

PROPOSITION 4.1.6. Let $H \subseteq K \subseteq L$ be difference (in particular, inversive difference) field extensions with the same basic set $\sigma$. Then $\sigma$-trdeg$_H L = \sigma$-trdeg$_K L + \sigma$-trdeg$_H K$.

4.2. Dimension polynomials of difference and inversive difference field extensions

The results of this section first appeared in [99–101,103], and [3–5]. Most of the proofs can be also found in [88, Chapter 6].

Let us consider $\mathbb{N}^n$ as a partially ordered set relative to the product order $\leq_P$ such that $(a_1, \ldots, a_n) \leq_P (b_1, \ldots, b_n)$ if and only if $a_i \leq b_i$ for $i = 1, \ldots, n$. For any $A \subseteq \mathbb{N}^n$ and $r \in \mathbb{N}$, let $A(r) = \{(e_1, \ldots, e_n) \in A \mid \sum_{i=1}^n e_i \leq r\}$. Furthermore, let $V_A = \{v = (v_1, \ldots, v_n) \in \mathbb{N}^n \mid$ there is no $a \in A$ such that $a \leq_P v\}$.

The following result is due to E. Kolchin (see [86, Chapter 0, Lemma 16]).

LEMMA 4.2.1. With the above notation, there exists a polynomial $\omega_A(t) \in \mathbb{Q}[t]$ with the following properties.

(i) $\omega_A(r) = \text{Card } V_A(r)$ for all sufficiently large $r \in \mathbb{N}$.

(ii) $\deg \omega_A \leq n$. 
The polynomial \( \omega_A(t) \) is called the Kolchin polynomial of the set \( A \subseteq \mathbb{N}^n \). Some methods and examples of computation of Kolchin polynomials can be found in [87] and [88, Chapter 2].

In [139] W. Sit proved that the set \( W \) of all Kolchin polynomials is well-ordered with respect to the order \( \preceq \) on \( \mathbb{Q}[t] \) such that \( f(t) \preceq g(t) \) if and only if \( f(r) \preceq g(r) \) for all sufficiently large \( r \in \mathbb{N} \). Furthermore, this set coincides with the set of all Hilbert polynomials of standard graded algebras over a field, as well as with the set of all differential dimension polynomials of finitely generated differential field extensions (see [108]).

In the rest of this section we assume that all fields have zero characteristic. The following theorem is a difference version of the Kolchin theorem on differential dimension polynomials [84].

**Theorem 4.2.2** [99]. Let \( K \) be a difference field with a basic set \( \sigma = \{ \alpha_1, \ldots, \alpha_n \} \), \( T \) the free commutative semigroup generated by \( \sigma \), and for any \( r \in \mathbb{N} \), \( T(r) = \{ \tau \in T \mid \text{ord } \tau \leq r \} \). Furthermore, let \( L = K(\eta_1, \ldots, \eta_s) \) be a \( \sigma \)-overfield of \( K \) generated by a finite family \( \eta = \{ \eta_1, \ldots, \eta_s \} \). Then there exists a polynomial \( \phi_{\eta|K}(t) \in \mathbb{Q}[t] \) with the following properties.

(i) \( \phi_{\eta|K}(r) = \text{trdeg}_K K(\{ \tau \eta_j \mid \tau \in T(r), 1 \leq j \leq s \}) \) for all sufficiently large \( r \in \mathbb{N} \).

(ii) \( \deg \phi_{\eta|K}(t) \leq n \) and the polynomial \( \phi_{\eta|K}(t) \) can be written as \( \phi_{\eta|K}(t) = \sum_{i=0}^n a_i t^{i+n} \) where \( a_0, \ldots, a_n \in \mathbb{Z} \).

(iii) The integers \( a_n, d = \deg \phi_{\eta|K}(t) \) and \( a_d \) are invariants of \( \phi_{\eta|K}(t) \), that is, they do not depend on the choice of a system of \( \sigma \)-generators \( \eta \). Furthermore, \( a_n = \sigma \)-trdeg \( K \).

(iv) Let \( P \) be the defining \( \sigma \)-ideal of \( \langle \eta_1, \ldots, \eta_s \rangle \) in the ring of \( \sigma \)-polynomials \( K[\{ y_1, \ldots, y_s \}] \) and let \( A \) be a characteristic set of \( P \) with respect to some orderly ranking of \( \{ y_1, \ldots, y_s \} \). Furthermore, for every \( j = 1, \ldots, s \), let \( E_j = \{ (k_1, \ldots, k_n) \in \mathbb{N}^n \mid \alpha_1^{k_1} \ldots \alpha_n^{k_n} y_j \text{ is a leader of a } \sigma \text{-polynomial from } A \} \). Then \( \phi_{\eta|K}(t) = \sum_{i=1}^s \omega_{E_j} (t) \) where \( \omega_{E_j}(t) \) is the Kolchin polynomial of the set \( E_j \).

The polynomial \( \phi_{\eta|K}(t) \) whose existence is established by Theorem 4.2.2 is called the difference (or \( \sigma \)-) dimension polynomial of the difference field extension \( L \) of \( K \) associated with the system of \( \sigma \)-generators \( \eta \). The integers \( d = \deg \phi_{\eta|K}(t) \) and \( a_d \) are called, respectively, the difference (or \( \sigma \)-) type and typical difference (or \( \sigma \)-) transcendence degree of \( L \) over \( K \). These invariants of \( \phi_{\eta|K}(t) \) are denoted by \( \sigma \)-type \( L \) and \( \sigma \cdot t \)-trdeg \( L \), respectively.

**Theorem 4.2.3.** Let \( K \) be a difference field with a basic set \( \sigma = \{ \alpha_1, \ldots, \alpha_n \} \) and let \( L \) be a finitely generated \( \sigma \)-field extension of \( K \) with a set of \( \sigma \)-generators \( \eta = \{ \eta_1, \ldots, \eta_s \} \) such that \( \{ \eta_1, \ldots, \eta_d \} \) is a \( \sigma \)-transcendence basis of \( L \) over \( K \) \((1 \leq d \leq s)\). Then \( \phi_{\{ \eta_d+1, \ldots, \eta_s \}|K}(\eta_1, \ldots, \eta_d)(t) \leq \phi_{\eta|K}(t) - d^{(1+n)} \).
**Theorem 4.2.4.** Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $L = K(\eta_1, \ldots, \eta_s)$, and $\phi_{\eta|K}(t)$ the $\sigma$-dimension polynomial of the extension $L/K$ associated with the family of $\sigma$-generators $\eta = \{\eta_1, \ldots, \eta_s\}$. Then $\phi_{\eta|K}(t) = m^{(\sigma^n)}_m$ ($m \in \mathbb{N}$ if and only if $\sigma \cdot \text{trdeg}_K L = \text{trdeg}_K K(\eta_1, \ldots, \eta_s) = m$.

The following theorem gives versions of Theorems 4.2.2–4.2.4 for finitely generated inverse difference field extensions.

**Theorem 4.2.5** ([101], [88, Theorems 6.4.8, 6.4.16 and 6.4.17]). Let $K$ be an inverse difference field with basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $L$ the free commutative group generated by $\sigma$ and for any $r \in \mathbb{N}$, and $\Gamma(r) = \{\gamma = \alpha_1^{k_1} \ldots \alpha_n^{k_n} \in \Gamma \mid \text{ord} \gamma = \sum_{i=1}^n |k_i| \leq r\}$. Furthermore, let $L = K(\eta_1, \ldots, \eta_s)$ be a $\sigma^*$-overfield of $K$ generated by a finite family $\eta = \{\eta_1, \ldots, \eta_s\}$. Then there exists a polynomial $\psi_{\eta|K}(t) \in \mathbb{Q}[t]$ with the following properties.

(i) $\psi_{\eta|K}(r) = \text{trdeg}_K K(\{\gamma \eta_j \mid \gamma \in \Gamma(r), 1 \leq j \leq s\})$ for all sufficiently large $r \in \mathbb{N}$.

(ii) $\deg \psi_{\eta|K}(t) \leq n$ and the polynomial $\psi_{\eta|K}(t)$ can be written as $\psi_{\eta|K}(t) = \sum_{i=0}^n a_i t^{(i+1)}$ where $a_0, \ldots, a_n \in \mathbb{Z}$.

(iii) The integers $a_i, d = \deg \phi_{\eta|K}(t)$ and $a_d$ do not depend on the choice of a system of $\sigma$-generators $\eta$. Furthermore, $a_0 = \sigma \cdot \text{trdeg}_K L$.

(iv) If $\{\eta_1, \ldots, \eta_d\}$ is a $\sigma$-transcendence basis of $L$ over $K$ ($1 \leq d \leq s$), then $\psi_{\eta_1(\ldots), \eta_d(\ldots)}(t) \leq \psi_{\eta|K}(t) - d \sum_{i=0}^n (-1)^{n-i} 2^{n-i} t^{(i+1)}$.

(v) $\psi_{\eta|K}(t) = m \sum_{i=0}^n (-1)^n - 2^{n-i} t^{(i+1)}$ for some $m \in \mathbb{N}$ if and only if

$$\sigma \cdot \text{trdeg}_K L = \text{trdeg}_K K(\eta_1, \ldots, \eta_s) = m.$$

The polynomial $\psi_{\eta|K}(t)$ is called the $\sigma^*$-dimension polynomial of the $\sigma^*$-field extension $L/K$ associated with the system of $\sigma^*$-generators $\eta$. The integers $d = \deg \psi_{\eta|K}(t)$ and $a_d$ are called, respectively, the $\sigma^*$-type and typical $\sigma^*$-transcendence degree of $L$ over $K$.

These invariants of $\psi_{\eta|K}(t)$ are denoted by $\sigma^*$-type$_L L$ and $\sigma^*$-t.trdeg$_L L$, respectively.

As in the last part of Theorem 4.2.2, the polynomial $\psi_{\eta|K}(t)$ can be expressed as a sum of certain analogs of Kolchin polynomials. Let $A \subseteq \mathbb{Z}^n$ ($n \geq 1$) and let $\mathbb{Z}_1, \ldots, \mathbb{Z}_2^n$ be the orthants of $\mathbb{Z}^n$ (introduced just before Definition 2.3.5). Let us consider the following partial order $\preceq$ on $\mathbb{Z}^n$: $(a_1, \ldots, a_n) \preceq (b_1, \ldots, b_n)$ if and only if $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ belong to the same orthant and $|a_i| \leq |b_i|$ for $i = 1, \ldots, n$. As in the case of subsets of $\mathbb{N}^n$, one can show that if $W_A = \{w \in W \mid a \npreceq w \text{ for any } a \in A\}$ and $W_A(r) = \{|w_1, \ldots, w_n| \in W \mid \sum_{i=1}^n |w_i| \leq r \} (r \in \mathbb{N})$, then there exists a polynomial $\psi_A(t) \in \mathbb{Q}[t]$ such that $\psi_A(r) = \text{Card} W_A(r)$ for all sufficiently large $r \in \mathbb{N}$. Furthermore, $\deg \psi_A \leq n$, and $\deg \psi_A = n$ if and only if $A = \emptyset$; in the last case $\psi_A(t) = \sum_{i=0}^n (-1)^{n-i} 2^{n-i} t^{(i+1)}$.

Some properties and methods of computation of polynomials $\psi_A(t)$ ($A \subseteq \mathbb{Z}^n$) can be found in [87] and [88, Chapter 2].

**Theorem 4.2.6.** With the notation of Theorem 4.2.5, let $P$ be the defining $\sigma^*$-ideal of the $s$-tuple $(\eta_1, \ldots, \eta_s)$ in the ring of $\sigma^*$-polynomials $K\{y_1, \ldots, y_s\}$ and let $A$ be a characteristic set of $P$ with respect to some orderly ranking of $\{y_1, \ldots, y_s\}$. Furthermore, for every
Theorem 4.2.7. With the above notation, let $\Omega_K(L)_r (r \in \mathbb{N})$ denote the vector $L$-subspace of $\Omega_K(L)$ generated by the set $\{d\gamma(\eta_j) \mid \gamma \in \Gamma(r), 1 \leq i \leq s\}$ and let $\Omega_K(L)_r = 0$ for $r < 0$. Then

(i) $(\Omega_K(L)_r)_{r \in \mathbb{Z}}$ is an excellent filtration of the $\sigma^*$-module $\Omega_K(L)$.

(ii) $\dim_K \Omega_K(L)_r = \text{trdeg}_K K(\{\eta^j_j \mid \gamma \in \Gamma(r), 1 \leq j \leq s\})$ for all $r \in \mathbb{N}$.

(iii) The $\sigma^*$-dimension polynomial $\phi_{\eta_j}K(t)$ is equal to the $\sigma^*$-dimension polynomial of $\Omega_K(L)$ associated with the filtration $((\Omega_K(L)_r)_{r \in \mathbb{Z}}$.

The last theorem allows one to reduce the computation of a $\sigma^*$-dimension polynomial of an inversive $\sigma^*$-field extension $L/K$ to the computation of a $\sigma^*$-dimension polynomial of the vector $\sigma^*$-space $\Omega_K(L)$. The corresponding algorithms can be found in [88, Chapters 6 and 9], [89–91,119], and [127].

Remark 4.2.8. By analogy with the theory of differential fields, one can expect that if $L$ is a finitely generated $\sigma^*$-field extension of an inversive difference field $K$ with a basic set $\sigma$ and $\sigma$-trdeg$_K L = 0$, then there is a set $\sigma_1 = \{\beta_1, \ldots, \beta_n\}$ of automorphisms of $L$ such that $\sigma_1$ is equivalent to $\sigma$ and $L$ is a finitely generated as a $\{\beta_1, \ldots, \beta_{n-1}\}^*$-overfield of $K$. This is not true (see [88, Example 6.4.18]), but we have a weaker version of such a statement proved in [88, Section 6.4].
Theorem 4.2.9. Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $L$ a finitely generated $\sigma^*$-field extension of $K$, and $d = \sigma^*$-type$_L$ $L$. Then there exists a set $\sigma_1 = \{\beta_1, \ldots, \beta_n\}$ of automorphisms of $L$ and a finite family $\{\xi_1, \ldots, \xi_q\} \in L$ such that

(a) $\sigma_1$ is equivalent to $\sigma$;

(b) if $K$ is treated as an inversive difference field with the basic set $\sigma_2 = \{\beta_1, \ldots, \beta_d\}$ and $H$ is the $\sigma_2^*$-field extension of $K$ generated by $\xi_1, \ldots, \xi_q$, then $L$ is an algebraic extension of $H$.

We conclude this section with a result on multivariable dimension polynomials that generalizes Theorem 4.2.2.

Let $K$ be a difference field whose basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ is a union of $p$ disjoint finite sets ($p \geq 1$): $\sigma = \sigma_1 \cup \cdots \cup \sigma_p$, where $\sigma_1 = \{\alpha_1, \ldots, \alpha_{n_1}\}$, $\sigma_2 = \{\alpha_{n_1+1}, \ldots, \alpha_{n_1+n_2}\}$, $\sigma_p = \{\alpha_{n_1+\cdots+n_{p-1}+1}, \ldots, \alpha_n\}$, $n_1, \ldots, n_p \in \mathbb{N}$. Let $T$ be the free commutative semigroup generated by $\sigma$, and for any $(r_1, \ldots, r_p) \in \mathbb{N}^p$, let $T(r_1, \ldots, r_p) = \{\theta \in T \mid \text{ord}_i \theta \leq r_i \text{ for } i = 1, \ldots, p\}$.

Theorem 4.2.10. With the above notation, let $L = K(\eta_1, \ldots, \eta_s)$ be a $\sigma^*$-field extension of $K$ generated by a finite set $\eta = \{\eta_1, \ldots, \eta_s\}$. Then there exists a polynomial $\Phi_\eta(t_1, \ldots, t_p)$ in $p$ variables $t_1, \ldots, t_p$ with rational coefficients such that

(i) $\Phi_\eta(r_1, \ldots, r_p) = \text{trdeg}_K K(\bigcup_{j=1}^{\eta_i} T(r_1, \ldots, r_p)\eta_j)$ for all sufficiently large $(r_1, \ldots, r_p) \in \mathbb{N}^p$;

(ii) $\text{deg}_T \Phi_\eta \leq n_i \ (i = 1, \ldots, p)$ and the polynomial $\Phi_\eta$ can be written as $\Phi_\eta(t_1, \ldots, t_p) = \sum t_1 \cdots t_p \alpha_{i_1\cdots i_p}(t_1^{i_1} \cdots t_p^{i_p}) \ (i_1, \ldots, i_p) \in \mathbb{Z}$ for all $i_1, \ldots, i_p$.

(iii) Let $E_\eta = \{(i_1, \ldots, i_p) \in \mathbb{N}^p \mid 0 \leq i_k \leq n_k \ (k = 1, \ldots, p) \text{ and } a_{i_1\cdots i_p} \neq 0\}$. Then the total degree $d$ of the polynomial $\Phi$, $a_{i_1\cdots i_p}$-tuples $(j_1, \ldots, j_p) \in E_\eta$ (we use the notation of Theorem 3.3.8), the corresponding coefficients $a_{j_1, \ldots, j_p}$, and the coefficients of the terms of total degree $d$ do not depend on the choice of the system of $\sigma^*$-generators $\eta$.

Theorem 4.2.10, as well as its analog for finitely generated inversive difference field extensions, can be proven in the same way as the corresponding results on multivariable differential and difference-differential dimension polynomials obtained in [107] and [109]. These papers also show that multivariable dimension polynomials can carry essentially more invariants of a difference field extension than the dimension polynomials obtained in Theorems 4.2.2 and 4.2.5.

Dimension polynomials and the strength of a system of difference equations

Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $E$ the corresponding ring of $\sigma^*$-operators, $K[y_1, \ldots, y_s]^*$ an algebra of $\sigma^*$-polynomials in $\sigma^*$-indeterminates $y_1, \ldots, y_s$, and $P$ a prime $\sigma^*$-ideal of $K[y_1, \ldots, y_s]^*$. If $\eta = (\eta_1, \ldots, \eta_s)$ is a generic zero of $P$, then the dimension polynomial $\psi_\eta|_K(t)$ associated with the $\sigma^*$-field extension $K(\eta_1, \ldots, \eta_s)/K$ is called the $\sigma^*$-dimension polynomial of the ideal $P$; it is denoted by $\psi_P(t)$. (Clearly, if $\eta$ and $\xi$ are two generic zeros of $P$, then $\psi_\eta|_K(t) = \psi_\xi|_K(t)$, so
that the $\sigma^*$-dimension polynomial of $P$ is well-defined.) It can be shown (see [88, Proposition 6.2.4]) that if $P_1$ and $P_2$ are prime $\sigma^*$-ideals of $K[y_1, \ldots, y_s]^*$ such that $P_1 \subseteq P_2$, then $\psi_p(t) < \psi_p(t)$ (that is, $\psi_p(r) < \psi_p(r)$ for all sufficiently large $r \in \mathbb{N}$).

If $\Phi = \{A_\lambda \mid \lambda \in \Lambda\}$ is a family of $\sigma^*$-polynomials in $K[y_1, \ldots, y_s]^*$, then an $s$-tuple $\eta$ that annihilates every $A_\lambda$ is said to be a solution of the system of algebraic difference (or $\sigma^*$-) equations $A_\lambda(y_1, \ldots, y_s) = 0 (\lambda \in \Lambda)$. By Theorem 2.4.5, the last system is equivalent to some its finite subsystem, that is, there is a finite set $\Phi_0 = \{A_1, \ldots, A_m\} \subseteq \Phi$ such that the set of solutions of the original system coincides with the set of solutions of the system

$$A_i(y_1, \ldots, y_s) = 0 \quad (i = 1, \ldots, m). \quad (4.2.1)$$

A system of algebraic $\sigma^*$-equations (4.2.1) is called prime if the perfect $\sigma^*$-ideal $\{A_1, \ldots, A_m\}$ of the $\sigma^*$-ring $K[y_1, \ldots, y_s]^*$ is prime. Note that any linear homogeneous system of difference equations (that is, a system of the form $\sum_{j=1}^s w_{ij} y_j = 0$ ($i = 1, \ldots, m$) where $w_{ij} \in \mathcal{E}$) is prime.

If (4.2.1) is a prime system of $\sigma^*$-equations, then the $\sigma^*$-dimension polynomial $\psi_p(t)$ of the prime $\sigma^*$-ideal $P = \{A_1, \ldots, A_m\}$ is called the $\sigma^*$-dimension polynomial of the system. This polynomial is an algebraic version of the concept of strength of a system of equations in finite differences defined as follows (by analogy with the similar notion for a system of differential equations introduced and studied by A. Einstein, [45]). Let us consider a system of equations in finite differences with respect to $s$ unknown grid functions $f_1, \ldots, f_s$ of $n$ real variables with coefficients from a function field $F$. Suppose that the difference grid, whose nodes form the domain of the considered functions, has equal cells of dimension $h_1 \times \cdots \times h_n$ ($h_1, \ldots, h_n \in \mathbb{R}$) and fills the whole space $\mathbb{R}^n$. Furthermore, let us fix some node $X_0$ and say that a node $X$ of the grid is a node of order $i$ (with respect to $X_0$) if the shortest route between $X$ and $X_0$ passing along the edges of the grid consists of $i$ steps ($i = 0, 1, \ldots$). (By a step we mean a path from a node to a neighboring node along the edge between them.)

Let us consider the values of the grid functions $f_1, \ldots, f_s$ at the nodes whose order does not exceed $i$ ($i \in \mathbb{N}$). If the functions $f_j$ ($1 \leq j \leq s$) do not satisfy any system of equations, their values at the nodes of any order may be chosen arbitrarily. Because of the system of equations in finite differences (and equations obtained from ones of the system by transformations of the form $f_j(x_1, \ldots, x_n) \mapsto f_j(x_1 + r_1 h_1, \ldots, x_n + r_n h_n)$ ($r_1, \ldots, r_n \in \mathbb{Z}$), the number of independent values of the functions $f_1, \ldots, f_s$ at the nodes of order less than or equal to $i$ decreases. This number $S_i$ is a function of $i$ called the strength of the system of equations.

Suppose that the mappings $\alpha_j : f(x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_j + h_j, \ldots, x_n)$ ($j = 1, \ldots, n$) are automorphisms of the field of coefficients $F$. Then $F$ can be treated as an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$. If we replace the unknown functions $f_k$ by the $\sigma^*$-indeterminates $y_k$ ($1 \leq k \leq s$) from the algebra of $\sigma^*$-polynomials $F[y_1, \ldots, y_s]^*$, then the given system of equations in finite differences generates a system of algebraic $\sigma^*$-equations of the form (4.2.1). Suppose that the last system is prime (e.g., linear), so that the $\sigma^*$-polynomials in its left-hand sides generate a prime $\sigma^*$-ideal $P$ in $F[y_1, \ldots, y_s]^*$. Then the $\sigma^*$-dimension polynomial $\psi_p(t)$ of this ideal is said to be the $\sigma^*$-dimension polynomial of the original system of equations in finite differences. It is
easy to see that $\psi_P(i) = S_i$ for all $i \in \mathbb{N}$, so the strength of the system of equations in finite differences can be determined if one can find the polynomial $\psi_P(t)$ (that is, the dimension polynomial of the $\sigma^*$-field extension $F(\eta_1, \ldots, \eta_s)/F$ where $(\eta_1, \ldots, \eta_s)$ is a generic zero of the ideal $P$). A number of examples of computation of the strength of systems of difference equations can be found in [88, Chapters 6 and 9], [90, 91, 119], and [127].

4.3. Finitely generated difference and inversive difference field extensions. Limit degree.

Finitely generated difference algebras

Let $K$ be an ordinary difference ring with a basic set $\sigma = \{\alpha\}$ and $L = K(\eta_1, \ldots, \eta_s)$ a $\sigma$-overfield of $K$ generated by a finite set $\eta = \{\eta_1, \ldots, \eta_s\}$. Furthermore, for any $r = 1, 2, \ldots$, let $L_r = K(\alpha'(\eta_j) \mid 1 \leq j \leq s, 0 \leq i \leq r))$ and $d_r = L_r : L_{r-1}$, the dimension of $L_r$ as a vector $L_{r-1}$-space. It is easy to see that $d_1 \geq d_2 \geq \cdots$. Moreover, as it is shown in [30], $\min\{d_r \mid r = 1, 2, \ldots\}$ does not depend on the choice of the system of $\sigma$-generators $\eta$. This minimum value is called the limit degree of the $\sigma$-field extension $L/K$, it is denoted by $\text{ld} L/K$.

If a $\sigma$-field extension $L$ of a difference ($\sigma$-) field $K$ is not finitely generated, we define its limit degree $\text{ld} L/K$ as the maximum of limit degrees of all finitely generated $\sigma$-field subextensions of $L/K$, if this maximum exists or $\infty$ if it does not. Clearly, if $\sigma$-trdeg$_K L > 0$, then $\text{ld} L/K = \infty$. If $\sigma$-trdeg$_K L = 0$ and $L$ is a finitely generated $\sigma$-field extension of $K$, then $\text{ld} L/K < \infty$.

The concept of limit degree was introduced by R.M. Cohn, [30]. The proofs of the following results on limit degrees of ordinary difference field extensions can be found in [32, Chapter 5].

**Theorem 4.3.1.** Let $K$ be an ordinary difference field with a basic set $\sigma$, $M$ a $\sigma$-overfield of $K$ and $L/K$ a $\sigma$-field subextension of $M/K$. Then $\text{ld} M/K = (\text{ld} M/L)(\text{ld} L/K)$.

**Theorem 4.3.2.** Let $K$ be an ordinary difference field and $L$ a difference overfield of $K$. Furthermore, let $K^*$ and $L^*$ denote the inversive closures of $K$ and $L$, respectively. Then $\text{ld} L^*/K = \text{ld} L^*/K^* = \text{ld} L/K$.

**Remark 4.3.3.** In the case of ordinary difference fields, there are several concepts related to the notion of limit degree. In the case of fields of characteristic $p > 0$, one can define an analog of limit degree using separable factor of degree in place of degree of $L_r$ over $L_{r-1}$ (we refer to the notation from the beginning of this section). The corresponding invariant of the difference field extension is called the reduced limit degree of $L/K$ and denoted by $\text{rd} L/K$; it has the same properties that are established for $\text{ld} L/K$ in Theorems 4.3.1 and 4.3.2.

If $K$ is an inversive difference field with a basic set $\sigma = \{\alpha\}$, then $K$ can be also treated as a difference field with the basic set $\sigma' = \{\alpha^{-1}\}$ called the inverse difference field of $K$. It is denoted by $K'$. Let $L$ be a $\sigma$-field extension of $K$ and $L'$ the inverse difference field of $L$ (so that $L'$ is a $\sigma'$-field extension of $K'$). The inverse limit degree of $L$ over $K$ is defined
to be $ld L'/K'$ (it is denoted by $ild L/K$), and the inverse reduced limit degree of $L$ over $K$ is defined to be $rld L'/K'$ (it is denoted by $irld L/K$). The analogs of Theorems 4.3.1 and 4.3.2 are valid for these concepts, as well.

**Theorem 4.3.4.** Let $K$ be an ordinary difference field with a basic set $\sigma$.

(i) If $L$ is a finitely generated $\sigma$-field extension of $K$, then $ld L/K = 1$ if and only if $L = K(S)$ for some finite set $S \subseteq L$.

(ii) The following two statements are equivalent:

(a) $L/K$ is a finitely generated $\sigma$-field extension, $L$ is algebraic over $K$, and $ld L/K = 1$.

(b) $L : K$ is finite.

**Theorem 4.3.5.** Let $K$ be an ordinary difference field and $L$ a difference field extension of $K$ which is algebraic over $K$. Then

(i) $ld L/K = ild L/K$ and $rld L/K = irld L/K$.

(ii) If $ld L/K = 1$ and $K$ is inversive, then $L$ is inversive.

The natural generalizations of the concept of limit degree to the case of partial difference fields were obtained in [50].

Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and $L$ a finitely generated $\sigma$-field extension of $K$: $L = K(S)$ where $S$ is a finite subset of $L$. In what follows we adopt the following notation. If $\sigma_1 = \{\alpha_1, \ldots, \alpha_{i_1}\} \subseteq \sigma$, then the set $\{\alpha_1^{j_1} \ldots \alpha_{i_1}^{j_{p_1}}(s) \mid j_1, \ldots, j_{p_1} \in \mathbb{N}, s \in S\}$ is denoted by $S^{(\alpha_1, \ldots, \alpha_{i_1}-1)}$. Furthermore, for any $k = 0, 1, \ldots$, we set $S_k = \bigcup_{j=0}^{k} \alpha_1^j (S^{(\alpha_1, \ldots, \alpha_{n-1})})$ (if $n = 1$, $S^{(\alpha_1, \ldots, \alpha_{n-1})} = S$), and for any positive integers $i_1, \ldots, i_t$ ($1 \leq t \leq n$), we set

$$S^*(i_n, i_{n-1} \ldots i_t) = \bigcup_{i=0}^{i_n-1} \alpha_1^{i_n}(S^{(\alpha_1, \ldots, \alpha_{n-1})}) \cup \bigcup_{i=0}^{i_{n-1}-1} \alpha_1^{i_n} \alpha_2^{i_1}(S^{(\alpha_1, \ldots, \alpha_{n-2})}) \cup \ldots$$

$$\cup \bigcup_{i=0}^{i_{t-1}-1} \alpha_1^{i_n} \ldots \alpha_t^{i_1}(S^{(\alpha_1, \ldots, \alpha_{t-1})})$$

and $S(i_n, i_{n-1} \ldots i_t) = S^*(i_n, i_{n-1} \ldots i_{t+1}, i_{t+1} + 1)$. Finally, if $K$ is treated as a difference field with a basic set $\sigma_1$, it is denoted by $(K; \sigma_1)$ or $(K; \alpha_1, \ldots, \alpha_{i_1})$; the difference transcendence degree of $(K(S_{k_1}); \alpha_1, \ldots, \alpha_{n-1})$ over $(K(S_{k-1}); \alpha_1, \ldots, \alpha_{n-1})$ is denoted by $\delta_k$. (If $n = 1$, then $\delta_k = \operatorname{trdeg}_K K(S_{k-1})$.)

With the above notation, the limit transforal transcendence degree $\sigma_1$-$\operatorname{trdeg}_K L$ of $L$ over $K$ is defined by $\sigma_1$-$\operatorname{trdeg}_K L = \min\{\delta_k \mid k = 1, 2, \ldots\}$.

**Proposition 4.3.6** [50]. Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and $L$ a finitely generated $\sigma$-field extension of $K$. Then

(i) $\sigma_1$-$\operatorname{trdeg}_K L = \sigma$-$\operatorname{trdeg}_K L$. Thus, $\sigma_1$-$\operatorname{trdeg}_K L$ is independent of the finite set of $\sigma$-generators and the translation from $\sigma$ chosen as $\alpha_n$ to define $\sigma_1$-$\operatorname{trdeg}_K L$. 

(ii) If $S$ is a finite set of $\sigma$-generators of $L$ over $K$, then there exists a finite subset $Z \subseteq S$ and positive integers $k_1, \ldots, k_n$ with the following properties:

(a) $Z$ is a $\sigma$-transcendence basis of $L$ over $K$.

(b) If $t \in \{1, \ldots, n\}$, $i_j \geq k_j$ for $j = 1, \ldots, n$, and $\alpha' = [\alpha_1, \ldots, \alpha_{t-1}]$, then $\alpha_n^t \cdots \alpha_1^t(Z)$ is a $\sigma'$-transcendence basis of $(K(S_1, \ldots, i_1); \alpha_1, \ldots, \alpha_{t-1})$ over $(K(S_1, \ldots, i_1); \alpha_1, \ldots, \alpha_{t-1})$ (and $\sigma$-trdeg$_K L$ is the $\sigma'$-transcendence degree of this extension).

With the notation of the last proposition, a set $Z$ that satisfies conditions (a) and (b) for some positive integers $k_1, \ldots, k_n$ is called a limit basis of transfinite transcendence of $L$ over $K$.

In [50] P. Evanovich introduced an invariant $ld_n(M/K)$ of a partial difference field extension $M/K$ that can be viewed as a generalization of the concept of limit degree. $ld_n(M/K)$ is inductively defined as an element of the set $\mathbb{N} \cup \{\infty\}$ that satisfies the following conditions (ld1)–(ld5). Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $M$ a $\sigma$-overfield of $K$ and $L/K$ a $\sigma$-field subextension of $M/K$. Then

(1) If there exists a finite set $S \subseteq M$ such that $M = L(S)$, then there exists a finitely generated $\sigma$-overfield $K'$ of $K$ contained in $L$ such that $ld_n(M/L) = ld_n(K'(S)/K')$.

(2) If $S \subseteq M$, then $ld_n(L(S)/K(S)) \leq ld_n(L/K)$ and $ld_n(L(S)/L) \leq ld_n(K(S)/K)$.

Equality will hold in both if $S$ is $\sigma$-algebraically independent over $L$.

(3) If there is a $\sigma$-isomorphism $\phi$ of $L$ onto a $\sigma$-field $L'$ and $K'$ is a $\sigma$-subfield of $L'$ such that $\phi(K) = K'$, then $ld_n(L(K)/K) = ld_n(L'/K')$.

(4) If the $\sigma$-field extension $L/K$ is finitely generated, then $\sigma$-trdeg$_K L = 0$ if and only if $ld_n(L/K) < \infty$.

(5) $ld_n(M/K) = ld_n(M/L) \cdot ld_n(L/K)$.

If $n = 1$, $ld_1$ is defined to be the limit degree $ld$ for ordinary difference fields. Suppose that $ld_{n-1}$ is defined for difference field extensions whose basic sets consist of $n - 1$ translations. Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and $L$ a $\sigma$-field extension of $K$. Assume first that $L/K$ is finitely generated, say $L = K(S)$ for some finite set $S \subseteq L$. For any $m \in \mathbb{N}$, let $L_m = (K(S_1, \ldots, \alpha_{n-1})$ and $L_m$ is finitely generated extension of $L_{m-1}$ (we set $K(S_1) = K$). Applying (ld3) and (ld2) we obtain that $ld_n(L_m/L_m) = ld_{n-1}(L_m/L_{m-1})$, whence there exists the limit $\lim_{m \to \infty} ld_{n-1}(L_m/L_{m-1}) = a$ where $a \in \mathbb{N}$ or $a = \infty$. As in the case of limit degree of ordinary difference fields, [30], one can show that $a$ is independent on the choice of $\sigma$-generators of $L/K$. Now we define $ld_n(L/K) = a$.

If $L/K$ is not finitely generated, $ld_n(L/K)$ is defined to be the maximum of $ld_n(K'/K)$ where $K'/K$ is a finitely generated $\sigma$-field subextension of $L/K$, if the maximum exists and $\infty$ if it does not. As in the case of limit degrees of ordinary difference fields, if $L/K$ itself is finitely generated, then $ld_n(L/K)$ is the maximum of $ld_n(K'/K)$ where $K'/K$ is a finitely generated $\sigma$-field subextension of $L/K$. The proof of the fact that the so defined $ld_n$ satisfies (ld1)–(ld5) can be found in [50, section 3]. In the same paper P. Evanovich used the properties of $ld_n$ to prove the following fundamental result that was first established by R.M. Cohn, [32, Chapter 5], for ordinary difference fields.
Theorem 4.3.7. Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $M$ a finitely generated $\sigma$-overfield of $K$ and $L/K$ a $\sigma$-field subextension of $M/K$. Then the $\sigma$-field extension $L/K$ is finitely generated.

We conclude this section with some results on finitely generated difference algebras.

Let $K$ be a difference (inversive difference) ring with a basic set $\sigma$. A $K$-algebra $R$ is said to be a difference algebra over $K$ or a $\sigma$-algebra (respectively, an inversive difference algebra over $K$ or a $\sigma^*$-algebra) if elements of $\sigma$ (respectively, $\sigma^*$) act on $R$ in such a way that $R$ is a $\sigma$- (respectively, $\sigma^*$-) ring and $\alpha(au) = \alpha(a)\alpha(u)$ for any $a \in K, u \in R$, $\alpha \in \sigma$ (for any $\alpha \in \sigma^*$ if $R$ is a $\sigma^*$-algebra). A $\sigma$-algebra (respectively, $\sigma^*$-algebra) $R$ is said to be finitely generated if there exists a finite family $\{\eta_1, \ldots, \eta_s\}$ of elements of $R$ such that $R = K[\eta_1, \ldots, \eta_s]$ (respectively, $R = K[\eta_1, \ldots, \eta_s]^*$). If a $\sigma$-algebra (or $\sigma^*$-algebra) $R$ is an integral domain, then the $\sigma$-transcendence degree of the corresponding $\sigma$- (respectively, $\sigma^*$-) field of quotients of $R$ over $K$.

In what follows we consider inversive difference algebras over inversive fields. All results formulated below for such algebras remain valid for difference algebras over difference fields (with replacement of the prefix $\sigma^*$- by $\sigma$-).

Let $R$ be an inversive difference algebra over an inversive difference field $K$ with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$ and let $U$ denote the set of all prime $\sigma^*$-ideals of $R$. As at the end of section 3.1, one can consider the set $B_U = \{(P, Q) \in U \times U \mid P \supseteq Q\}$ and the uniquely defined mapping $\mu_U : B_U \to \mathbb{Z}$ such that

(i) $\mu_U(P, Q) \geq -1$ for every pair $(P, Q) \in B_U$;
(ii) for any $d \in \mathbb{N}$, the inequality $\mu_U(P, Q) \geq d$ holds if and only if $P \neq Q$ and there exists an infinite chain $P = P_0 \supseteq P_1 \supseteq \cdots \supseteq Q$ such that $P_i \in U$ and $\mu_U(P_{i-1}, P_i) \geq d - 1$ for $i = 1, 2, \ldots$.

The type and dimension of the $\sigma^*$-algebra $R$ over $U$ are defined, respectively, as $\text{sup}\{\mu_U(P, Q) \mid (P, Q) \in B_U\}$ and the least upper bound of the lengths $k$ of chains $P_0 \supseteq P_1 \supseteq \cdots \supseteq P_k$ such that $P_0, \ldots, P_k \in U$ and $\mu_U(P_{i-1}, P_i) = \text{type}_UR$ for $i = 1, \ldots, k$.

These characteristics are denoted by $\text{type}_UR$ and $\text{dim}_UR$, respectively.

Theorem 4.3.8 [102]. Let $K$ be an inversive difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$, $R = K[\eta_1, \ldots, \eta_s]^*$ a $\sigma^*$-algebra without zero divisors generated by a finite set $\eta = \{\eta_1, \ldots, \eta_s\}$, and $U$ the family of all prime $\sigma^*$-ideals of $R$. Then:

(i) $\text{type}_UR \leq n$.
(ii) If $\sigma^*$-trdeg$_K R = 0$, then $\text{type}_UR < n$.
(iii) If $\text{type}_UR = n$, then $\text{dim}_UR \leq \sigma^*$-trdeg$_K R$.
(iv) If $\eta_1, \ldots, \eta_s$ are $\sigma$-algebraically independent over $K$, then $\text{type}_UR = n$ and $\text{dim}_UR = s$.

Let $K$ be an inversive difference field with a basic set $\sigma$. A $\sigma^*$-algebra $R$ is said to be a local $\sigma^*$-algebra if $R$ is a local ring (in this case its maximal ideal is a $\sigma^*$-ideal). We say that $R$ is a local $\sigma^*$-algebra of finitely generated type if $R$ is a local $\sigma^*$-algebra and there exist a finite set $\{\eta_1, \ldots, \eta_s\} \subseteq R$ such that $R = K[\eta_1, \ldots, \eta_s]_m$ where $m$ is the maximal ideal of $R$. 
Difference algebra

THEOREM 4.3.9 [102]. Let $K$ be an inversive difference field of zero characteristic with a basic set $\sigma$. Let an integral domain $R$ be a local $\sigma^*-K$-algebra of finitely generated type, $m$ the maximal ideal of $R$, and $k = R/m$ the corresponding $\sigma^*$-field of residue classes. Then:

(i) $m/m^2$ is a finitely generated vector $\sigma^*-k$-space.

(ii) $\dim_k m/m^2 \geq \trdeg_k R - \trdeg_k k.$

4.4. Difference kernels. Realizations

Let $F$ be an ordinary inversive difference field with a basic set $\sigma = \{a\}$. A difference kernel of length $r$ over $F$ is an ordered pair $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ where each $a_i$ is itself an $s$-tuple $(a_i^{(1)}, \ldots, a_i^{(s)})$ over $F$ (a positive integer $s$ is fixed) and $\tau$ is an extension of $a$ to an isomorphism of $F(a_0, \ldots, a_{r-1})$ onto $F(a_1, \ldots, a_r)$ such that $\tau a_i = a_{i+1}$ for $i = 0, \ldots, r - 1$. (In other words, $\tau(a_i^{(j)}) = a_{i+1}^{(j)}$ for $0 \leq i \leq r - 1, 1 \leq j \leq s$.) If $r = 0$, then $\tau = a$. The degree of transcendence of the difference kernel $\mathcal{R}$ is defined to be $\trdeg F(a_0, \ldots, a_{r-1}) F(a_0, \ldots, a_r)$; it is denoted by $\delta \mathcal{R}$. Below, while considering difference kernels over a difference field $F$, we always assume that $F$ is inversive.

A prolongation $\mathcal{R}'$ of a difference kernel $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ is a difference kernel of length $r + 1$ consisting of an overfield $F(a_0, \ldots, a_r, a_{r+1})$ of $F(a_0, \ldots, a_r)$ and an extension $\tau'$ of $\tau$ to an isomorphism of $F(a_0, \ldots, a_r, a_{r+1})$ onto $F(a_1, \ldots, a_{r+1})$.

In what follows, a set of the form $a = \{a^{(i)} | i \in I\}$ will be referred to as an indexing of $a$ (with the index set $I$). If $J \subseteq I$, the set $\{a^{(i)} | i \in J\}$ will be called a subindexing of $a$. If $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ is a difference kernel and $\tilde{a}_0$ is a subindexing of $a_0$ (that is, $\tilde{a}_0 = (a_0^{(i_1)}, \ldots, a_0^{(i_q)}), 1 \leq i_1 < \cdots < i_q \leq s$), then $\tilde{a}_k$ ($k = 1, \ldots$) will denote the corresponding subindexing of $a_k$. The proofs of the following results on prolongations of difference kernels over ordinary difference fields can be found in [32, Chapter 6].

THEOREM 4.4.1. Every difference kernel $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ over an ordinary difference field $F$ has a prolongation $\mathcal{R}' = (F(a_0, \ldots, a_r, a_{r+1}), \tau')$. Moreover, one can choose a prolongation $\mathcal{R}'$ with the following properties.

(i) If $\tilde{a}_0$ is a subindexing of $a_0$ such that $\tilde{a}_r$ is algebraically independent over $F(a_0, \ldots, a_{r-1})$, then $\tilde{a}_{r+1}$ is algebraically independent over $F(a_0, \ldots, a_r)$ (so that $\delta \mathcal{R} = \delta \mathcal{R}'$).

(ii) If $\tilde{a}_0$ is as in (i), then $\bigcup_{i=0}^{r+1} \tilde{a}_i$ is algebraically independent over $F$.

With the above notation, a prolongation $\mathcal{R}'$ of a difference kernel $\mathcal{R}$ is called generic if $\delta \mathcal{R} = \delta \mathcal{R}'$.

A generic prolongation of a difference kernel $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ over an ordinary difference field can be constructed as follows. Let $P$ be the prime ideal with generic zero $a_r$ of a polynomial ring $F(a_0, \ldots, a_{r-1})[X_1, \ldots, X_s]$ in $s$ indeterminates $X_1, \ldots, X_s$. Let $P'$ be obtained from $P$ by replacing the coefficients of the polynomials of $P$ by their images under $\tau$. Then $P'$ is a prime ideal of $F(a_1, \ldots, a_r)[X_1, \ldots, X_s]$ and generates an ideal $\tilde{P}$ in $F(a_0, \ldots, a_r)[X_1, \ldots, X_s]$. Let $Q$ be an essential prime divisor of $\tilde{P}$ in the
last ring and let \( a_{r+1} \) be a generic zero of \( Q \). Then \( a_{r+1} \) is also a generic zero of \( P' \) and there is an isomorphism \( \tau' : F(a_0, \ldots, a_r) \to F(a_1, \ldots, a_{r+1}) \) that extends \( \tau \). We obtain the desired generic prolongation. Conversely, if \( \mathcal{R}' = (F(a_0, \ldots, a_{r+1}), \tau') \) is any generic prolongation of \( \mathcal{R} \), then \( a_{r+1} \) is a solution of the ideal \( \tilde{P} \) and hence of one of its essential prime divisors \( Q \). It follows that \( \dim \mathcal{R}' = \dim P = \text{trdeg}_\mathcal{P}(F(a_0, \ldots, a_r)) = \text{trdeg}_\mathcal{P}(F(a_0, \ldots, a_r)) = F(a_0, \ldots, a_{r+1}) \), so that \( a_{r+1} \) is a generic zero of \( Q \).

**Example 4.4.2** [32, Chapter 6, Section 2]. Let \( F \) be an ordinary difference field with a basic set \( \sigma = \{a_1, F\{y\} \} \) the ring of \( \sigma \)-polynomials in one \( \sigma \)-indeterminate \( y \) over \( F \), and \( A \) an algebraically irreducible \( \sigma \)-polynomial from \( F\{y\} \) (that is, \( A \) is irreducible as a polynomial in \( y, a(y), a^2(y), \ldots \)). Assuming that \( A \) contains \( y \) and \( a^m y, m > 0 \), is the highest transform of \( y \) in \( A \), we shall use prolongations of difference kernels to construct a solution of \( A \). First, let us consider an \( m \)-tuple \( a = (a^{(1)}, \ldots, a^{(m)}) \) whose coordinates constitute an algebraically independent set over \( F \). Now we define an \( m \)-tuple \( a_1 \) as follows: we set \( a_1^{(i)} = a^{(i+1)} \) for \( i = 1, \ldots, m - 1 \), replace \( y_{i-1} \) by \( a^{(i)} \) in \( A \) (\( 1 \leq i \leq m \)), find a solution of the resulting polynomial in one unknown \( y_i \), and take it as \( a^{(m)} \). Since \( A \) involves \( y \), \( a^{(1)} \) will be algebraically dependent on \( a^{(2)}, \ldots, a^{(m)} \), \( a^{(m)}_1 \), \( a^{(m)}_2 \) over \( F \). Therefore, \( a^{(1)}, \ldots, a^{(m)}_1 \) are algebraically independent over \( F \) whereas there is a difference kernel \( \mathcal{R}_1 \) defined over \( F \) by the extension of the translation \( a \) to an isomorphism \( \tau_0 : F(a) \to F(a_1) \). By successive applications of Theorem 4.4.1 we find a sequence \( a_0 = a, a_1, \ldots \) such that a kernel \( \mathcal{R}_{k+1} \) is defined by an isomorphism \( \tau_k : F(a_0, \ldots, a_k) \to F(a_1, \ldots, a_{k+1}) \) \( (k = 0, 1, \ldots) \) and \( \mathcal{R}_0 \) is a prolongation of \( \mathcal{R}_k \). Then \( F(a_0, a_1, \ldots) \) becomes a difference overfield of \( F \) where the extension of \( a \) (denoted by the same letter) is defined by \( a(b) = \tau_k(b) \) whenever \( b \in F(a_0, \ldots, a_k) \). It is clear that this field coincides with \( F(a^{(1)}) \) and the element \( a^{(1)} \) is a solution of \( A \).

**Proposition 4.4.3.** Let \( \mathcal{R} = (F(a_0, \ldots, a_r), \tau) \) be a difference kernel.

(i) There are only finitely many distinct (that is, pairwise non-isomorphic) generic prolongations of \( \mathcal{R} \).

(ii) Let \( \mathcal{R}' \) be a generic prolongation of \( \mathcal{R} \) and let \( \tilde{a}_0 \) be a subindexing of \( a_0 \) which is algebraically independent over \( F(a_1, \ldots, a_r) \). Then \( \tilde{a}_0 \) is algebraically independent over \( F(a_1, \ldots, a_{r+1}) \).

Let \( \mathcal{R} = (F(a_0, \ldots, a_r), \tau) \) be a difference kernel and let \( \tilde{a}_r \) be a subindexing of \( a_0 \). If \( \tilde{a}_r \) is a transcendence basis of \( a_r \) over \( F(a_0, \ldots, a_{r-1}) \), then \( \tilde{a}_r \) is called a special set. Clearly, such a set consists of \( \delta \mathcal{R} \) elements and it is also a special set for any generic prolongation \( \mathcal{R}' \) of \( \mathcal{R} \). If \( b \) denotes a subindexing \( b_0 \) of \( a_0 \) such that \( b \) contains a special set and \( \text{trdeg}_{F(b_0, \ldots, b_r)} F(b, \ldots, b_r) = \delta \mathcal{R} \), then the order of \( \mathcal{R} \) with respect to \( b \) is defined as \( \text{ord}_b \mathcal{R} = \text{trdeg}_{F(b_0, \ldots, b_r)} F(a_0, \ldots, a_{r-1}) \). (In this case, \( b \) is said to be a subindexing of \( a_0 \) for which \( \text{ord}_b \mathcal{R} \) is defined.) Furthermore, if \( b \) itself is a special set, we define the degree \( \text{ord}_b \mathcal{R} \) and reduced degree \( \text{ord}_b \mathcal{R} \) of \( \mathcal{R} \) with respect to \( b \) to be \( F(a_0, \ldots, a_r) / F(a_0, \ldots, a_{r-1}; b_r) \) and \( [F(a_0, \ldots, a_r) : F(a_0, \ldots, a_{r-1}; b_r)]_{\mathcal{R}} \), respectively. (As usual, if \( L \) is a field extension of a field \( K \), \( [L : K]_{\mathcal{R}} \) denotes the separable factor of the degree of \( L \) over \( K \).)
**Difference algebra**

**Proposition 4.4.4.** With the above notation, let $b$ be a subindexing of $a_0$ for which $\text{ord}_b \mathcal{R}$ is defined.

(i) If $\mathcal{R}'$ is a generic prolongation of $\mathcal{R}$, then $\text{ord}_b \mathcal{R}'$ is defined and $\text{ord}_b \mathcal{R} = \text{ord}_b \mathcal{R}'$.

(ii) Suppose that $b$ itself is a special set. Then $d_b \mathcal{R}$ and $rd_b \mathcal{R}$ are finite. Furthermore, if $\mathcal{R}'_1, \ldots, \mathcal{R}'_h$ are all distinct (pairwise non-isomorphic) prolongations of $\mathcal{R}$, then $\sum_{i=1}^h rd_b \mathcal{R}'_i = rd_b \mathcal{R}$ and $\sum_{i=1}^h d_b \mathcal{R}'_i \leq d_b \mathcal{R}$. In the case of characteristic 0 the last inequality becomes an equality.

Let $F$ be a difference field with a basic set $\sigma$ and let $a = \{a^{(i)} \mid i \in I\}$ be an indexing of elements in a $\sigma$-overfield of $F$. A specialization of $a$ over $F$ is a $\sigma$-homomorphism $\phi$ of $F\{a\}$ into a $\sigma$-overfield of $F$ that leaves $F$ fixed. The image $\phi a = \{\phi a^{(i)} \mid i \in I\}$ is also called a specialization of $a$ over $F$. A specialization $\phi$ is called generic if it is a $\sigma$-homomorphism. Otherwise it is called proper.

Let $F$ be an ordinary difference field with a basic set $\sigma = \{\alpha\}$ and $G = F(\eta_1, \ldots, \eta_s)$ a $\sigma$-overfield of $F$ generated by an $s$-tuple $\eta = (\eta_1, \ldots, \eta_s)$. Then the contraction of $\alpha$ to an isomorphism $\tau_r : F(\eta, \alpha \eta, \ldots, \alpha^{r-1} \eta) \to F(\alpha \eta, \ldots, \alpha^r \eta)$ ($r = 1, 2, \ldots$) defines a difference kernel $\mathcal{R} = (F(a_0, \ldots, a_r), \tau_r)$ of length $r$ over $F$. Conversely, let $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ be a difference kernel with $a_0 = (a_0^{(1)}, \ldots, a_0^{(i)})$. An $s$-tuple $\eta = (\eta_1, \ldots, \eta_s)$ with coordinates from a $\sigma$-overfield of $F$ is called a realization of $\mathcal{R}$ if $\eta, \alpha \eta, \ldots, \alpha^r \eta$ is a specialization of $a_0, \ldots, a_r$ over $F$. If this specialization is generic, the realization is called regular. If there exists a sequence $\mathcal{R}^{(0)} = \mathcal{R}, \mathcal{R}^{(1)}, \mathcal{R}^{(2)}, \ldots$ of kernels, each a generic prolongation of the preceding, such that $\eta$ is a regular realization of each $\mathcal{R}^{(i)}$, then $\eta$ is called a principal realization of $\mathcal{R}$.

The proofs of the following two statements can be found in [32, Chapter 6, Section 6].

**Proposition 4.4.5.** Let $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ be a difference kernel over an ordinary difference field $F$.

(i) There exists a principal realization of the kernel $\mathcal{R}$. If $\eta$ is such a realization, then $\sigma - \text{trdeg}_F F(\eta) = \delta \mathcal{R}$.

(ii) Let $b$ be a subindexing of a such that $\text{ord}_b \mathcal{R}$ is defined and let $\zeta$ is the corresponding subindexing of a principal realization $\eta$ of $\mathcal{R}$. Then $\text{trdeg}_{F(\zeta)} F(\eta) = \text{ord}_b \mathcal{R}$. Furthermore, if $b$ is a special set, then $\zeta$ is a $\sigma$-transcendence basis of $F(\eta)$ over $F$.

(iii) The number of distinct principal realizations of the kernel $\mathcal{R}$ is finite. Let us denote them by $(^{(i)}\eta, \ldots, ^{(i)}\eta)$. If $b$ is a special set and $(^{(i)}\eta$ is the corresponding subset of the components of $(^{(i)}\eta$ ($1 \leq i \leq h$), then $\sum_{i=1}^h rd F^{(i)}(\eta) F^{(i)}(\eta) / F^{(i)}(\eta) \leq d_b \mathcal{R}$ and $\sum_{i=1}^h (rd F^{(i)}(\eta) / F^{(i)}(\eta)) \leq d_b \mathcal{R}$.

(iv) If $\eta$ is a regular realization of the kernel $\mathcal{R}$, but not a principal realization, then $\sigma - \text{trdeg}_F F(\eta) < \delta \mathcal{R}$.

(v) A realization of the kernel $\mathcal{R}$ which specializes over $F$ to a principal realization is a principal realization, and the specialization is generic.

**Proposition 4.4.6.** Let $\eta = (\eta_1, \ldots, \eta_s)$ be an $s$-tuple over an ordinary difference field $F$. Then $\eta$ is the unique principal realization of a kernel $\mathcal{R}$ over $F$ such that every realization of $\mathcal{R}$ is a specialization of $\eta$ over $F$. 


Note that every kernel $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ over an ordinary difference field $F$ with a basic set $\sigma = \{a\}$ is equivalent to a kernel of length 0 or 1 in the sense that its realizations generate precisely the same extensions. Indeed, if $r > 1$, we can define $sr$-tuples $b_0$ and $b_1$ with components of the $s$-tuples $a_0, \ldots, a_{r-1}$ and $s$-tuples $a_1, \ldots, a_r$, respectively. Then the difference kernel $\mathcal{R}^* = (F(b_0, b_1), \tau)$ of length 1 is equivalent to $\mathcal{R}$.

Many concepts and results of this section formulated for the ordinary case can be extended to partial difference fields. In what follows we review such generalizations most of which can be found in [7].

Let $F$ be an inversive difference field with a basic set $\sigma = \{a_1, \ldots, a_n\}$, $\sigma_q = \sigma \setminus \{a_q\}$ for some $a_q \in \sigma$ ($1 \leq q \leq n$), and let $F^q$ denote the $\sigma_q$-field $F$ (that is, the field $F$ treated as a difference field with the basic set $\sigma_q$). A difference kernel $\mathcal{R}$ of length $r$ over $F$ ($r \in \mathbb{N}, r > 0$) is an ordered pair $(F^q \langle a_0, \ldots, a_r \rangle, \tau)$, where $F^q \langle a_0, \ldots, a_r \rangle$ is a $\sigma_q$-overfield of $F^q$ generated by a set of $s$-tuples $a_i = (a_i^{(1)}, \ldots, a_i^{(s)}), 0 \leq i \leq r$ (a positive integer $s$ is fixed), and $\tau$ is a $\sigma_q$-isomorphism of $F^q \langle a_0, \ldots, a_r \rangle$ onto $F^q \langle a_1, \ldots, a_r \rangle$ such that $\tau a_i^{(k)} = a_i^{(k)}$ for $i = 0, \ldots, r-1; k = 1, \ldots, s$ and the restriction of $\tau$ on $F^q$ coincides with $a_q$. (If $r = 0$, then $\tau = a_q$.) A prolongation of $\mathcal{R}$ is a difference kernel $\mathcal{R}'$ consisting of a $\sigma_q$-overfield $F^q \langle a_0, \ldots, a_{r+1} \rangle$ of $F^q \langle a_0, \ldots, a_r \rangle$ together with the extension $\tau'$ of $\tau$ such that $\tau': a_i^{(k)} \mapsto a_i^{(k)} + 1 \leq k \leq s$. (If a kernel is of length 1 or 0, then $F^q \langle a_1, \ldots, a_r \rangle$ and $F^q \langle a_0, \ldots, a_{r-1} \rangle$ respectively are interpreted as $F^q$.) A prolongation $\mathcal{R}'$ of $\mathcal{R}$ is called generic if $F^q \langle a_0, \ldots, a_{r+1} \rangle^*$ is a free join of $F^q \langle a_0, \ldots, a_r \rangle^*$ and $F^q \langle a_1, \ldots, a_{r+1} \rangle^*$ over $F^q \langle a_1, \ldots, a_r \rangle^*$. It follows from Proposition 2.1.4 that if $\mathcal{R}'$ is a prolongation of a kernel $\mathcal{R} = (F^q \langle a_0, \ldots, a_r \rangle, \tau)$ such that $F^q \langle a_0, \ldots, a_r \rangle$ and $F^q \langle a_1, \ldots, a_{r+1} \rangle$ are free over $F^q \langle a_1, \ldots, a_r \rangle$, then $\mathcal{R}'$ is a generic prolongation of $\mathcal{R}$.

**Definition 4.4.7.** We say that a kernel $\mathcal{R}$ satisfies property $\mathcal{P}$ if there exists a $\sigma_q$-overfield $E_1$ of $F^q \langle a_0, \ldots, a_r \rangle$, a $\sigma_q$-subfield $E$ of $E_1$ which contains $F^q \langle a_0, \ldots, a_{r-1} \rangle$, and an extension of $\tau$ to a $\sigma_q$-isomorphism $\tilde{\tau}$ of $E$ into $E_1$ such that

(a) $E_1/E$ is primary (that is, the algebraic closure of $E$ in $E_1$ is purely inseparable over $E$);

(b) $E$ and $F^q \langle a_0, \ldots, a_r \rangle$ are free over $F^q \langle a_0, \ldots, a_{r-1} \rangle$;

(c) if $r = 0$, then $\tilde{\tau} E \subseteq E$. If $r > 0$, then $E_1 = \langle E, \tilde{\tau} E \rangle$ (we use the notation of Proposition 2.1.4).

**Definition 4.4.8.** A kernel $\mathcal{R}$ is said to satisfy property $\mathcal{P}^*$ if there exists an inversive $\sigma_q$-overfield $G_1$ of $F^q \langle a_0, \ldots, a_r \rangle$, a $\sigma_q^*$-subfield $G$ of $G_1$ which contains $F^q \langle a_0, \ldots, a_{r-1} \rangle$, and an extension of $\tau$ to a $\sigma_q$-isomorphism $\tilde{\tau}$ of $G$ into $G_1$ such that

(a) $G_1/G$ is primary;

(b) $G$ and $F^q \langle a_0, \ldots, a_r \rangle^*$ are free over $F^q \langle a_0, \ldots, a_{r-1} \rangle^*$;

(c) if $r = 0$, then $\tilde{\tau} G \subseteq G$. If $r > 0$, then $G_1 = \langle G, \tilde{\tau} G \rangle$.

If the word “free” in (b) is replaced by “quasi-linearly disjoint”, the difference kernel $\mathcal{R}$ is said to satisfy property $\mathcal{L}^*$.

**Proposition 4.4.9.**

(i) The properties $\mathcal{P}$ and $\mathcal{P}^*$ are equivalent.
(ii) With the above notation, a difference kernel \( \mathcal{R} \) satisfies \( \mathcal{L}^* \) if and only if the field extension \( F^q(\sigma_0, \ldots, a_r)^*/F^q(\sigma_0, \ldots, a_{r-1})^* \) is primary.

(iii) If a difference kernel \( \mathcal{R} \) satisfies \( \mathcal{P}^* \) then \( \mathcal{R} \) has a generic prolongation \( \mathcal{R}' \) which satisfies \( \mathcal{P}^* \).

(iv) If a generic prolongation \( \mathcal{R}' \) of a difference kernel \( \mathcal{R} \) satisfies \( \mathcal{P}^* \), then there exists a triple \( (G, G_1, \tilde{\tau}) \) with respect to which \( \mathcal{R} \) satisfies \( \mathcal{P}^* \) and through which a generic prolongation \( \mathcal{R}'' \) of \( \mathcal{R} \) can be obtained such that \( \mathcal{R}'' \) is equivalent to \( \mathcal{R}' \) in sense of isomorphism.

(v) If a difference kernel satisfies \( \mathcal{L}^* \), then all its generic prolongations are equivalent and satisfy \( \mathcal{L}^* \).

Let \( \mathcal{R} = (F^q(\alpha_0, \ldots, a_r), \tau) \) be a difference kernel over an inversive difference field with a basic set \( \sigma = \{\alpha_1, \ldots, \alpha_n\} \) (we use the above notation; in particular, \( a_0, \ldots, a_r \) are \( s \)-tuples and \( \sigma_q = \sigma \setminus \{\alpha_q\} \) with \( 1 \leq q \leq n \)). Let \( \eta \) denote an \( s \)-tuple in a \( \sigma^* \)-overfield \( H \) of \( F \) and let \( \beta \) be the translation of \( H \) which is the extension of \( \alpha_q \). Then with the notation \( \eta_0 = \eta, \eta_j = \beta^j \eta \) \( (j = 1, 2, \ldots) \), we say that \( \eta \) is a realization of \( \mathcal{R} \) in \( H \) over \( F \) if \( (\eta_0, \ldots, \eta_r) \) is a specialization of \( (a_0, \ldots, a_r) \) over \( F^q \) with \( a_j^{(k)} \mapsto \eta_j^{(k)} (1 \leq k \leq s) \). We also say that \( \beta \) is a realization of \( \tau \) in \( H \) over \( F \). If the specialization is generic, \( \eta \) is called a regular realization of \( \mathcal{R} \). If there exists a sequence of kernels \( \mathcal{R}_0 = \mathcal{R}, \mathcal{R}_1, \ldots \) such that for each \( h \in \mathbb{N} \), \( \mathcal{R}_{h+1} \) is a generic prolongation of \( \mathcal{R}_h \) and \( \eta \) is a realized of \( \mathcal{R}_h \) over \( F \), then \( \eta \) is said to be a principal realization of \( \mathcal{R} \).

Two realizations \( \eta \) and \( \zeta \) of a difference kernel \( \mathcal{R} \) are said to be equivalent if the \( \sigma \)-field extensions \( F(\eta)/F \) and \( F(\zeta)/F \) are \( \sigma \)-isomorphic with \( \eta^{(k)} \mapsto \zeta^{(k)} \) \( (1 \leq k \leq s) \).

**Proposition 4.4.10.** The following statements about a difference kernel \( \mathcal{R} \) are equivalent.

(i) \( \mathcal{R} \) satisfies \( \mathcal{P}^* \).

(ii) \( \mathcal{R} \) has a principal realization.

(iii) \( \mathcal{R} \) has a regular realization.

**Proposition 4.4.11.** If a difference kernel \( \mathcal{R} \) over an inversive difference field \( F \) satisfies \( \mathcal{L}^* \), then \( \mathcal{R} \) has a principal realization over \( F \) and all principal realizations of \( \mathcal{R} \) are equivalent.

**Theorem 4.4.12.** If \( \eta \) is a principal realization of a difference kernel \( \mathcal{R} \) over an inversive difference field \( F \), then \( \eta \) is not a proper specialization over \( F \) of any other realization of \( \mathcal{R} \).

One can also prove an analog of Proposition 4.4.6 for partial difference kernels: with the above notation, if \( \eta \) is an \( s \)-tuple over a partial inversive difference field \( F \), there exists a difference kernel \( \mathcal{R} \) such that \( \eta \) is the unique principal realization of \( \mathcal{R} \) over \( F \) and every realization of \( \mathcal{R} \) is a specialization of \( \eta \) over \( F \).

As before, if \( \mathcal{R} = (F^q(\alpha_0, \ldots, a_r), \tau) \) is a difference kernel over an inversive difference \( (\sigma^* \)-field \( F \) (we use the above notation) and \( \tilde{a}_0 \) a subindexing of \( a_0 \), then \( \tilde{a}_j \) will denote the corresponding subindexing of \( a_j \) \( (1 \leq j \leq r) \). Furthermore, we say that an indexing
Let $\mathcal{R}$ be a generic prolongation of a kernel $\mathcal{R} = (F^q \langle a_0, \ldots, a_r \rangle, \tau)$. 

(i) If $\tilde{a}_0$ is a subindexing of $a_0$ such that $\tilde{a}_r$ is $\sigma_q$-algebraically independent over $F^q \langle a_0, \ldots, a_{r-1} \rangle$, then $\tilde{a}_{r+1}$ is $\sigma_q$-algebraically independent over $F^q \langle a_0, \ldots, a_r \rangle$. Furthermore, the set $\bigcup_{i=1}^{r+1} \tilde{a}_i$ is $\sigma_q$-algebraically independent over $F^q$. 

(ii) If $\tilde{a}_0$ is a subindexing of $a_0$ which is $\sigma_q$-algebraically independent over $F^q \langle a_1, \ldots, a_r \rangle$, then $\tilde{a}_0$ is $\sigma_q$-algebraically independent over $F^q \langle a_1, \ldots, a_{r+1} \rangle$ and $\bigcup_{i=1}^{r+1} \tilde{a}_i$ is $\sigma_q$-algebraically independent over $F^q$.

**Theorem 4.4.14.** Let $\eta$ be a principal realization of a kernel $\mathcal{R} = (F^q \langle a_0, \ldots, a_r \rangle, \tau)$.

(i) If $\tilde{a}_0$ is a subindexing of $a_0$ such that $\tilde{a}_r$ is $\sigma_q$-algebraically independent over $F^q \langle a_0, \ldots, a_{r-1} \rangle$, then the corresponding subindexing of $\eta$ is $\sigma_q$-algebraically independent over $F$. 

(ii) If $\tilde{a}_0$ is a subindexing of $a_0$ which is $\sigma_q$-algebraically independent over $F^q \langle a_1, \ldots, a_r \rangle$, then the corresponding subindexing of $\eta$ is $\sigma_q$-algebraically independent over $F$.

### 4.5. Ordinary difference polynomials. Existence theorem

Let $F$ be an ordinary difference field with basic set $\sigma = \{a\}$ and let $R = F \langle y_1, \ldots, y_s \rangle$ be a ring of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$. Throughout this section a $k$-th transform $a^k$ of an element of a $\sigma$-ring will be also denoted by $(k)g$.

Suppose that a $\sigma$-polynomial $A \in R$ contains one or more transforms of $y_i$, $1 \leq i \leq s$ (that is, one or more transforms of $y_i$ appear in the irreducible representation of $A$ as a linear combination of monomials in $y_1, \ldots, y_s$ with coefficients from $F$; we treat $y_i$ as its transform $(0)y_i$). Let $(p)y_i$ and $(q)y_i$ be the transforms of $y_i$ of lowest and highest order, respectively, contained in $A$. Then $q$ and $q - p$ are called the order and effective order of $A$ in $y_i$; they are denoted by ord$_{y_i} A$ and Eord$_{y_i} A$, respectively. If $A$ does not contain transforms of $y_i$, both the order and effective order of $A$ in $y_i$ are defined to be 0.

**Proposition 4.5.1 [32, Chapter 2, Theorem VIII].** Let $F$ be an ordinary difference field with a basic set $\sigma$ and $F[y]$ a ring of $\sigma$-polynomials in one $\sigma$-indeterminate $y$ over $F$. Let an element $\eta$ from some $\sigma$-overfield of $F$ be $\sigma$-algebra over $F$ and let $r \in \mathbb{N}$ be the smallest integer such that $F[y]$ contains a non-zero $\sigma$-polynomial of order $r$ with the solution $\eta$. Then trdeg$_F F(\eta) = r$.

Let $G$ be a difference overfield of an ordinary difference ($\sigma$-) field $F$. In the theory of ordinary difference fields the transcendence degrees trdeg$_F G$ and trdeg$_{F^*} G^*$ are also called the order and effective order of the $\sigma$-field extension $G/F$; they are denoted by ord $G/F$ and Eord $G/F$, respectively. (As usual, $K^*$ denotes the inversive closure of a difference field $K$.) Clearly, Eord $G/F \leq$ ord $G/F$, Eord $G/F =$ ord $G/F$ if $F$ is inversive, and ord $G/F = \infty$ if $\sigma$-trdeg$_F G > 0$. Furthermore, the properties of transcendence
degree imply that for any chain $F \subseteq G \subseteq H$ of ordinary difference field extensions we have $\text{ord} H/F = \text{ord} H/G + \text{ord} G/F$ and $\text{Eord} H/F = \text{Eord} H/G + \text{Eord} G/F$.

The proofs of the statements in the rest of this section can be found in [32, Chapters 6–10].

**Proposition 4.5.2.** Let $F$ be an ordinary difference $(\sigma,\cdot)$-field and $\phi$ a specialization of an $s$-tuple $a = (a^{(1)}, \ldots, a^{(s)})$ over $F$. Let $\{i_1, \ldots, i_q\}$ be a subset of $\{1, \ldots, s\}$, $\check{a}$ denote the set $\{a^{(i_1)}, \ldots, a^{(i_q)}\}$, and $\phi \check{a} = \{\phi a^{(i_1)}, \ldots, \phi a^{(i_q)}\}$.

(i) If $\phi \check{a}$ is $\sigma$-algebraically independent over $F$, so is $\check{a}$. Thus, $\sigma$-trdeg$_F F(\phi a^{(1)}, \ldots, \phi a^{(s)}) \leq \sigma$-trdeg$_F F(a^{(1)}, \ldots, a^{(s)})$.

(ii) If $\phi \check{a}$ is a transcendence basis of $\{a^{(1)}, \ldots, a^{(s)}\}$ over $F$, then $\text{ord} F(\phi a^{(1)}, \ldots, \phi a^{(s)})/\text{ord} F(\check{a}) \leq \text{ord} F(a^{(1)}, \ldots, a^{(s)})/\text{ord} F(\check{a})$, $\text{Eord} F(\phi a^{(1)}, \ldots, \phi a^{(s)})/\text{Eord} F(\check{a}) \leq \text{Eord} F(a^{(1)}, \ldots, a^{(s)})/\text{Eord} F(\check{a})$, and the equality occurs if and only if the specialization $\phi$ is generic.

**Proposition 4.5.3.** Let $F$ be an ordinary difference field with a basic set $\sigma$, $F[y]$ a ring of $\sigma$-polynomials in one $\sigma$-indeterminate $y$ over $F$, and $\eta$ an element of some $\sigma$-overfield of $F$ which is $\sigma$-algebraic over $F$. Let $r \in \mathbb{N}$ be the smallest non-negative integer such that $F[y]$ contains a non-zero $\sigma$-polynomial of effective order $r$ with the solution $\eta$. Then $\text{Eord} F(\eta)/r = r$.

Let $F$ be an ordinary difference field with a basic set $\sigma$, $F[y_1, \ldots, y_k]$ a ring of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_k$ over $F$, and $\mathcal{M}$ a non-empty irreducible variety over $F[y_1, \ldots, y_k]$. If $(\eta_1, \ldots, \eta_k)$ is a generic zero of $\mathcal{M}$, then $\sigma$-trdeg$_F F(\eta_1, \ldots, \eta_k)$, $\text{ord} F(\eta_1, \ldots, \eta_k)/F$, $\text{Eord} F(\eta_1, \ldots, \eta_k)/F$, and $\text{id} F(\eta_1, \ldots, \eta_k)/F$ are called the dimension, order, effective order, and limit degree of the variety $\mathcal{M}$, respectively. They are denoted by $\text{dim} \mathcal{M}$, $\text{ord} \mathcal{M}$, $\text{Eord} \mathcal{M}$, and $\text{id} \mathcal{M}$, respectively. If $\mathcal{M} = \emptyset$, we set $\text{dim} \mathcal{M} = \text{ord} \mathcal{M} = \text{Eord} \mathcal{M} = -1$ and $\text{id} \mathcal{M} = 0$. Obviously, if $\text{dim} \mathcal{M} > 0$, then $\text{ord} \mathcal{M} = \infty$.

If $P$ is a prime inverse difference ideal of $F[y_1, \ldots, y_k]$ then the dimension, order, effective order, and limit degree of $P$ (they are denoted by $\text{dim} P$, $\text{ord} P$, $\text{Eord} P$, and $\text{id} P$, respectively) are defined as the corresponding values of the variety $\mathcal{M}(P)$. (Thus, these concepts are determined as above through the $\sigma$-field extension $F(\eta_1, \ldots, \eta_k)/F$ where $(\eta_1, \ldots, \eta_k)$ is a generic zero of $P$.)

Let $\mathcal{M}$ be a non-empty irreducible variety over $F[y_1, \ldots, y_k]$, $(\eta_1, \ldots, \eta_k)$ a generic zero of $\mathcal{M}$, and $\{y_{i_1}, \ldots, y_{i_q}\}$ is a subset of the set of $\sigma$-indeterminates $\{y_1, \ldots, y_k\}$. Then the dimension, order, effective order, and limit degree of $\mathcal{M}$ relative to $y_{i_1}, \ldots, y_{i_q}$ are defined as $\sigma$-trdeg$_F F(\eta_{i_1}, \ldots, \eta_{i_q})/F(\eta_1, \ldots, \eta_k)$, $\text{ord} F(\eta_{i_1}, \ldots, \eta_{i_q})/F(\eta_1, \ldots, \eta_k)$, $\text{Eord} F(\eta_{i_1}, \ldots, \eta_{i_q})/F(\eta_1, \ldots, \eta_k)$, and $\text{id} F(\eta_{i_1}, \ldots, \eta_{i_q})/F(\eta_1, \ldots, \eta_k)$, respectively. These characteristics of $\mathcal{M}$ are denoted by $\text{dim}(y_{i_1}, \ldots, y_{i_q})\mathcal{M}$, $\text{ord}(y_{i_1}, \ldots, y_{i_q})\mathcal{M}$, $\text{Eord}(y_{i_1}, \ldots, y_{i_q})\mathcal{M}$, and $\text{id}(y_{i_1}, \ldots, y_{i_q})\mathcal{M}$, respectively. If $P$ is a prime inverse difference ideal of $F[y_1, \ldots, y_k]$ then the dimension, order, effective order, and limit degree of $P$ relative to $y_{i_1}, \ldots, y_{i_q}$ are defined as the corresponding characteristics of $\mathcal{M}(P)$ (the notation is the same: $\text{dim}(y_{i_1}, \ldots, y_{i_q})P$, $\text{ord}(y_{i_1}, \ldots, y_{i_q})P$, $\text{Eord}(y_{i_1}, \ldots, y_{i_q})(P)$, and $\text{id}(y_{i_1}, \ldots, y_{i_q})(P)$, respectively).
A subset $\{y_1, \ldots, y_s\}$ of $\{y_1, \ldots, y_s\}$ is called a set of parameters of $P$ (or $\mathcal{M}(P)$) if $P$ contains no non-zero $\sigma$-polynomial in $F[y_1, \ldots, y_s]$. A set of parameters of $P$ which is not a proper subset of any set of parameters of $P$ is called complete. Clearly, every reflexive prime $\sigma$-ideal of $F[y_1, \ldots, y_s]$ has at least one complete set of parameters, and every set of parameters can be extended to a complete one.

**PROPOSITION 4.5.4.** With the above notation, a set of parameters of a proper inversive difference ideal $P$ of $F[y_1, \ldots, y_s]$ is complete if and only if it contains $\dim P$ elements.

**PROPOSITION 4.5.5.** Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two irreducible varieties over $F[y_1, \ldots, y_s]$ such that $\mathcal{M}_1 \subseteq \mathcal{M}_2$. Then

(i) $\dim \mathcal{M}_1 \leq \dim \mathcal{M}_2$.

(ii) If $\{y_{i_1}, \ldots, y_{i_q}\}$ is a complete set of parameters of $\Phi(\mathcal{M}_1)$ (we use the notation of section 2.5), then $\text{ord}(y_{i_1}, \ldots, y_{i_q}) \mathcal{M}_1 \leq \text{ord}(y_{i_1}, \ldots, y_{i_q}) \mathcal{M}_2$, and the equality occurs if and only if $\mathcal{M}_1 = \mathcal{M}_2$. Furthermore, $\text{Eord}(y_{i_1}, \ldots, y_{i_q}) \mathcal{M}_1 \subseteq \mathcal{M}_2$, and the equality occurs if and only if $\mathcal{M}_1 = \mathcal{M}_2$.

Let $F$ be an ordinary difference field with a basic set $\sigma = \{\alpha\}$ and $R = F[y_1, \ldots, y_s]$ a ring of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$. Let $A \subseteq R$ be an algebraically irreducible $\sigma$-polynomial, that is, $A \notin F$ and $A$ is irreducible as a polynomial in the indeterminates $(^{(j)}y_i)$ $(1 \leq i \leq s; j = 0, 1, \ldots)$ (in this case $R$ is treated as a polynomial ring in this denumerable set of indeterminates over $F$; as above, $(^{(j)}y_i)$ denotes $\alpha^j y_i$). An irreducible component $\mathcal{M}$ of $\mathcal{M}(A)$ is called a principal component of the variety $\mathcal{M}(A)$ if whenever $A$ contains a transform of $y_i$ $(1 \leq i \leq s)$, the family $\{y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s\}$ is a complete set of parameters of $\mathcal{M}$ and $\text{Eord}(y_{i-1}, y_{i+1}, \ldots, y_s) \mathcal{M}$ is the effective order of $A$ in $y_i$. (This implies that $\dim \mathcal{M} = s - 1$.) Irreducible components of $\mathcal{M}$ which are not principal are called singular.

The following proposition shows that no component of $\mathcal{M}(A)$ is “larger” than a principal component.

**PROPOSITION 4.5.6.**

(i) With the above notation, let $\eta = (\eta_1, \ldots, \eta_s)$ be a solution of the non-zero $\sigma$-polynomial $A \in R$. Then either $\sigma \text{-trdeg}_F F(\eta_1, \ldots, \eta_s) < s - 1$ or $\sigma \text{-trdeg}_F F(\eta_1, \ldots, \eta_s) = s - 1$ and for each $k$ $(1 \leq k \leq s)$, such that the set $\tilde{\eta}_k = \{\eta_1, \ldots, \eta_{k-1}, \eta_{k+1}, \ldots, \eta_s\}$ is $\sigma$-algebraically independent over $F$, $A$ contains a transform of $y_k$ and $\text{ord} F(\eta_1, \ldots, \eta_s(s))/F(\tilde{\eta}_k) \leq \text{ord}_{y_k} A$, $\text{Eord} F(\eta_1, \ldots, \eta_s(s))/F(\tilde{\eta}_k) \leq \text{Eord}_{y_k} A$.

(ii) Let a $\sigma$-polynomial $A \in R$ be algebraically irreducible, $\sigma \text{-trdeg}_F F(\eta_1, \ldots, \eta_s) = s - 1$, and for each $k$ $(1 \leq k \leq s)$ such that $A$ contains a transform of $y_k$, $\text{Eord} F(\eta_1, \ldots, \eta_s(s))/F(\tilde{\eta}_k) = \text{Eord}_{y_k} A$. Then $\eta$ is a generic zero of a principal component of $\mathcal{M}(A)$.

The following fundamental result is an abstract form of an existence theorem for ordinary algebraic difference equations.
Theorem 4.5.7. Let $F$ be an ordinary difference field with a basic set $\sigma$, $R = F\{y_1, \ldots, y_s\}$ a ring of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$, and $A \in R$ an algebraically irreducible $\sigma$-polynomial. Then

(i) The variety $M(A)$ has principal components.

(ii) Let $A$ contain a transform of some $y_i$ ($1 \leq i \leq s$). Then

(a) If $A$ contains a transform of order 0 of some $\sigma$-indeterminate $y_j$ and $M$ is a principal component of $M(A)$, then $\text{ord}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s)M = \text{ord}_{y_i} A$.

(b) If $M_1, \ldots, M_k$ are the principal components of $M(A)$ and $d$ and $e$ denote, respectively, the degree and reduced degree of $A$ in the highest transform of $y_i$, which it contains, then $\sum_{j=1}^k \text{id}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s)M_j \leq d$ and $\sum_{j=1}^k \text{irld}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s)M_j = e$.

(c) If $d_1$ and $e_1$ denote, respectively, the degree and reduced degree of $A$ in the lowest transform of $y_i$ contained in $A$, then $\sum_{j=1}^k \text{id}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s)M_j \leq d_1$ and $\sum_{j=1}^k \text{irld}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s)M_j = e_1$.

(d) Let $q = \text{Eord}_y A$ and let $M$ be a component of $M(A)$ such that $\{y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s\}$ is a complete set of parameters of $M$ and $\text{Eord}(y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s)M = q$. Then is $M$ a principal component of $M(A)$.

Corollary 4.5.8. Every ordinary difference field $F$ has an algebraic closure $G$ which is a difference overfield of $F$.

Proposition 4.5.9. Let $F$ be an ordinary difference field with a basic set $\sigma$, $F\{y\}$ a ring of $\sigma$-polynomials in one $\sigma$-indeterminate $y$ over $F$, and $A$ an algebraically irreducible $\sigma$-polynomial of order 0. Then, if any component of $M(A)$ has limit degree 1, so do all components. Similar statements hold for $\text{rld}$, $\text{id}$, and $\text{irld}$.

Theorem 4.5.10. Let $F$ be an ordinary difference field with a basic set $\sigma$ and $R = F\{y_1, \ldots, y_s\}$ a ring of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$.

(i) If $M$ is an irreducible variety over $F\{y_1, \ldots, y_s\}$, then $\dim M = s - 1$ if and only if $M$ is a principal component of the variety of an algebraically irreducible $\sigma$-polynomial $A$.

(ii) Let $A \in R \setminus F$ and let $M$ be an irreducible component of the variety $M(A)$ of the $\sigma$-polynomial $A$. Then $\dim M = s - 1$.

(iii) Let $A, B \in R$ be $\sigma$-polynomials of order at most 1 in each $y_i$ ($1 \leq i \leq s$). If $M(A, B)$ is not empty, it has a component of dimension not less than $s - 2$.

R.M. Cohn [32, Chapter 10] conjectured that if $\eta$ is a realization of a kernel $\mathcal{R}$ over an ordinary difference field $F$, then $\eta$ is a specialization over $F$ of a realization $\xi$ such that $\sigma\text{-trdeg}_F F(\xi) \geq \delta\mathcal{R}$. If this statement is true, it would strengthen the conclusion of Theorem 4.5.10(iii) to the statement that every component of $M(A, B)$ has dimension at least $s - 2$.

Let $F$ be an ordinary difference field with a basic set $\sigma = \{\alpha\}$ and let $R = F\{y_1, \ldots, y_s\}$ be a ring of $\sigma$-polynomials in a set of $\sigma$-indeterminates $Y = \{y_1, \ldots, y_s\}$ over $F$. Let $\Phi \subseteq
with the kernel is the specialization of a principal realization of $\Phi$ in each $y_i \in Y \setminus Y'$ such that $y_i \in Y \setminus Y'$. For every $y_i \in Y \setminus Y'$, let $r_i$ denote the maximum of the orders of the $\sigma$-polynomials of $\Phi$ in $y_i$. Then the number $\mathcal{R}(Y')\Phi = \sum_i r_i$ (the summation extends over all values of the index $i$ such that $y_i \in Y \setminus Y'$) is called the Ritt number of the system $\Phi$ associated with the set $Y' \subseteq Y$. If $Y' = \emptyset$, we set $\mathcal{R}(Y')\Phi = \sum_{i=1}^s r_i$.

Let $0 \neq A \in \Phi$ and let $h(A)$ be the greatest non-negative integer such that for each $i$ with $y_i \in Y \setminus Y'$, $\hat{a}^{h(A)}(A)$ is of order at most $r_i$ in $y_i$. Let $h = \max \{h(A) \mid 0 \neq A \in \Phi\}$, that is, $h$ is the greatest integer such that some polynomial in $\Phi$ may be replaced by its $h$-th transform without altering the Ritt number of $\Phi$. The number $\mathcal{G}(Y')\Phi = \mathcal{R}(Y')\Phi - h$ is called the Greenspan number of the system $\Phi$ associated with the set $Y' \subseteq Y$. If $Y' = \emptyset$, we write $\mathcal{G}\Phi$ for $\mathcal{G}(Y')\Phi$.

Now, let the system of $\sigma$-polynomials $\Phi$ be finite, $\Phi = \{A_1, \ldots, A_m\}$, and let $r_{ij} = \text{ord}_{y_i} A_j \ (1 \leq i \leq m, 1 \leq j \leq s)$. Then the number $\mathcal{J}(\Phi) = \max \{\sum_{i=1}^s r_{ij} \mid (j_1, \ldots, j_s)\}$ is a permutation of $1, \ldots, s$ is called the Jacobi number of the system $\Phi$. The following theorem (where we use the above notation) gives some bounds on the effective orders with respect to $Y'$ that involve the Ritt, Greenspan and Jacobi numbers.

**Theorem 4.5.11.**

(i) Suppose that the Ritt number $\mathcal{R}(Y')\Phi$ for a system of $\sigma$-polynomials $\Phi$ (and a set $Y' \subseteq Y$) is defined. If $\mathcal{M}$ is a component of $\mathcal{M}(\Phi)$ for which $Y'$ contains a complete set of parameters, then $\text{Eord}(Y')\mathcal{M} \leq \mathcal{R}(Y')\Phi$.

(ii) If $\mathcal{G}\Phi$ is defined, $\mathcal{M}$ is a component of $\mathcal{M}(\Phi)$ and $Y'$ contains a complete set of parameters of every component of $\mathcal{M}(\Phi)$, then $\text{Eord}(Y')\mathcal{M} \leq \mathcal{G}(Y')\Phi$.

(iii) Let $\Phi = \{A_1, \ldots, A_m\}$ where $A_1, \ldots, A_m$ are first-order $\sigma$-polynomials (that is $r_{ij} \leq 1$ for $i = 1, \ldots, m; j = 1, \ldots, s$ and at least one $r_{ij}$ is equal to 1). If $\mathcal{M}$ is an irreducible component of $\mathcal{M}(\Phi)$ of dimension 0, then $\text{Eord}\mathcal{M} \leq \mathcal{J}(\Phi)$.

A number of results on the Jacobi bound for systems of algebraic differential equations (see, for example, [41,42] and [88, Section 5.8]) give a hope that the last theorem can be essentially strengthened and generalized to the case of partial difference polynomials.

In [32, Chapter 10, Example 2] R.M. Cohn showed that it is possible for a difference kernel $\mathcal{R}$ to have a realization $\eta$ such that $\sigma$-trdeg$_F F(\eta) > \delta \mathcal{R}$. His example also shows that if $\mathcal{R}'$ is a kernel over an ordinary inversive difference field $F$, $\xi$ is a principal realization of $\mathcal{R}'$, and $\mathcal{R}$ is a kernel over $F$ which specializes to $\mathcal{R}'$, there may be no principal realization of $\mathcal{R}$ which specializes to $\xi$. The following theorem presents conditions that imply the existence of such a realization. It also implies that every regular realization of a kernel is the specialization of a principal realization.

**Theorem 4.5.12** [96]. Let $F$ be an ordinary inversive difference field. Let $\mathcal{R} = (F(a_0, \ldots, a_r), \tau)$ and $\mathcal{R}' = (F(a'_0, \ldots, a'_r), \tau) \ (r \in \mathbb{N})$ be two difference kernels over $F$ such that there is an $F$-isomorphism of $F(a_0, \ldots, a_{r-1})$ onto $F(a'_0, \ldots, a'_{r-1})$. Let $\xi$ be a principal realization of $\mathcal{R}'$. If $\mathcal{R}$ specializes to $\mathcal{R}'$, then there exists a principal realization $\eta$ of $\mathcal{R}$ which specializes to $\xi$. 


In the rest of this section we use the notation and conventions of Section 4.4. Let \( F \) be an ordinary inversive difference field with a basic set \( \sigma = \{ \alpha \} \) and \( R = (F(a_0, \ldots, a_r), \tau) \) a difference kernel \( (a_0 = (a_0^{(1)}, \ldots, a_0^{(s)})) \). A realization \( \eta = (\eta_1, \ldots, \eta_s) \) of the kernel \( R \) is called singular if it is not a specialization of a principal realization of \( R \). A realization of \( R \) is called multiple if it is a specialization of two principal realizations which are distinct in the sense of isomorphism. The following two examples (see [32, Chapter 6, Section 21]) illustrate these concepts.

**Example 4.5.13.** Let us consider the algebraically irreducible \( \sigma \)-polynomial \( A = (2)yy + (1)y \) in the ring of \( \sigma \)-polynomials \( F[y] \) in one difference indeterminate \( y \) (as before, \( k \) stands for \( \alpha^k y \)). Then \( y\alpha(A) - A = (1)(3)yy - 1 \). If \( \eta \neq 0 \) is a generic zero of an irreducible component \( M \) of \( M(A) \), then \( \eta \) must annul \( (3)yy - 1 \), whence \( (3)yy - 1 \in \Phi(M), 0 \notin M \). Thus, the solution 0 of \( A \) itself constitutes an irreducible component of \( M(A) \). Clearly, it is a singular component and furnishes a singular realization of the kernel produced for \( A \) by the procedure described in Example 4.4.2. It should be noted (see [19, Theorem V]) that the variety of a first-order \( \sigma \)-polynomial cannot have a singular component.

**Example 4.5.14.** With the notation of the previous example, let \( B = (1)y^2 + y^2 \in F[y] \) where \( F = \mathbb{Q} \) (and \( \alpha = \text{id}_\mathbb{Q} \)). Then \( \alpha(B) - B = (2)y - y(2)y - y \). It can be shown (see [32, Chapter 6, Theorem IV]) that \( M(B) \) has two principal components \( M_1 \) and \( M_2 \) which annul \( (2)y - y \) and \( (2)y - y \), respectively. Furthermore, 0 \( \in M_1 \cap M_2 \), so 0 is a multiple solution of the kernel \( R \) formed for \( B \) by the procedure described in Example 4.4.2.

**Proposition 4.5.15.** Let \( R = (F(a_0, \ldots, a_r), \tau) \) be a difference kernel over an ordinary inversive difference (\( \sigma \)-) field \( F \) of zero characteristic. Let \( S \) denote the polynomial ring in \( s(r + 1) \) indeterminates \( F[[y_i] \mid 1 \leq i \leq s, 0 \leq j \leq r] \subseteq F[y_1, \ldots, y_s] \), and let \( P \) be the prime ideal of \( S \) with the general zero \( (a_0, \ldots, a_r) \). Then:

(i) There exist \( \sigma \)-polynomials \( A, B \in S \setminus P \) such that

(a) Every singular realization of \( R \) annuls \( A \).

(b) Every multiple realization of \( R \) annuls \( B \).

(ii) If \( r = 0 \), then \( R \) has no singular realization (even if \( \text{Char } F \neq 0 \)).

Let \( F \) be an ordinary difference (\( \sigma \)-) field and \( R = F[y_1, \ldots, y_s] \) a ring of \( \sigma \)-polynomials over \( F \). Suppose that a \( \sigma \)-polynomial \( A \in R \) contains one or more transforms of a \( \sigma \)-indeterminate \( y_i \) \( (1 \leq i \leq s) \) and let \( (p)y_i \) and \( (q)y_i \) be such transforms of lowest and highest orders, respectively. Then the formal partial derivatives \( \partial A/\partial^{(p)}y_i \) and \( \partial A/\partial^{(q)}y_i \) are called the separants of \( A \) with respect to \( y_i \). If \( A \) is written as a polynomial in \( (q)y_i \), then the coefficient of the highest power of \( (q)y_i \) is called the initial of \( A \) with respect to \( y_i \).

**Proposition 4.5.16.** With the above notation, let \( F \) be a \( \sigma \)-field of zero characteristic and \( A \in R \) an algebraically irreducible \( \sigma \)-polynomial. Then

(i) Every singular component of \( M(A) \) and every solution common to two principal components of \( M(A) \) annuls the separants of \( A \).
More precisely, we have the following statement.

PROPOSITION 4.5.17. Let \( F \) be an ordinary difference field and \( P \) a reflexive prime difference ideal in a ring of difference polynomials \( F[y_1, \ldots, y_s] \). Unless \( \dim P = 0 \), \( \text{Eord} \ P = 0 \), there exist difference overfields of \( F \) containing arbitrarily many generic zeros of \( P \).

An element \( a \) of an ordinary difference ring \( R \) with a basic set \( \sigma = \{ \alpha \} \) is called periodic if \( \alpha^k(a) = a \) for some \( k \in \mathbb{N} \). (In particular, every constant \( c \in R \) is periodic.) Clearly, the set of all periodic elements of \( R \) is a \( \sigma \)-subring of \( R \); it is called the \( \sigma \)-subring of periodic elements of \( R \).

A difference ring \( R \) with a basic set \( \sigma = \{ \alpha \} \) is said to be periodic if there exists \( m \in \mathbb{N} \) such that \( \alpha^m(a) = a \) for all \( a \in R \). In particular, if \( m = 1 \), the \( \sigma \)-ring \( R \) is said to be invariant.

PROPOSITION 4.5.18. Let \( F \) be an ordinary difference field with a basic set \( \sigma \) and let \( F' \) and \( F'' \) be its \( \sigma \)-subfields of constants and periodic elements, respectively. Then \( F'' \) is the algebraic closure of \( F' \) in \( F \).

An ordinary difference field with a basic set \( \sigma = \{ \alpha \} \) is called completely aperiodic if either \( \text{Char} \ F = 0 \) or \( \text{Char} \ F = p > 0 \) and for any \( i, j, k, l \in \mathbb{N} \), \( i \neq j, k \neq l \), no element of \( F \) satisfies the equation \( (\alpha^i(y))^{pl} = (\alpha^j(y))^{pl} \). (It is easy to see that if an ordinary difference field \( F \) of positive characteristic is aperiodic and contains infinitely many constants, then \( F \) is completely aperiodic.)

THEOREM 4.5.19 [25]. Let \( F \) be a completely aperiodic ordinary difference \((\sigma \cdot \cdot)\) field, \( G \) a \( \sigma \)-overfield of \( F \), \( R = G[y_1, \ldots, y_s] \) a ring of \( \sigma \)-polynomials in \( \sigma \)-indeterminates \( y_1, \ldots, y_s \) over \( G \), and \( 0 \neq A \in R \). Then

(i) There exists an \( s \)-tuple \( (\eta_1, \ldots, \eta_s) \) of elements of \( F \) which is not a solution of \( A \).

(ii) Suppose that \( G/F \) is a finitely generated \( \sigma \)-field extension, \( \sigma \text{-trdeg}_F G = 0 \), and either \( \text{Char} \ F = 0 \) or \( \text{Char} \ F > 0 \) and \( \text{rlc} G.F = \text{ld} G/F \). Then there exist \( \zeta \in G \), \( k \in \mathbb{N} \) such that \( \alpha^k(a) \in F(\zeta) \) for every element \( a \in G \). Moreover, such an element \( \zeta \) can be chosen as a linear combination of the members of any finite set of \( \sigma \)-generators of \( G/F \) with coefficients from any pre-assigned completely aperiodic \( \sigma \)-subfield of \( F \).

Using the last theorem, R.M. Cohn [32, Chapter 8] showed that the solutions of any irreducible variety \( \mathcal{M} \) over an ordinary difference field \( F \) can be obtained by rational operations, transforming, and the inverse of transforming from the solutions of a principal component \( \mathcal{N} \) of the variety of an algebraically irreducible difference polynomial \( \mathcal{N} \) is the variety over \( F \), but not necessarily over the same ring of difference polynomials, as \( \mathcal{M} \).

More precisely, we have the following statement.

THEOREM 4.5.20. Let \( F \) be an ordinary difference field with a basic set \( \sigma = \{ \alpha \} \), \( F[y_1, \ldots, y_s] \) a ring of \( \sigma \)-polynomials in \( \sigma \)-indeterminates \( y_1, \ldots, y_s \) over \( F \), and \( \mathcal{M} \)
an irreducible variety over $F\{y_1, \ldots, y_s\}$. Suppose that $F$ is completely aperiodic or that $\dim \mathcal{M} > 0$, and also that $\mathcal{M}$ possesses a complete set of parameters $\{y_1, \ldots, y_k\}$ such that $\text{rld}(y_1, \ldots, y_k)\mathcal{M} = \text{ld}(y_1, \ldots, y_k)\mathcal{M}$. (We assume that the $\sigma$-indeterminates are so numbered that the first $k$ of them constitute the complete set of parameters. Note also that the last equality always holds if $\text{Char} F = 0$.) Then there exist:

(a) an irreducible variety $\mathcal{N}$ over the ring of $\sigma$-polynomials $F\{y_1, \ldots, y_k; w\}$ ($w$ is the $(k + 1)$-th $\sigma$-indeterminate of this ring);

(b) $\sigma$-polynomials $A_{k+1}, \ldots, A_s \in F\{y_1, \ldots, y_k\}$, $\sigma$-polynomials $B_{k+1}, \ldots, B_s, C \in F\{y_1, \ldots, y_k; w\}$, $C \notin \mathcal{N}$, and an integer $\tau \geq 0$ such that

(i) $\mathcal{N}$ is a principal component of the variety of an algebraically irreducible $\sigma$-polynomial of $F\{y_1, \ldots, y_k; w\}$, $y_1, \ldots, y_k$ constitute a complete set of parameters of $\mathcal{N}$, and $\text{Eord}(y_1, \ldots, y_k)\mathcal{N} = \text{Eord}(y_1, \ldots, y_k)\mathcal{M}$.

(ii) If $(\eta_1, \ldots, \eta_s)$ is any solution in $\mathcal{M}$, there is a solution in $\mathcal{N}$ with $y_i = \eta_i$ for $i = 1, \ldots, k$, and $w$ given by the result of substituting $\eta_j$ for $y_j$ in $\sum_{i=k+1}^s A_i y_i$.

(iii) If $y_i = \zeta_i, i = 1, \ldots, k$, $w = 0$ is a solution in $\mathcal{N}$ which does not annul $C$, then there is a solution in $\mathcal{M}$ with $y_i = \zeta_i, i = 1, \ldots, k$, and $y_j, k+1 \leq j \leq s$, given by applying $\sigma^{-1}$ to the result of substituting $\zeta_{k+1}, \ldots, \zeta_s, \theta$ for $y_{k+1}, \ldots, y_s, w$, respectively, in $B_j/C$.

(iv) If $D \in F\{y_1, \ldots, y_k\} \setminus \Phi(\mathcal{M})$, then there exists a $\sigma$-polynomial $E \in F\{y_1, \ldots, y_k; w\} \setminus \Phi(\mathcal{N})$ such that any solution in $\mathcal{N}$ not annulling $E$ gives rise by a procedure described in (iii) to a solution in $\mathcal{M}$ not annulling $D$.

(v) The procedures of (ii) and (iii) carry generic zeros of $\mathcal{M}$ or $\mathcal{N}$ into generic zeros of $\mathcal{N}$ or $\mathcal{M}$. Whenever (iii) is defined, these procedures, applied to elements of $\mathcal{M}$ or $\mathcal{N}$, are inverses of each other.

With the notation of the last theorem, $W = \sum_{i=k+1}^s A_i y_i$ is called a resolvent for $\mathcal{M}$ or for $\Phi(\mathcal{M})$, and $\Phi(\mathcal{N})$ is called a resolvent ideal for $\mathcal{M}$ or for $\Phi(\mathcal{M})$. One can say that $\mathcal{M}$ is obtained from the solutions of its resolvent ideal by the relations $\sigma^t y_i = B_j/C$ $(k + 1 \leq i \leq s)$ and the solutions of the resolvent ideal are obtained from those of $\mathcal{M}$ by the relation $w = W$.

We conclude this section with a summary of the basic properties of a variety of one ordinary difference polynomial (see [32, Chapters 6 and 10]).

Let $F$ be an ordinary difference ($\sigma$-) field (we assume $\text{Char} F = 0$, except in (i) and (iii)), $R = F\{y_1, \ldots, y_s\}$ a ring of $\sigma$-polynomials over $F$ and $A \in R \setminus F$ an algebraically irreducible $\sigma$-polynomial. Then

(i) The variety $\mathcal{M}(A)$ consists of one or more principal components and (possibly) of singular components. Each component of $\mathcal{M}(A)$ has dimension $s - 1$.

(ii) The relative effective orders and (with certain limitations) the relative orders of principal components of $\mathcal{M}(A)$ are determined by the effective orders and orders of $A$. Furthermore, it is sufficient for a component to have dimension $s - 1$ and one of these effective orders for it to be a principal component.

(iii) Except in the case that $s = 1$ and $A$ is of effective order 0, every principal component of $\mathcal{M}(A)$ contains infinitely many generic zeros.

(iv) The relative effective orders of the singular components of $\mathcal{M}(A)$ are less by at least 2 than those of the principal components. More precisely, if $\text{Eord}_y A = r_i$
(1 ≤ i ≤ s) and \( \mathcal{M} \) is a singular component of \( \mathcal{M}(A) \), then either \( y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_s \) do not constitute a set of parameters of \( \mathcal{M} \) or they do constitute such a set and \( \text{Eord}(y_1, \ldots, y_{i-1}, y_s, y_{i+1}, \ldots, y_s) \mathcal{M} \subseteq r_i - 2 \).

(v) Any singular component of \( \mathcal{M}(A) \) is itself a principal component of \( \mathcal{M}(B) \) for some algebraically irreducible \( \sigma \)-polynomial \( B \in R \). It follows from (iv), for each \( i, 1 \leq i \leq s \), either \( B \) contains no transforms of \( y_i \) or \( \text{Eord} y_i B \leq r_i - 2 \).

(vi) If \( s = 1 \) and \( A \) is a first-order \( \sigma \)-polynomial in \( R \), then \( \mathcal{M}(A) \) has no singular components.

A number of additional important results on varieties of ordinary difference polynomials can be found in \([133,19,20,22,23,26,28,33,34]\), and \([32\text{, Chapter 10}\).]

### 4.6. Compatibility of difference field extensions. Specializations

Let \( F \) be a difference field with a basic set \( \sigma \) and let \( G \) and \( H \) be two \( \sigma \)-overfields of \( F \). The difference field extensions \( G/F \) and \( H/F \) are said to be compatible if there exists a \( \sigma \)-field extension \( E \) of \( F \) such that \( G/F \) and \( H/F \) have \( \sigma \)-isomorphisms into \( E/F \). (As usual, this means that there exist \( \sigma \)-isomorphisms of \( G \) and \( H \) into \( E \) that leave the \( \sigma \)-field \( F \) fixed.) Otherwise, the \( \sigma \)-field extensions \( G/F \) and \( H/F \) are called incompatible.

**Example 4.6.1** \([32\text{, Chapter 1, Example 4}\). Let us consider \( \mathbb{Q} \) as an ordinary difference field whose basic set \( \sigma \) consists of the identity automorphism \( \alpha \). If one adjoins to \( \mathbb{Q} \) an element \( i \) such that \( i^2 = -1 \), then the resulting field \( \mathbb{Q}(i) \) has two automorphisms that extend \( \alpha \): one of them is the identity mapping (we denote it by the same letter \( \alpha \)) and the other (denoted by \( \beta \)) sends an element \( a + bi \in \mathbb{Q}(i) \) (\( a, b \in \mathbb{Q} \)) to \( a - bi \) (complex conjugation). Then \( \mathbb{Q}(i) \) can be treated as a difference field with the basic set \( \{\alpha\} \), as well as a difference field with the basic set \( \{\beta\} \). Denoting these two difference fields by \( G \) and \( H \), respectively, we can naturally consider them as \( \sigma \)-field extensions of \( \mathbb{Q} \). Let us show that \( G/\mathbb{Q} \) and \( H/\mathbb{Q} \) are incompatible \( \sigma \)-field extensions. Indeed, suppose that there exists a \( \sigma \)-field extension \( E \) of \( \mathbb{Q} \) and \( \sigma \)-isomorphisms \( \phi \) and \( \psi \), respectively, of \( G/\mathbb{Q} \) and \( H/\mathbb{Q} \) into \( E/\mathbb{Q} \). Let \( j = \phi(i) \), \( k = \psi(i) \), and let \( \gamma \) denote the translation of \( E \) that extends \( \alpha \) and \( \beta \). Then \( j^2 = k^2 = -1 \) whence either \( j = k \) or \( j = -k \). Since \( \gamma(j) = j \) and \( \gamma(k) = -k \), in both cases we obtain that \( j = -j \), that is, \( j = 0 \). This contradiction implies that the \( \sigma \)-field extensions \( G/\mathbb{Q} \) and \( H/\mathbb{Q} \) are incompatible.

The existence of incompatible extensions plays an important part in the development of the theory of difference algebra. Of particular concern here is the fact that the presence of incompatible extensions can inhibit the extension of difference isomorphisms for difference field extensions.

**Example 4.6.2** \([32\text{, Chapter 9, Example 1}\). As in the preceding example, let us consider \( \mathbb{Q} \) as an ordinary difference field with a basic set \( \sigma = \{\alpha\} \) where \( \alpha \) is the identity automorphism. Let \( a \) denote the positive fourth root of 2 (we assume \( a \in \mathbb{C} \)) and let \( i \in \mathbb{C} \) be the square root of \(-1\). Then the field \( \mathbb{Q}(a, i) \) can be treated as a \( \sigma \)-field extension
of \( Q \) such that \( \alpha(i) = -i \) and \( \alpha(a) = -a \). Obviously, \( Q(i)/Q \) can be treated as a \( \sigma \)-field subextension of \( Q(a)/Q \). Let \( \phi \) be a \( \sigma \)-isomorphism of \( Q(a,i)/Q(i) \) into some \( \sigma \)-field extension \( M/Q(i) \) where \( M \) is a \( \sigma \)-overfield of \( Q(a,i) \). Since the field extension \( Q(a,i)/Q(i) \) is normal, \( \phi \) is an automorphism of \( Q(a,i) \). Clearly, \( \phi(a) = i^k a \) for some \( k \in \mathbb{N} \). Since \( \alpha(a) = -a \), \( \alpha(i^k a) = -i^k a \). But \( \alpha(i^k a) = (-1)^{k+1}(i^k a) \), so \( k \) is even and either \( \phi(a) = a \) or \( \phi(a) = -a \). It follows that every \( \sigma \)-isomorphism of \( Q(a,i)/Q(i) \) leaves the elements of \( Q(i)(a^2) \) fixed. On the other hand, there exists a \( \sigma \)-automorphism \( \psi \) of \( Q(i)(a^2)/Q(i) \) such that \( \psi(a^2) = -a^2 \). Clearly, such a \( \sigma \)-automorphism has no extension to a \( \sigma \)-isomorphism of \( Q(a,i)/Q(i) \).

Let \( F \) be a difference field with a basic set \( \sigma \) and \( G/F, H/F \) two \( \sigma \)-field extensions of \( F \). Then the number of \( \sigma \)-isomorphisms of \( G/F \) into \( H/F \) is called the replicability of \( G/F \) in \( H/F \). The replicability of the \( \sigma \)-field extension \( G/F \) is defined as the maximum of replicabilities of \( G/F \) in all \( \sigma \)-field extensions of \( F \), if this maximum exists, or \( \infty \) if it does not. Clearly, if \( G/F \) is finitely generated (as a \( \sigma \)-field extension), then it is sufficient to define the replicability of \( G/F \) considering only \( \sigma \)-overfields in the universal system over \( F \).

**Theorem 4.6.3** [32, Chapter 7, Theorem II]. A necessary condition for finite replicability of an ordinary difference field extension \( G/F \) is that every element of \( G \) have a transform of some order that is algebraic over \( F \). This condition is sufficient if \( G \) is finitely generated over \( F \).

Let \( F \) be an ordinary difference field with a basic set \( \sigma \) and \( a \) an indexing of elements lying in a \( \sigma \)-overfield of \( F \). We say that almost every specialization of \( a \) over \( F \) has property \( P \) if there exists a non-zero element \( u \in F[a] \) such that every specialization of \( a \) over \( F \) which does not specialize \( u \) to 0 has property \( P \). The statement that almost every specialization which has property \( P \) has property \( Q \) means that almost every specialization has the property “not \( P \) or \( Q \)”.

In [32, Chapter 7] R.M. Cohn showed that a set of elements and a specialization of the set can generate incompatible difference field extensions over a given difference field. At the same time, the following statement implies that “in most cases” a finite indexing and one of its specializations generate compatible extensions.

**Theorem 4.6.4.** Let \( F \) be an ordinary difference field with a basic set \( \sigma \) and \( a \) a finite indexing of elements from some \( \sigma \)-overfield of \( F \). Then almost every specialization of \( a \) over \( F \) generates a \( \sigma \)-field extension of \( F \) compatible with \( F\langle a \rangle/F \).

The last theorem implies that if \( M \) is a non-empty irreducible variety over an ordinary difference field \( F \), then almost every solution in \( M \) has the property of compatibility with a generic zero of \( M \). That is, there exists a difference polynomial \( A \notin \Phi(M) \) such that if \( (\eta_1, \ldots, \eta_s) \) is a generic zero of \( M \) and \( (\eta'_1, \ldots, \eta'_s) \) a solution in \( M \) not annulling \( A \), then \( F\langle \eta_1, \ldots, \eta_s \rangle/F \) and \( F\langle \eta'_1, \ldots, \eta'_s \rangle/F \) are compatible.

**Theorem 4.6.5** [32, Chapter 7, Theorem IV]. Let \( F \) be an ordinary difference field with a basic set \( \sigma \). Let \( a \) and \( b \) be finite indexings of elements lying in a \( \sigma \)-overfield of \( F \), \( u \) a non-
zero element of $F\{a, b\}$, and $B$ a $\sigma$-transcendence basis of $b$ over $F\{a\}$. If $F\{a, b\}/F\{a\}$ is primary, then almost every specialization $a' = \phi a$ of $a$ over $F$ can be extended to a specialization $a', b'$ of $a, b$ in such a way that

(i) The specialization of $u$ is not 0.
(ii) Every $\sigma$-transcendence basis of $b$ over $F\{a\}$ specializes to a $\sigma$-transcendence basis of $b'$ over $F\{a'\}$.
(iii) If $B'$ denote the specialization of $B$, then $\text{Eord } F\{a', b'\}/F\{a', B'\} = \text{Eord } F\{a, b\}/F\{a, B\}$.

Let $G$ be a difference overfield of an ordinary difference field $F$ with a basic set $\sigma$. The core $G_F$ of $G$ over $F$ is defined to be the set of elements $a \in G$ algebraic and separable over $F$ and such that $\text{ld } F\{a\}/F = 1$. (It follows from Theorem 4.3.1 that $G_F$ is a $\sigma$-field and $\text{ld } G_F/F = 1$.) Example 4.6.1 shows that a core $G_F$ need not to be $F$. Furthermore, if $\text{Char } F = 0$ or $G/F$ is separable, it follows from Theorem 4.3.4 that $G = G_F$ if and only if $G : F$ is finite.

Suppose now that $F$ is an inversive ordinary difference field with a basic set $\sigma$ and $G$ an algebraic $\sigma$-overfield of $F$ (that is, $G$ is algebraic over $F$ in the usual sense). Since it is possible to define a structure of a $\sigma$-overfield of $G$ on the algebraic closure of this field (see Corollary 4.5.8), one can define a structure of a $\sigma$-overfield of $G$ on a normal closure $G'$ of $G$ over $F$. In what follows, this difference field will be called a normal closure of the $\sigma$-field $G$ over $F$. Clearly, the normal closures of $G$ over $F$ are generated by the $\sigma$-generators of $G/F$ and their conjugates with respect to $F$. Hence, if $G/F$ is a finitely generated $\sigma$-field extension, then any normal closure of $G$ over $F$ is a finitely generated $\sigma$-field extension of $F$. If $G$ is inversive, the normal closures of $G$ over $F$ are inversive.

**Proposition 4.6.6.** Let $F$ be an inversive ordinary difference field with a basic set $\sigma$, $G$ an algebraic $\sigma$-overfield of $F$, and $G'$ a normal closure of $G$ over $F$. Then $G'_F$ contains a normal closure of $G_F$ over $F$.

Let $F$ be an ordinary difference field with a basic set $\sigma = \{\alpha\}$ and let $G$ be a finitely generated $\sigma$-field extension of $F$ which is an algebraic, normal and separable overfield of $F$. It is easy to see that one can choose a finite set of $\sigma$-generators $S$ of $G$ over $F$ such that $F(S) = F(v)$ (hence $G = F(v)$) and $v$ is normal over $F$. Furthermore, there exists $k \in \mathbb{N}$ such that $F(v, \ldots, \alpha^k v) : F(v, \ldots, \alpha^{k-1} v) = \text{ld } F(v)/F$. Let $w$ be such that $F(w) = F(v, \ldots, \alpha^{k-1} v)$. Then $w$ is normal over $F$, $F(w) = G$, and $F(w, \alpha w) : F(w) = \text{ld } F(w)/F$. An element with these properties is called the standard generator of $G$ over $F$.

If $w$ is a standard generator such that $F(w) : F$ is as small as possible, then $w$ is called a minimal standard generator.

With the above assumptions, if there exists an element $u \in G$ such that $G = F\{u\}$, $u$ is normal over $F$, and $F(u) : F = \text{ld } G/F$, then $G/F$ is said to be a benign $\sigma$-field extension. (Clearly, a $\sigma$-generator of a benign extension has the properties ascribed to $u$ if and only if it is a minimal standard generator.)

**Proposition 4.6.7.** Let $G/F$ be a benign ordinary difference field extension with a basic set $\sigma$ and let $u$ be a minimal standard generator of $G$ over $F$. Then
(i) $G/F$ is compatible with every $\sigma$-field extension of $F$.
(ii) The replicability of $G/F$ is $\text{ld } G/F$.
(iii) If $K$ is a $\sigma$-subfield of $F$, $\eta$ a set of $\sigma$-generators of $F$ over $K$, and $0 \neq v \in G$, then almost every specialization of $\eta$ over $K$ can be extended to a specialization of $\eta,u,v$ over $K$ such that the specialization of $v$ is not 0.
(iv) Let $F(u')$ be a $\sigma$-field extension of $F$ such that the field extensions $F(u')/F$ and $F(u)/F$ are isomorphic with $u$ corresponding to $u'$. Then $F(u')/F$ and $F(u)$ are isomorphic $\sigma$-field extensions with $u$ corresponding to $u'$.

Let $K$ be an inversive difference field with a basic set $\sigma$ and let $K_1$ and $K_2$ be $\sigma$-subfields of this field whose inversive closures in $K$ coincide with $K$. Let $L_1$ and $L_2$ be $\sigma$-field extensions of $K_1$ and $K_2$, respectively, and let $L_1'$ and $L_2'$ be the inversive closures of $L_1$ and $L_2$, respectively. The $\sigma$-field extensions $L_1/K_1$ and $L_2/K_2$ are called equivalent if $L_1'/K$ and $L_2'/K$ are $\sigma$-isomorphic.

**Theorem 4.6.8 (Babbitt’s decomposition, [2]).** Let $F$ be an ordinary difference field with a basic set $\sigma = \{\alpha\}$ and let $G$ be a finitely generated $\sigma$-field extension of $F$ which is an algebraic, normal and separable overfield of $F$. Then there exist $\sigma$-fields $G_1 \subseteq \cdots \subseteq G_r$, with $G_1 = G_F$, such that $G_i/F$ is equivalent to $G/F$ and for every $i = 2, \ldots, r$, the $\sigma$-field $G_i$ is inversive and $G_i/G_{i-1}$ is equivalent to a benign $\sigma$-field extension of $G_{i-1}$.

The following three propositions on compatibility are consequences of the last theorem.

**Proposition 4.6.9.** Let $F$ be an ordinary difference field with a basic set $\sigma$ and let $G$ and $H$ be two $\sigma$-field extensions of $F$. Then the following statements are equivalent.

(i) $G/F$ and $H/F$ are incompatible.
(ii) There exist finitely generated $\sigma$-field extensions $G'$ and $H'$ of $F$ such that $G' \subseteq G$, $H' \subseteq H$, and $G'/F$ and $H'/F$ are incompatible.
(iii) $G_F/F$ and $H_F/F$ are incompatible.
(iv) $G/F$ and $H/F$ are incompatible.

**Proposition 4.6.10.** Let $F$ be an ordinary difference field with a basic set $\sigma$ and $G$ a $\sigma$-overfield of $F$ such that the field extension $G/F$ is primary. Let $H$ and $K$ be $\sigma$-overfields of $G$ such that $H_G/G$ is equivalent to $G(H_F)/G$, and $K_G/G$ is equivalent to $G(K_F)/G$. Then $H/G$ and $K/G$ are compatible if and only if $H/F$ and $K/F$ are compatible.

**Proposition 4.6.11.** Let $F$ be an ordinary difference field with a basic set $\sigma$ and $G$ a $\sigma$-overfield of $F$ such that $G = F(S)$ for some $\sigma$-algebraically independent over $F$ set $S \subseteq G$.

(i) If $H$ is a $\sigma$-overfield of $G$, then the $\sigma$-field extensions $H_G/G$ and $G(H_F)/G$ are equivalent.
(ii) Let $L$ be a $\sigma$-overfield of $F$ such that $L/F$ is equivalent to $G/F$, and let $M$ and $N$ be two $\sigma$-overfields of $L$. Then $M/G$ and $N/G$ are compatible if and only if $M/F$ and $N/F$ are compatible.
PROPOSITION 4.6.12. Let $F$ be an ordinary difference field with a basic set $\sigma$, and let $a$ and $b$ be finite indexings of elements lying in some $\sigma$-overfield of $F$. Let $G = F(a)$, $H = F(a, b)$, and let $c$ be a finite set of $\sigma$-generators of $H_F$ over $G$. Let $u \neq 0$ be an element of $F(a, b)$ and $B$ a $\sigma$-transcendence basis of $b$ over $F(a)$. Then almost every specialization $a', c'$ of $a, c$ over $F$, can be extended to a specialization $a', b'$ of $a, b$ with the properties (i), (ii), (iii) of Theorem 4.6.5.

THEOREM 4.6.13. Let $F$ be an ordinary difference field of zero characteristic with a basic set $\sigma$. Let $a = (a^{(1)}, \ldots, a^{(s)})$ and $b$ be finite indexings of elements from some $\sigma$-overfield of $F$ such that $a^{(1)}, \ldots, a^{(s)}$ are $\sigma$-algebraically independent over $F$. Furthermore, let $B$ be a $\sigma$-transcendence basis of $b$ over $F(a)$ and $0 \neq u$ an element of $F(a, b)$. Finally, let $R = F\{y_1, \ldots, y_s\}$ be the ring of $\sigma$-polynomials in $\sigma$-indeterminates $y_1, \ldots, y_s$ over $F$. Then there exists a non-zero $\sigma$-polynomial $A \in R$ with the following property: if $c = (c^{(1)}, \ldots, c^{(s)})$ is an indexing in a $\sigma$-overfield of $F$, $c$ is not a solution of $A$, and $F(c)/F$ is compatible with $F(a, b)/F$, then there is a specialization $c, b'$ of $a, b$ with the properties (i), (ii), (iii) of Theorem 4.6.5.

The following three theorems generalize the correspondent statements for ordinary case (see [32, Chapter 7]) to partial difference field extensions.

THEOREM 4.6.14. Let $F$ be an inversive difference field with a basic set $\sigma$ and let $H/F$ and $H/F$ be two $\sigma$-field extensions of $F$. Then

(i) If the extension $G/F$ is primary, then there exists a $\sigma$-field extension $L/F$ such that $G/F$ and $H/F$ have $\sigma$-isomorphisms into $L/F$ with the images of $G$ and $H$ quasi-linearly disjoint over $F$.
(ii) Let $G'$ and $H'$ denote the separable parts of $G$ and $H$, respectively, over $F$. Then $G/F$ and $H/F$ are compatible if and only if the $\sigma$-field extensions $G'/F$ and $H'/F$ are compatible.

THEOREM 4.6.15 [7, Theorem 3.6]. Let $F$ be an inversive difference field with a basic set $\sigma$, $G/F$ a primary $\sigma$-field extension, and $\tau$ a $\sigma$-isomorphism of $F$ into $G$. Then

(i) $\tau$ has an extension to a $\sigma$-isomorphism $\tau_1$ of $G$ into a $\sigma$-overfield $E$ of $G$ such that $E = (G, \tau_1 G)$, $G$ and $\tau_1 G$ are quasi-linearly disjoint over $\tau F$, and $E/G$ is a primary extension. Furthermore, if $G$ is inversive, then $E$ is inversive.
(ii) Let $E'$ be a $\sigma$-overfield of $G$ such that $\tau$ has an extension to a $\sigma$-isomorphism $\tau'$ of $G$ into $E'$ and $E'$ is the free join of $G$ and $\tau' G$ over $\tau F$. Then there exists a unique $\sigma$-isomorphism $\psi$ of $E/G$ onto $E'/G$ such that $\psi \tau_1(a) = \tau' \psi(a)$ for any $a \in G$.

Let $K$ be a difference field with a basic set $\sigma = \{\alpha_1, \ldots, \alpha_n\}$. We say that $K$ satisfies the universal compatibility condition if every two $\sigma$-extensions of $K$ are compatible. $K$ is said to satisfy the stepwise compatibility condition if there exists a permutation $(i_1, \ldots, i_n)$ of $(1, \ldots, n)$ such that all difference fields $(K; \alpha_{i_1}, \ldots, \alpha_{i_k})$, $1 \leq k \leq n$, satisfy the universal compatibility condition. (As in Section 4.3, $(K; \alpha_{i_1}, \ldots, \alpha_{i_k})$ denotes the difference field $K$ with the basic set $\{\alpha_{i_1}, \ldots, \alpha_{i_k}\}$.)
Theorem 4.6.16 [7]. Let \( K \) be a difference field with a basic set \( \sigma, \alpha \in \sigma \), and let \( K' \) denote the field \( K \) treated as a difference field with the basic set \( \sigma \setminus \{\alpha\} \). If \( K' \) satisfies the stepwise compatibility condition, then there exists a \( \sigma \)-overfield \( L \) of \( K \) such that \( L \) is an algebraic closure of \( K \).

Although any ordinary difference field \( K \) has a difference overfield which is an algebraic closure of \( K \) (see Corollary 4.5.8), the following example shows that this result cannot be generalized to partial difference fields. Furthermore, the converse of theorem 4.6.16 is not true (see [7, Example 3.9]).

Example 4.6.17 [7]. Let us consider \( \mathbb{Q} \) as a difference field with a basic set \( \sigma = \{\alpha_1, \alpha_2\} \) where \( \alpha_1 \) and \( \alpha_2 \) are the identity automorphisms. Let \( b \) and \( i \) denote the positive square root of 2 and the square root of \(-1\), respectively (we assume \( b, i \in \mathbb{C} \)), and let \( F = \mathbb{Q}(b, i) \). Let us extend \( \alpha_1 \) and \( \alpha_2 \) to automorphisms of the field \( F \) by setting \( \alpha_1(b) = -b, \; \alpha_1(i) = i, \; \alpha_2(b) = b, \; \alpha_2(i) = -i \), and consider \( F \) as a \( \sigma \)-overfield of \( \mathbb{Q} \).

Then there exists no \( \sigma \)-overfield of \( F \) which is an algebraic closure of \( F \). Indeed, if there were one, then there exists a \( \sigma \)-overfield \( G \) of \( F \) which contains an element \( a \) such that \( a^2 = b \). Since \( (\alpha_1(a))^2 = \alpha_1(a^2) = -b \) and \( (\alpha_2(a))^2 = \alpha_2(a^2) = b \), we have \( \alpha_1(a) = \lambda a i \) and \( \alpha_2(a) = \mu a \) where \( \lambda \) and \( \mu \) denote plus or minus 1. Then \( \alpha_1 \alpha_2(a) = \lambda \mu i a \) and \( \alpha_2 \alpha_1(a) = -\lambda \mu i a \). Thus, \( \alpha_1 \) and \( \alpha_2 \) do not commute at \( a \), which contradicts the assumption that \( G \) is a \( \sigma \)-overfield of \( F \).

Remark 4.6.18. As we have seen in Section 4.5, every ordinary algebraically irreducible difference polynomial has an abstract solution. One can use the last example to show that this result cannot be extended to partial difference polynomials. Indeed, let \( F \) and \( b \) be as in the example, and let \( F[y] \) be a ring of \( \sigma \)-polynomials in one \( \sigma \)-indeterminate \( y \) over \( F \). If \( a \) is a solution for \( A = y^2 - b \in F[y] \), then \( a^2 = b \) which, as demonstrated in Example 4.6.17, is impossible.

The following statement is the version of the existence theorem for difference fields with two translations.

Theorem 4.6.19 [7, Theorem 6.1]. Let \( F \) be an inversive difference field with a basic set \( \sigma \) consisting of two translations. Let \( R = F\{y_1, \ldots, y_s\} \) be the ring of \( \sigma \)-polynomials in \( \sigma \)-indeterminates \( y_1, \ldots, y_s \) over \( F \). Furthermore, suppose that there exists a translation \( \alpha \in \sigma \) such that if \( F \) is treated as a difference field with the basic set \( \sigma' = \sigma \setminus \{\alpha\} \) (we denote this \( \sigma' \)-field by \( F' \)), then every two \( \sigma' \)-field extensions of \( F' \) are compatible.

Then every algebraically irreducible \( \sigma \)-polynomial \( A \in R \setminus F \) has a solution \( \eta = (\eta_1, \ldots, \eta_s) \) with the following properties.

(i) \( \eta \) is not a proper specialization over \( F \) of any solution of \( A \).

(ii) If \( A \) contains a transform of some \( y_k \) \((1 \leq k \leq s)\), then the elements \( \eta_1, \ldots, \eta_{k-1}, \eta_{k+1}, \ldots, \eta_s \) are \( \sigma \)-algebraically independent over \( F \). Furthermore, if a \( \sigma \)-polynomial \( B \in R \setminus F \) is annulled by \( \eta \) and contains only those transforms of \( y_k \) which are contained in \( A \), then \( B \) is a multiple of \( A \).
COROLLARY 4.6.20.
(i) If \( F \) and \( A \) are is in the last theorem, then \( A \) has at most a finite number of isomorphically distinct solutions of the type described in the theorem.
(ii) The conclusion of Theorem 4.6.19 holds if \( F \) is an inversive difference field with two translations and \( F \) is separably algebraically closed or algebraically closed.

COROLLARY 4.6.21. Let \( F \) be a difference field whose basic set \( \sigma \) consists of two translations, and let \( R = F\{y_1, \ldots, y_s\} \) be the ring of \( \sigma \)-polynomials in \( \sigma \)-indeterminates \( y_1, \ldots, y_s \) over \( F \). Then the following statements are equivalent:
(a) \( F \) has an algebraic closure which is a \( \sigma \)-field extension of \( F \).
(b) Every algebraically irreducible \( \sigma \)-polynomial \( A \in R \setminus F \) has a solution \( \eta \) with the properties stated in Theorem 4.6.19.
(c) If \( a \) is any element that is separably algebraic and normal over \( F \), then \( F \) may be extended to the \( \sigma \)-field \( F(\langle a \rangle) \).

4.7. Isomorphisms of difference fields. Monadicity

We have already seen that difference field isomorphisms cannot be extended as freely as field isomorphisms. The following theorem proved in [32, Chapter 9] gives some conditions under which \( \sigma \)-isomorphisms of a given difference (\( \sigma \)-) field can be extended.

THEOREM 4.7.1. Let \( F \) be an ordinary difference field with a basic set \( \sigma \) and \( G \) a \( \sigma \)-overfield of \( F \). If \( G/F \) is compatible with every \( \sigma \)-field extension of \( F \) (in particular, if \( GF = F \)), then every \( \sigma \)-isomorphism of \( F \) into a \( \sigma \)-overfield \( H \) of \( G \) extends to a \( \sigma \)-isomorphism of \( G \) into a \( \sigma \)-overfield of \( H \).

COROLLARY 4.7.2. Let \( F \) be an ordinary difference field with a basic set \( \sigma \), \( G \) a \( \sigma \)-overfield of \( F \), and \( H \) a \( \sigma \)-field extension of \( G \). Then the replicability of \( H/F \) is not less than the replicability of \( HG/F \).

Let \( F \) be a difference field with a basic set \( \sigma \). A \( \sigma \)-field extension \( G/F \) is called monadic if its replicability is 1, that is, \( G/F \) has at most one \( \sigma \)-isomorphism into any \( \sigma \)-field extension of \( F \). It follows from Corollary 4.7.2 that a monadic extension has a monadic core.

A monadic extension \( G/F \) is called properly monadic if \( G \neq F \) and there exists an element \( a \in G \) such that no transform of \( a \) belongs to \( F \).

PROPOSITION 4.7.3 [2]. Let \( F \) be an inversive difference field with a basic set \( \sigma \) and let \( G \) and \( H \) be \( \sigma \)-overfields of \( F \). Then
(i) \( G/F \) and \( H/F \) are incompatible if and only if the inversive closures \( G^* \) and \( H^* \) are incompatible \( \sigma \)-field extensions of \( F \).
(ii) The \( \sigma \)-field extension \( G/F \) is monadic if and only if \( G^*/F \) is monadic.

EXAMPLE 4.7.4 [32, Chapter 1, Example 6]. Let \( G = \mathbb{C}(x) \) be the field of rational fractions in one variable \( x \) over \( \mathbb{C} \). Let us consider \( G \) as an ordinary difference field with a
basic set $\sigma = \{a\}$ where $\alpha f(x) = f(x^2)$ for any $f(x) \in F$. If $t = x^3$ and $F = \mathbb{C}(t)$, then the field $F$ can be viewed as a $\sigma$-subfield of $G$. Let us show that $G/F$ is a monadic $\sigma$-field extension.

Let $\phi$ and $\psi$ be two $\sigma$-isomorphisms of $G/F$ into some $\sigma$-field extension $H/F$. Let $y = \phi(x)$ and $z = \psi(x)$. Then $y^3 = t$, $\alpha(y) = y^2$, $z^3 = t$, and $\alpha(z) = z^2$. Clearly, in order to prove that $\phi = \psi$, one should show that $z = y$. Since $t = y^3 = z^3$, we find $z = \omega y$ where $\omega^3 = 1$, $\omega \in \mathbb{C} \subseteq F$. It follows that $\alpha(\omega) = \omega$ and $\omega \alpha(y) = \omega^2 y^2$, $\alpha(y) = \omega y^2$ (we use the fact that $\alpha(z) = z^2$). Since $\alpha(y) \neq 0$, $\omega = 1$ whence $z = y$ and $\phi = \psi$.

In what follows we show that a finitely generated separable monadic extension coincides with its core (generally speaking, this is not true for extensions of finite replicability). The proofs of the results of the rest of this section can be found in [2] and in [32, Chapter 9].

**Proposition 4.7.5.** Let $F$ be an ordinary difference field with a basic set $\sigma$, $G$ an algebraic $\sigma$-overfield of $F$, and $G'$ the separable part of $G$. Then $G/F$ is monadic if and only if $G'/F$ is monadic.

**Theorem 4.7.6.** Let $F$ be an inversive ordinary difference field with a basic set $\sigma$ and $G/F$ a finitely generated monadic $\sigma$-field extension of $F$. Then:

(i) $G$ is purely inseparable over its core.
(ii) If $G/F$ is separable, $G = G_F$.
(iii) $\text{ord } G/F = 0$ and $\text{rld } G/F = 1$. (If $\text{Char } F = 0$, then $\text{ld } G/F = 1$.)

**Theorem 4.7.7.** Let $F$ be an inversive ordinary difference ($\sigma$-)field, $B$ a $\sigma$-algebraically independent set over $F$, and $G$ the inversive closure of $F(B)$. If $H/G$ is a finitely generated, separable monadic $\sigma$-field extension of $G$, then there exists a finitely generated monadic $\sigma$-field extension $K/F$ such that $H = G(K)$. Conversely, if $K/F$ is monadic (whether finitely generated or not), then $H/G$ is monadic, where $H = G(K)$.

**Theorem 4.7.8.** Let $F$ be an ordinary difference field with a basic set $\sigma$ and $G$ a $\sigma$-overfield of $F$ of finite degree (that is $|G:F| < \infty$). Then $G/F$ is compatible with every $\sigma$-field extension of $F$ if and only if it is monadic.

**Corollary 4.7.9.** A difference subextension of a finitely generated monadic extension of an ordinary difference field is monadic.

Notice that the restriction to finitely generated difference field extensions made in the last statement is essential (see [32, Chapter 9, Example 4]).

A difference ($\sigma$-) field extension $G/F$ is called pathological if either it is incompatible with some other $\sigma$-field extension of $F$ or $G/F$ is monadic.

**Example 4.7.10.** Let $F = \mathbb{C}(x)$ be the field of rational fractions in one variable $x$ over $\mathbb{C}$ considered as an ordinary difference field with a basic set $\sigma = \{\alpha\}$ where $\alpha f(x) = f(x + 1)$ for any $f(x) \in F$. Then $F$ has no finitely generated pathological extensions (see [2, Theorem 2.9]).
Example 4.7.11. Let us consider the field $\mathbb{C}(x)$ from the previous example as an ordinary difference field $K$ with a translation $\sigma: f(x) \mapsto f(qx)$ ($f(x) \in \mathbb{C}(x)$), where $q$ is a non-zero complex number such that $q^n \neq 1$ for every $n \in \mathbb{N}$. By [32, Chapter 9, Theorem XX], a $\sigma$-field extension $G/K$ is incompatible with some other $\sigma$-field extension of $K$ if and only if $G$ contains a $k$-th root of $x$ for some $k \in \mathbb{N}, k > 1$. Therefore, $K$ has no finitely generated properly monadic extensions.

Theorem 4.7.12 [2]. If an ordinary inversive difference field admits a pathological extension, it admits a pathological extension of finite degree of the same type.

4.8. Difference valuation rings and extensions of difference specializations

Let $K$ be a difference field with a basic set $\sigma$. A maximal difference (or $\sigma$-) specialization of $K$ is a $\sigma$-homomorphism $\phi$ of a $\sigma$-subring $R$ of $K$ onto a difference ($\sigma$-) domain $\Lambda$ such that $\phi$ cannot be extended to a $\sigma$-homomorphism of a larger $\sigma$-subring of $K$ onto a domain which is a $\sigma$-overring of $\Lambda$. (It can be easily shown that the $\sigma$-domain $\Lambda$ is, in fact, a $\sigma$-field.) The $\sigma$-domain $R \subseteq K$ is called the maximal difference (or $\sigma$-) ring of $K$. If $K$ is the quotient field of $R$, we say that $R$ is a difference (or $\sigma$-) valuation ring of $K$, and $\phi$ is called a difference (or $\sigma$-) place of $K$. It is easy to see that if the $\sigma$-field $K$ is inversive, then every its maximal $\sigma$-ring is also inversive.

A difference ring $R$ with a basic set $\sigma$ is called a local difference (or $\sigma$-) ring if the non-units of $R$ form a $\sigma$-ideal. This ideal will be denoted by $M(R)$. If $R_0$ is a local $\sigma$-subring of $R$ and $M(R) \cap R_0 = M(R_0)$, we say that $R$ dominates $R_0$. Clearly, a localization $R_P$ of a difference ($\sigma$-) ring $R$ by a prime $\sigma$-ideal $P$ of $R$ is a local $\sigma$-ring.

In what follows, we present some results on extensions of difference specializations that are natural generalizations of the corresponding statements proved in [97] for ordinary difference rings and fields.

Proposition 4.8.1. Let $K$ be a difference field with a basic set $\sigma$ and $R$ a local $\sigma$-subring of $K$. Then the following statements are equivalent.

(i) $R$ is a maximal $\sigma$-ring of $K$.

(ii) $R$ is maximal among local $\sigma$-subrings of $K$ ordered by domination.

(iii) If $x \in K$ and $x \notin R$, then $1 \in \{R[x]M(R)\}$, the perfect $\sigma$-ideal generated by $M(R)$ in $R[x]$.

Let $K$ be a difference field with a basic set $\sigma$ and $R$ a difference valuation ring of $K$. Then the set $U$ of units of $R$ forms a subgroup of $K' = K \setminus \{0\}$ and one may define the natural homomorphism $v: K' \rightarrow K'/U$. Let $K'/U$ be denoted by $F$, with the operation written as addition. Then $v$ is called a difference (or $\sigma$-) valuation of $K$. Let $\gamma^+ = v(M(R) \setminus \{0\})$; then for $a \in \gamma^+$ we have $-a \notin \gamma^+$. For any $a, b \in \gamma$ we define $a < b$ if $b - a \in \gamma^+$. Then $\gamma$ becomes a partially ordered group which is not necessarily linearly ordered. Clearly, $x \in R \setminus \{0\}$ if and only if $v(x) \geq 0$, and $x \in M(R) \setminus \{0\}$ if and only if $v(x) > 0$.

Notice that there are difference valuation rings which are not valuation rings (and, thus, difference valuations which are not valuations), see the example in [97, Section 1].
THEOREM 4.8.2. Let $R$ be a local difference subring of a difference field $K$ with a basic set $\sigma$. $M(R)$ the maximal ideal of $R$, and $x \in K$. Then the natural $\sigma$-homomorphism $\phi: R \to R/M(R)$ extends to one sending $x$ to 0 if and only if $1 \notin [x]$ in $R[x]$.

If $R$ is a maximal difference ring of $K$, then $x \in M(R)$ if and only if $1 \notin [x]$.

COROLLARY 4.8.3. Let $K$ be a difference field with a basic set $\sigma$ and let $R$ be a maximal difference ring of $K$.

(i) If $S$ is another maximal difference ring of $K$ such that $R \subseteq S$, then $M(S) \subseteq M(R)$.

(ii) Let $P$ be a prime $\sigma^*$-ideal of $R$. Then there is a maximal difference ring $R_1$ of $K$ such that $R \subseteq R_1$ and $M(R_1) = P$.

(iii) The prime $\sigma^*$-ideals of $R$ are linearly ordered by inclusion.

(iv) Every perfect $\sigma$-ideal of $R$ is prime.

(v) Let $S$ be a maximal difference ring of $K$ with a specialization $\phi: S \to \Lambda$, and let $R \subseteq S$. Then $\phi(R)$ is a maximal difference ring of $\Lambda$.

Let $K$ be a difference field with a basic set $\sigma$ and $K[\gamma]$ a ring of $\sigma$-polynomials in one $\sigma$-indeterminate $\gamma$ over $K$. Let $R$ be a $\sigma$-subring of $K$ and $g$ a $\sigma$-polynomial in $R[\gamma]$ with a constant term $b \in R$. Furthermore, let $\{g\}_R$ and $\{g\}_K$ denote perfect $\sigma$-ideals generated by $g$ in the $\sigma$-rings $R[\gamma]$ and $K[\gamma]$, respectively.

PROPOSITION 4.8.4. With the above notation, let the $\sigma$-ideal $\{g\}_K$ be prime, $\eta$ a generic zero of $\{g\}_K$, and $\phi: R \to \Lambda$ a difference specialization of $K$ with $\phi(b) = 0$. If $\{g\}_K \cap R[\gamma] = \{g\}_R$, then $\phi$ can be extended to $R[\eta]$ with $\phi(\eta) = 0$.

PROPOSITION 4.8.5. Let $K$ be a difference field with a basic set $\sigma$, $R_0$ a $\sigma$-subring of $K$ with prime $\sigma^*$-ideals $P$ and $Q$ such that $P \subseteq Q$. Let $S$ be a proper maximal $\sigma$-ring of $K$ with $R_0 \subseteq S$ and $M(S) \cap R_0 = P$. Then there exists a proper maximal $\sigma$-ring $R$ of $K$ such that $R_0 \subseteq R$ and $M(R) \cap R_0 = Q$. Furthermore, if $S$ is a $\sigma$-valuation ring of $K$ then $R$ is also.

COROLLARY 4.8.6. Let $K$ be a difference field with a basic set $\sigma$ and $R_0$ a local $\sigma$-subring of $K$. Let $L$ be a $\sigma$-overfield of $K$ and $S$ a proper maximal $\sigma$-ring (valuation ring) of $L$ containing $R_0$. Then there exists a proper maximal $\sigma$-ring (valuation ring) $R$ of $L$ dominating $R_0$.

The existence of $S$ in the corollary is equivalent to the condition that $L$ has a subring $S_0$, $S_0 \subseteq R_0$, which contains a proper non-zero prime $\sigma^*$-ideal. This condition does not always hold. For example, if $Q$ is treated as an ordinary difference ($\sigma$-) field with the identity translation $\alpha$ and $Q\{b\}$ is a $\sigma$-overring of $Q$ such that $b$ is transcendental over $Q$ and $\alpha(b) = b$, then the $\sigma$-specialization $b \to 1$ does not extend to a $\sigma$-place $Q(a)$ where $a^2 = b$ and $\alpha(a) = -a$ (see [32, Chapter 7, Example 3]). However, the condition does hold in the situation of the following proposition.

PROPOSITION 4.8.7. Let $R_0$ be a local difference ring with a basic set $\sigma$ and $K$ the difference quotient field of $R_0$. 
(i) Suppose that $R_0$ does not contain minimal non-zero prime $\sigma^*$-ideals. If $L$ is a primary finitely generated $\sigma$-field extension of $K$, $L = K \langle \eta_1, \ldots, \eta_\ell \rangle$, then there exists a difference valuation ring $R$ of $L$ dominating $R_0$.

(ii) Let $\zeta$ be an element from some $\sigma$-overfield of $K$ which is $\sigma$-algebraically independent over $K$. If $N = K \langle \zeta, \eta_1, \ldots, \eta_\ell \rangle$ is a primary $\sigma$-field extension of $K \langle \zeta \rangle$, then there exists a difference valuation ring of $N$ dominating $R_0$.

5. Difference Galois theory

5.1. Algebraic difference field extensions. Galois correspondence

Let $K$ be an inversive difference field with a basic set $\sigma = \{ \alpha_1, \ldots, \alpha_n \}$ and $L$ a $\sigma^*$-overfield of $K$. As usual, $\text{Gal}(L/K)$ denotes the corresponding Galois group, that is, the group of all automorphisms (not necessarily $\sigma$-automorphisms) that leave the field $K$ fixed. It is easy to see that the mappings $\bar{\alpha}_i : \theta \mapsto \alpha_i^{-1} \theta \alpha_i$ ($\theta \in \text{Gal}(L/K)$, $1 \leq i \leq n$) are automorphisms of the group $\text{Gal}(L/K)$; they are called the induced automorphisms of $\text{Gal}(L/K)$. A subgroup $B$ of $\text{Gal}(L/K)$ is called $\sigma$-stable if $\bar{\alpha}_i(B) = B$ for $i = 1, \ldots, n$. If $\bar{\alpha}_i(b) = b$ for every $b \in B, \alpha_i \in \sigma$, the subgroup $B$ is called $\sigma$-invariant. The largest $\sigma$-invariant subgroup of $\text{Gal}(L/K)$ consists of all $\sigma$-automorphisms of $L$ that leave the field $K$ fixed. This group is called the difference or $\sigma$-Galois group of $L/K$; it is denoted by $\text{Gal}_\sigma(L/K)$.

If $M/K$ is a $\sigma^*$-field subextension of $L/K$, then $\{ \theta \in \text{Gal}(L/K) \mid \theta(a) = a$ for every $a \in M \}$ is a $\sigma$-stable subgroup of $\text{Gal}(L/K)$ denoted by $M'$. Also, if $B$ is an $\sigma$-stable subgroup of $\text{Gal}(L/K)$, then $\{ a \in L \mid \theta(a) = a$ for every $\theta \in B \}$ is a $\sigma^*$-overfield of $K$ denoted by $B'$. As usual, we denote the field $(M')'$ by $M''$ and the group $(B')'$ by $B''$.

In what follows, the group $\text{Gal}(L/K)$ is considered as a topological group with the Krull topology. A fundamental system of neighborhoods of the identity in this topology is the set of all groups $\text{Gal}(L/M) \subseteq \text{Gal}(L/K)$ such that $M$ is a subfield of $L$ which is a Galois extension of $K$ of finite degree. (When $L$ is of finite degree over $K$, the Krull topology is discrete.) The topological group $G = \text{Gal}(L/K)$ is compact, Hausdorff, and has a basis at the identity consisting of the collection of invariant, open (and hence closed and of finite index in $G$) subgroups of $G$. If $H$ is a closed invariant stable subgroup of $G$, then each $\bar{\alpha}_i$ induces a topological automorphism $\beta_i$ on $G/H$ such that $\beta_i(gH) = \bar{\alpha}_i(g)H$ for any $g \in G$.

**Theorem 5.1.1 [49].** Let $L$, $K$, and $G = \text{Gal}(L/K)$ be as above. Then

(i) The mapping $M \mapsto M'$ establishes a 1-1 correspondence between the set of $\sigma^*$-field subextensions of $L/K$ and the closed stable subgroups of $G$. If $M/K$ is a $\sigma^*$-field subextension of $L/K$, then $M'' = M$, and if $H$ is a closed stable subgroup of $G$, then $H'' = H$.

(ii) Let $M/K$ be a $\sigma^*$-field subextension of $L/K$ such that $M$ is normal over $K$. Let $\gamma_1, \ldots, \gamma_n$ be the automorphisms of the group $\text{Gal}(M/K)$ induced by $\alpha_1, \ldots, \alpha_n$ (treated as automorphisms of $M/K$), respectively. Then there is a natural isomorphism of topological groups $\phi : G/M' \rightarrow \text{Gal}(M/K)$ such that $\phi \beta_i = \gamma_i \phi$ ($1 \leq i \leq n$) where $\beta_i$ is the automorphism of $G/M'$ induced by $\bar{\alpha}_i$. 
DEFINITION 5.1.2. Let $M$ be an ordinary inversive difference field with a basic set $\sigma$ and $N$ a $\sigma^*$-overfield of $M$. The extension $N/M$ is said to be universally compatible if given a $\sigma^*$-field extension $Q/M$, there exist a $\sigma^*$-field extension $R/M$ and $\sigma^*$-monomorphisms $\phi: N/M \to R/M$ and $\psi: Q/M \to R/M$.

The following two theorems combine the main results on difference Galois groups, universal compatibility and monadicity obtained in [49].

THEOREM 5.1.3. Let $K$ be an ordinary inversive difference field with a basic set $\sigma = \{\alpha\}$ and $L$ a $\sigma^*$-overfield of $K$ such that $L/K$ is a Galois extension. Then:

(i) $\text{Gal}_\sigma(L/K)$ is a closed subgroup of $\text{Gal}(L/K)$.

(ii) Let $\lambda$ be the mapping of $\text{Gal}(L/K)$ into itself such that $\lambda(g) = g^{-1}\bar{\alpha}(g)$ (as before, $\bar{\alpha}$ denotes the automorphism of $\text{Gal}(L/K)$ induced by $\alpha$). Then

(a) $\lambda$ is a continuous function;
(b) $L/K$ is universally compatible if and only if $\lambda$ maps $\text{Gal}(L/K)$ onto itself;
(c) $L/K$ is monadic if and only if $\lambda$ is one-to-one on $\text{Gal}(L/K)$.

(iii) If $L/K$ is universally compatible and $M/K$ a $\sigma^*$-subextension of $L/K$, then $M/K$ is universally compatible.

(iv) Let $M/K$ be a $\sigma^*$-subextension of $L/K$ such that $M/K$ is a Galois extension. If $L/M$ and $M/K$ are universally compatible then $L/K$ is universally compatible. If $L/K$ is monadic and $L/M$ is universally compatible, then $M/K$ is monadic.

(v) $L/K$ is universally compatible if and only if every $\sigma^*$-subextension of $L/K$ which is finite-dimensional Galois over $K$ is universally compatible.

THEOREM 5.1.4. Let $K$ and $L$ be as in Theorem 5.1.3 and let $M/K$ be a $\sigma^*$-subextension of $L/K$ such that $M$ is a Galois extension of $K$. Then

(i) $\text{Gal}_\sigma(M/K)$ is a closed normal stable subgroup of $\text{Gal}_\sigma(L/K)$.

(ii) There exists a natural monomorphism $\psi: \text{Gal}_\sigma(L/K)/\text{Gal}_\sigma(L/M) \to \text{Gal}_\sigma(M/K)$. If $L/M$ is universally compatible then $\psi$ is an isomorphism.

(iii) If $L/K$ is universally compatible and $\psi$ is an isomorphism, then $L/M$ is universally compatible.

5.2. Picard–Vessiot theory of linear homogeneous difference equations

Throughout this section all fields have characteristic zero and all difference fields are inversive and ordinary. If $K$ is such a difference field with a basic set $\sigma = \{\alpha\}$, then its field of constants $\{c \in K \mid \alpha(c) = c\}$ will be denoted by $C_K$. All topological statements of this section will refer to the Zariski topology. As before, $K(y)$ will denote the ring of $\sigma$-polynomials in one $\sigma$-variable $y$ over $K$. Furthermore, for any $n$-tuple $b = (b_1, \ldots, b_n)$, $C^*(b)$ will denote the determinant of the matrix $(\alpha^j b_k)_{0 \leq i \leq n-1, 1 \leq j \leq n}$.

Let us consider a linear homogeneous difference equation of order $n$ over $K$, that is an algebraic difference equation of the form

$$\alpha^n y + a_{n-1}\alpha^{n-1} y + \cdots + a_0 y = 0, \tag{5.2.1}$$
where $a_0, \ldots, a_{n-1}, a_n \in K$ ($n > 0$) and $a_0 \neq 0$. A $\sigma^*$-overfield $M$ of $K$ is said to be a solution field over $K$ for equation (5.2.1) or for the $\sigma$-polynomial $f(y) = a_n y^n + \cdots + a_0 y$, if $M = K(b)^*$ for an $n$-tuple $b = (b_1, \ldots, b_n)$ such that $f(b_j) = 0$ for $j = 1, \ldots, n$ and $C^*(b) \neq 0$. Any such $n$-tuple $b$ is said to be a basis of $M/K$ or a fundamental system of solutions of (5.2.1). If, in addition, $C_K$ is algebraically closed and $C_M = C_K$, then $M$ is said to be a Picard–Vessiot extension (PVE) of $K$ (for equation (5.2.1)).

Note that if $C^*(b) \neq 0$, then the elements $b_1, \ldots, b_n$ are linearly independent over the constant field of any difference field containing them (see [32, Chapter 8, Lemma II]).

**Proposition 5.2.1** [56]. With the above notation, let $M$ be a $\sigma^*$-overfield of $K$ and $R \subseteq C_M$.

(i) A subset of $R$ that is linearly (algebraically) dependent over $K$ is linearly (respectively, algebraically) dependent over $C_K$.

(ii) If $N$ is a $\sigma^*$-overfield of $K$ with $C_N = C_K$, then $N$ and $K(R)$ are linearly disjoint over $K$.

(iii) $C_K(R) = C_K(R)$.

If $M$ is a solution field for equation (5.2.1) over $K$ with basis $b = (b_1, \ldots, b_n)$ and $b'$ is any solution of (5.2.1) in a $\sigma^*$-overfield $N$ of $M$, then $b' = \sum_{i=1}^n c_i b_j$ for some elements $c_1, \ldots, c_n \in C_N$ (see [32, Chapter 8, Theorem XII]). It follows that a $\sigma$-homomorphism $h$ of $K[b]/K$ into a $\sigma^*$-overfield $N$ of $M$ determines an $n \times n$-matrix $(c_{ij})$ over $C_N$ by the equations $h(b_i) = \sum_{j=1}^n c_{ij} b_j$. The following theorem and corollary proved in [56] show that the matrices corresponding to $\sigma$-homomorphisms satisfy a set of algebraic equations over $C_M$, and, in the case of a PVE, form an algebraic matrix group.

**Theorem 5.2.2.** If $M/K$ is a solution field with basis $b = (b_1, \ldots, b_n)$, then there is a set $S_b$ in the polynomial ring $C_M[x_{ij}]$ ($1 \leq i, j \leq n$) so that if $N$ is a $\sigma^*$-overfield of $M$ then the following hold.

(i) A $\sigma$-homomorphism of $K[b]/K$ to $N/K$ determines a matrix over $C_N$ that annihilates every polynomial of $S_b$. (In the last case we say “the matrix satisfies $S_b$”.)

(ii) A matrix over $C_N$ satisfying $S_b$ defines a $\sigma$-homomorphism of $K[b]/K$ to $N/K$.

(iii) If $C_M = C_K$ then a $\sigma$-homomorphism of $K[b]/K$ to $N/K$ determines a $\sigma$-isomorphism if and only if its matrix is non-singular.

**Corollary 5.2.3.** If $M/K$ is a PVE then the difference Galois group $Gal_\sigma(M/K)$ is an algebraic matrix group over $C_K$.

If $M$ is a solution field for equation (5.2.1) and $b = (b_1, \ldots, b_n)$ a basis of $M/K$, then $S_b$ will denote the set of polynomials in Theorem 5.2.2 ($S_b \subseteq C_M[x_{ij}]$, $1 \leq i, j \leq n$), and $T_b$ will denote the variety of $S_b$ over the algebraic closure of $C_M$. The following example shows that a matrix in $T_b$ may not correspond to a difference homomorphism of $K[b]$.

**Example 5.2.4** [56]. Let $b$ be a solution of the difference equation $a\gamma + y = 0$ which is transcendental over $K$. Then the constant field of $K(b)$ contains $C_K(b^2)$, $S_b = \{0\}$, and $T_b$...
contains the algebraic closure of $C_K(b^2)$. Since no $\sigma^*$-overfield of $K(b)^*$ contains $b$ in its constant field, Theorem 5.2.2 does not apply to the matrix $(b)$. The algebraic isomorphism $h : K\{b\} \rightarrow K\{b\}$ defined by $h(b) = b^2$, is not a $\sigma$-homomorphism.

Let $L$ be an intermediate $\sigma^*$-field of a difference $(\sigma^*)$ field extension $M/K$ ($M$ is a solution field for equation (5.2.1) over $K$). A $\sigma^*$-overfield $N$ of $M$ is said to be a universal extension of $M$ for $L$ if every $\sigma$-isomorphism of $L$ over $K$ can be extended to a $\sigma$-isomorphism of $M$ into $N$. It follows from Theorem 4.7.1 that if $L$ is algebraically closed in $M$, then universal extensions of $M$ for $L$ exist. At the same time, Example 5.2.4 shows that even if $L$ itself is a solution field over $K$, $M$ need not be a universal $\sigma^*$-field extension for $L$.

**Proposition 5.2.5.** Let $M$ be a solution field for equation (5.2.1) over a difference field $K$ and $b$ a basis of $M/K$. If the field $K$ is algebraically closed in $M$, then the variety $T_b$ is irreducible and $\dim T_b = \text{trdeg}_K M$.

Let $M$ be a $\sigma^*$-overfield of a difference field $K$ with a basic set $\sigma$. We say that the extension $M/K$ is $\sigma$-normal if for every $x \in M \setminus K$, there exists a $\sigma$-automorphism $\phi$ of $M$ such that $\phi(x) \neq x$ and $\phi(a) = a$ for every $a \in K$. (Note that C. Franke, [56], called such extensions “normal” while similar differential field extensions are called “weakly normal”.) The existence of proper monadic algebraic difference extensions suggests the existence of solution fields that are not $\sigma$-normal extensions.

The following result is a version of the fundamental Galois theorem for PVE. As usual, primes indicate the Galois correspondence.

**Theorem 5.2.6** [56]. Let $M/K$ be a difference $(\sigma^*)$ PVE, $G = \text{Gal}_\sigma(M/K)$, $L$ an intermediate $\sigma^*$-field of $M/K$, and $H$ an algebraic subgroup of $G$. Then

(i) $L^\prime$ is an algebraic matrix group.
(ii) $H'' = H$.
(iii) If $L$ is algebraically closed in $M$, then $M$ is $\sigma$-normal over $L$ and $L'' = L$.
(iv) There is a one-to-one correspondence between intermediate $\sigma^*$-fields of $M/K$ that are algebraically closed in $M$ and connected algebraic subgroups of $G$.
(v) Let $\overline{K}$ denote the algebraic closure of $K$ in $M$. If $H$ is a connected normal subgroup of $G$, then $G/H$ is the full group of $H'$ over $K$ (that is, $G/H$ is isomorphic to $\text{Gal}_\sigma(H'/K)$) and $H'$ is $\sigma$-normal over $\overline{K}$.
(vi) If $L$ is algebraically closed in $M$ and $\sigma$-normal over $K$, then $L'$ is a normal subgroup of $G$ and $G/L'$ is the full group of $L$ over $K$.

Let $M$ be a $\sigma^*$-overfield of a difference $(\sigma^*)$ field $K$ and $H$ a subgroup of $\text{Gal}_\sigma(M/K)$. If $L$ is an intermediate $\sigma^*$-field of $M/K$, then $L'_H$ will denote the group $\{ h \in H \mid h(a) = a \text{ for all } a \in L \} \subseteq H$. A subgroup $A$ of $H$ is said to be Galois closed in $H$ if $(A'_H)^' = A$. An intermediate field $N$ of $M/K$ is said to be Galois closed with respect to $H$ if $(N'_H)^' = N$.

**Theorem 5.2.7** [56,58]. Let $M$ be a solution field for a difference equation (5.2.1) over a difference $(\sigma)$ field $K$. Let $b$ be a basis of $M/K$ and $H$ a subgroup of $\text{Gal}_\sigma(M/K)$ which is naturally isomorphic to the set of matrices $T_b$ corresponding to $H$. Then:
A.B. Levin

(i) Algebraic subgroups of $H$ are Galois closed in $H$.

(ii) Connected subgroups of $H$ correspond to intermediate $\sigma^*$-fields of $M/K$ algebraically closed in $M$.

(iii) Let $L$ be an intermediate $\sigma^*$-field of $M/K$ which is algebraically closed in $M$. Let $T^L_b$ be the variety obtained by considering $M$ as a solution field over $L$ with basis $b$. If $L'_H$ is dense in $T^L_b$, then $(L'_H)^\sigma = L$ and $L'_H$ is connected.

(iv) If the algebraic closure of $K(C_M)$ in $M$ coincides with $K$, then $M/K$ is a normal extension. In this case, there is a one-to-one correspondence between connected algebraic subgroups of $Gal_\sigma(M/K)$ and intermediate $\sigma^*$-fields of $M/K$ algebraically closed in $M$.

Some generalization of the last theorem was obtained in [59]. Let $K$ and $M$ be as in Theorem 5.2.7, $L$ an intermediate $\sigma^*$-field of $M/K$, and $N$ a $\sigma^*$-overfield of $M$. Furthermore, let $I_L$ denote the set of all $\sigma$-isomorphisms of $M$ into $N$ leaving $L$ fixed.

**Proposition 5.2.8.** If $L$ is algebraically closed in $M$, then $I_L$ is a connected algebraic matrix group. Furthermore, $Gal_\sigma(M/L)$ is dense in $I_L$ and $I_L$ is isomorphic to $Gal_\sigma(M(C_N)/L(C_N))$.

**Theorem 5.2.9.** Let $K$ and $M$ be as in Theorem 5.2.7, $H$ an algebraic subgroup of $Gal_\sigma(M/K)$, and $L$ an intermediate $\sigma^*$-field of $M/K$. Then

(i) $H'' = H$.

(ii) If the field $L$ is algebraically closed in $M$, then $L'' = L$ and $M$ is $\sigma$-normal over $L$.

(iii) There is a one-to-one correspondence between connected algebraic subgroups of $Gal_\sigma(M/K)$ and intermediate $\sigma^*$-fields of $M/K$ algebraically closed in $M$.

(iv) Assume that $H$ is connected and $L = H'$. In this case

(a) $H$ is a normal subgroup of $Gal_\sigma(M/K)$ if and only if $L$ is $\sigma$-normal over $K$.

(b) If $H$ is a normal subgroup of $Gal_\sigma(M/K)$ and $N$ is any universal extension of $M$ for $L$, then $I_L$ is a normal subgroup of $I_K$. The homomorphisms defined by restriction and extension determine natural isomorphisms $Gal_\sigma(M/K)/H \rightarrow Gal_\sigma(L/K) \rightarrow I_K/I_L$ and the image of $Gal_\sigma(M/K)/H$ is dense in $I_K/I_L$.

In general a full difference Galois group $G = Gal_\sigma(M/K)$ is not naturally isomorphic to a matrix group (if $g, h \in G$, then the matrix of the composite of $g$ and $h$ is the matrix of $g$ times the matrix obtained by applying $g$ to the entries of the matrix of $h$). However, if we adjoin $C_M$ to $K$ and consider $M$ as a solution field over $K(C_M)$, we obtain a group $D$ which is naturally isomorphic to a group of matrices contained in an algebraic variety $T$. Theorem 5.2.2 implies that $T$ consists only of isomorphisms and singular matrices. The Galois correspondence given in Theorem 5.2.7 for $D$ and fields between $K(C_M)$ and $M$ depends in part on whether a subgroup of $D$ is dense in a variety containing it. Examples where this is not the case are not known.

If $M$ is a solution field for (5.2.1) over $K$ with a basis $b$, then the subsets of $T_b$ and $D$ consisting of non-singular matrices with entries in $C_K$ are automorphism groups. The following two results obtained in [56] deal with these groups.
PROPOSITION 5.2.10. Let $M$ be a solution field for the difference equation (5.2.1) over a difference $(\sigma)$-field $K$. Let $b$ be a basis of $M/K$, $\Lambda$ a subfield of $C_M$, and $S_b$ the subset of the polynomial ring $C_M[x_{ij}]$ $(1 \leq i, j \leq n)$ whose existence is established by Theorem 5.2.2. Then there exists a set $S_b' \subseteq \Lambda[x_{ij}]$ so that the following hold.

(i) Every solution of the set $S_b'$ is a solution of $S_b$.
(ii) Every solution of $S_b$ that lies in $\Lambda$ is a solution of $S_b'$.
(iii) If $\Lambda$ is algebraically closed and contained in $K$, then the variety of $S_b'$ over $\Lambda$ is an algebraic matrix group of automorphisms of $M/K$ plus singular matrices.

Note that Theorem 5.2.7 can be applied to any group $G_b^{(1)}$ obtained by deleting the singular matrices from a variety $T_b^{(1)}$ determined as in Proposition 5.2.10 by a basis $b$ and a subfield $\Lambda$.

PROPOSITION 5.2.11. Let $K$, $M$ and $b$ be as in Proposition 5.2.10, and let $\Lambda$ be an algebraically closed field of constants of $K$. Let $G_b^{(1)}$ be the group determined by $b$ and $\Lambda$ as in Proposition 5.2.10 and let $C_b^{(1)}$ be the component of the identity of $G_b^{(1)}$. Finally, let $C_b$ be the irreducible subvariety of $T_b$ determined by $\overline{K}$ (the algebraic closure of $K$ in $M$). The following are equivalent and imply that $\overline{K}$ is Galois closed with respect to $C_b^{(1)}$.

(i) $C_b^{(1)}$ is dense in $C_b$.
(ii) $\dim C_b = \dim C_b^{(1)}$.
(iii) There is a basis for the ideal of $C_b$ in the polynomial ring $C[x_{ij}]$ $(1 \leq i, j \leq n)$.

Let $M$ be a solution field for difference equation (5.2.1) over a difference $(\sigma)$-field $K$. $M/K$ is called a generalized Picard–Vessiot extension (GPVE) if there is a basis $b$ of $M/K$ and an algebraically closed subfield $\Lambda$ of $C_K$ such that $C_b^{(1)}$ is dense in $C_b$. $M$ is said to be a generic solution field for equation (5.2.1) if $\text{trdeg}_K M = n^2$ ($n$ is the order of the difference equation).

PROPOSITION 5.2.12. Every linear homogeneous difference equation over a difference field $K$ has a generic solution field $M$. Therefore, if $C_K$ contains an algebraically closed subfield, then every linear homogeneous difference equation over $K$ has a solution field which is a GPVE.

THEOREM 5.2.13. If $L = K\langle a \rangle^*$ and $M = K\langle b \rangle^*$ are solution fields of equation (5.2.1) over a difference $(\sigma)$-field $K$, then $\text{trdeg}_K(C_L)\ L = \text{trdeg}_K(C_M)\ M$. Furthermore, if $L/K$ and $M/K$ are compatible, then

(i) There is a difference $(\sigma)$-field $M_1$ isomorphic to $M$ and a set of constants $R$ such that $L(R) = M_1(R)$.
(ii) If $L$ is a PVE, then there is a specialization $b \to b'$ with $L = K\langle b' \rangle^*$.
(iii) If $L$ and $M$ are PVE of $K$, then $L$ and $M$ are $\sigma$-isomorphic over $K$.

As in the corresponding theory for the differential case, three types of extensions are used in constructing solution fields for linear homogeneous difference equations over
a difference field $K$: solution fields for difference equations $\alpha y = Ay$ or $\alpha y - y = B$ ($A, B \in K$), and algebraic extensions. The following is a brief account of this approach (the proofs can be found in [56]).

**Proposition 5.2.14.** Let $K$ be a difference field with a basic set $\sigma = \{\alpha\}, 0 \neq B \in K$, and let $a$ be a solution of the difference equation

$$\alpha y - y = B. \quad (5.2.2)$$

Then $M = K(a)$ is a corresponding solution field over $K$ with basis $b = (a, 1)$.

If there is no solution of equation (5.2.2) in $K$, then $a$ is transcendental over $K$, $K(a)$ has no new constants, and there are no intermediate $\sigma^*$-fields different from $K$ and $K(a)$.

If there is a solution $f \in K$, then $K(a)$ is an extension of $K$ generated by a constant which may be either transcendental or algebraic over $K$. If $a$ is transcendental, then $T_b$ is the set of matrices $\left( \begin{array}{cc} 1 & c \\ 0 & 1 \end{array} \right)$ where $c$ lies in the algebraic closure of $C_M$. If $C_M = C_K$, then the full Galois group of $M/K$ is isomorphic to the additive group of $C_K$.

**Proposition 5.2.15.** Let $K$ be as before, $a$ a non-zero solution of the difference equation

$$\alpha y - Ay = 0 \quad (5.2.3)$$

$(A \in K)$, and there is no non-zero solution in $K$ of the equation

$$\alpha y - A^n y = 0 \quad (5.2.4)$$

for $n \in \mathbb{N}, n > 0$. Then $a$ is transcendental over $K$ and $K(a)$ has no new constants.

If $L$ is an intermediate $\sigma^*$-field, then $L = K(a^n)$ for some $n \in \mathbb{N}$.

If equation (5.2.4) has a solution for some $n \in \mathbb{N}, n > 0$, then $K(a)$ is obtained from $K$ by an extension by a constant, which may be either transcendental or algebraic over $K$, followed by an algebraic extension. If $a$ is transcendental over $K$, the variety $T_a$ is the full set of all constants. If $C_{K(a)} = C_K$, then the full Galois group of $K(a)/K$ is a multiplicative subgroup of $C_K$.

**Definition 5.2.16.** Let $K$ be a difference field with a basic set $\sigma$ and $N$ a $\sigma^*$-overfield of $K$. $N/K$ is said to be a Liouvillian extension (LE) if there exists a chain

$$K = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_t = N, \quad K_{j+1} = K_j(a_j)^* (j = 0, \ldots, t - 1), \quad (5.2.5)$$

where $a_j$ is one of the following.

(a) A solution of an equation (5.2.2) where $B \in K_j$ and there is no solution of (5.2.2) in the field $K_j$.

(b) A solution of an equation (5.2.3) where $A \in K_j$ and for any $n \in \mathbb{N}, n > 0$, there is no non-zero solution of (5.2.4) in the field $K_j$.

(c) An algebraic element over $K_j$. 
More generally, \( N/K \) is said to be a generalized Liouvillian extension (GLE) if there exists a chain (5.2.5) where \( a_j \) is either a solution of (5.2.2) with \( B \in K_j \) or a solution of (5.2.3) with \( A \in K_j \), or an algebraic element over \( K_j \).

It follows from the definition that if a difference field \( K \) has an algebraically closed field of constants and a solution field \( M \) for a difference equation (5.2.1) is contained in a Liouvillian extension \( N \) of \( K \), then \( M \) is a PVE of \( K \).

The following results connect the solvability of a difference equation with the solvability of a matrix group.

**Theorem 5.2.17.** Let \( M \) be a solution field for the difference equation (5.2.1) over a difference (\( \sigma \)-) field \( K \). Let \( H \) be a connected group of automorphisms of \( M/K \) with matrix entries with respect to some basis \( b \) in an algebraically closed subfield of \( C_M \). (It need not be isomorphic to the set of matrices corresponding to \( H \).)

(i) If \( H \) is solvable, then \( M/H' \) is a GLE.

(ii) If \( H \) is reducible to diagonal form, then \( M/H' \) can be obtained by solving equations of the type (5.2.3).

(iii) If \( H \) is reducible to special triangular form, then \( M/H' \) can be obtained by solving equations of the type (5.2.2).

**Theorem 5.2.18.** Let \( K \) be a difference field with a basic set \( \sigma = \{ \alpha \} \) and \( M \) a \( \sigma^* \)-overfield of \( K \).

(i) If \( M/K \) is a PVE, then \( M/K \) is a GLE if and only if the component of identity of the Galois group is solvable.

(ii) If \( M \) is a solution field of a difference equation (5.2.1) contained in a GLE \( N/K \), then the component of the identity of \( \text{Gal}(M/K(C_M)) \) is solvable. Furthermore, if \( M/K(C_M) \) is a GPVE, then \( M/K \) is a GLE.

(iii) Suppose that \( M \) and \( L \) are solution fields of a difference equation (5.2.1) over \( K \). If \( L \) is contained in a GLE \( N \) of \( K \) and \( M/K \) is compatible with \( N/K \), then \( M \) is contained in a GLE of \( K \).

(iv) If \( N \) is a generic solution field for (5.2.1) over \( K \) and a solution field \( L \) for (5.2.1) is contained in a GLE of \( K \), then \( N \) is contained in a GLE of \( K \).

The following example indicates that it is not satisfactory to consider equation (5.2.1) to be “solvable by elementary operations” only if its solution field is contained in a GLE.

**Example 5.2.19.** With the notation of Theorem 5.2.18, suppose that \( K \) contains an element \( j \) with \( \alpha(j) \neq j \) and \( \alpha^2(j) = j \), and an element \( u \) with the following property. If \( u^k = v\alpha(v) \) or \( u^k = \frac{\alpha(v)}{v} \) for some \( v \in K, k \in \mathbb{N} \), then \( k = 0 \).

If \( \eta \) is any non-zero solution of the difference equation \( \alpha^2 y - uy = 0 \), then \( M = K(\eta)^* \) is a solution field for this equation with basis \( (\eta, \alpha(\eta)) \), \( \text{trdeg}_K M = 2 \), \( C_M = C_K \) and \( \text{Gal}_\sigma(M/K) \) is commutative. However, as it is shown in [57, Example 1], \( M \) is not a GLE of \( K \).

In what follows we consider some results by C. Franke (see [57] and [62]) that characterize the solvability of a difference equation of the form (5.2.1) “by elementary operations”.
Throughout the rest of the section $K$ denotes an inversive difference field with a basic set $\sigma = \{\alpha\}$. If $L$ is a $\sigma^*$-overfield of $K$, then $K_L$ will denote the algebraic closure of $K(C_L)$ in $L$.

**Definition 5.2.20.** Let $N$ be a $\sigma^*$-overfield of $K$, and $q$ a positive integer. A $q$-chain from $K$ to $N$ is a sequence of $\sigma^*$-fields $K = K_1 \subseteq K_1 \subseteq \cdots \subseteq K_t = N$, $K_{i+1} = K_i \langle \eta_i \rangle^*$ where $\eta_i$ is one of the following.

(a) Algebraic over $K$.
(b) A solution of an equation $\alpha^q y = y + B$ for some $B \in K_i$.
(c) A solution of an equation $\alpha^q y = Ay$ for some $A \in K_i$.

If there is a $q$-chain from $K$ to $N$, then $N$ is called a $q$LE of $K$.

Let $K_q$ denote the field $K$ treated as an inversive difference field with basic set $\sigma_q = \{\alpha^q\}$ and let $N(q)$ be a $\sigma^*_q$-overfield $N$ of $K$ treated as a $\sigma^*_q$-overfield of $K(q)$. In this case $N$ is a $q$LE of $K$ if and only if $N(q)$ is a GLE of $K(q)$ (see [57, Proposition 2.1]).

**Theorem 5.2.21.** If $M$ is a $\sigma$-normal $\sigma^*$-overfield of $K$ such that $K = K_M$ and the group $\text{Gal}_\sigma(M/K)$ is solvable, then $M$ is contained in a $q$LE of $K$. If, in addition, $M/K$ is a $\sigma^*$-field extension generated by a fundamental system of solutions of a difference equation (5.2.1), then $M$ is contained in a GLE of $K$.

**Theorem 5.2.22.** Let $N$ be a $q$LE of $K$ and $L$ an intermediate $\sigma^*$-field of $N/K$. Then $\text{Gal}_\sigma(L/K_L)$ is solvable.

Let $M$ be a $\sigma^*$-field extension of $K$. We say that difference equation (5.2.1) is solvable by elementary operations in $M$ over $K$ if $M$ is a solution field for (5.2.1) over $K$ and $M$ is contained in a $q$LE of $K$. This concept is independent of the solution field $M$, as follows from the second statement of the next theorem.

**Theorem 5.2.23.** Let a $\sigma^*$-field $M$ be a solution field for (5.2.1) over $K$.

(i) Equation (5.2.1) is solvable by elementary operations in $M$ over $K$ if and only if the group $\text{Gal}_\sigma(M/K)$ has a subnormal series whose factors are either finite or commutative.

(ii) If (5.2.1) is solvable by elementary operations in $M$ over $K$ and $N$ is another solution field for (5.2.1) over $K$, then (5.2.1) is solvable by elementary operations in $N$ over $K$. (This property allows one to say that (5.2.1) is solvable by elementary operations over $K$ if it is solvable by elementary operations in some solution field $M \supseteq K$.)

(iii) If (5.2.1) is solvable by elementary operations over $K$ and $L$ a $\sigma^*$-overfield of $K$, then (5.2.1) is solvable by elementary operations over $L$.

A number of results that specify the results of this section for the case of second-order difference equations were obtained in [56] and [57]. C. Franke, [56], also showed that the properties of having algebraically closed field of constants and having full sets of solutions of difference equations can be incompatible. Indeed, if $K$ is a difference field with basic set
\[\sigma = \{\alpha\} \text{ (Char } K \neq 2)\] such that \(C_K\) is algebraically closed, then the difference equation \(\alpha y + y = 0\) has no non-zero solution in \(K\). (If \(b\) is such a solution, then \(\alpha(b^2) = b^2\), so \(b^2 \in C_K\). Since \(C_K\) is algebraically closed, \(b \in C_K\) contradicting the fact that \(\alpha(b) = -b\).)

This observation and the fact that one could not associate a Picard–Vessiot-type extension to every difference equation have led to a different approach to the Galois theory of difference equations. This approach, based on the study of simple difference rings rather than difference fields, was realized by M. van der Put and M.F. Singer in their monograph [142]. In the next section we give an outline of the corresponding theory.

We conclude this section with one more theorem on the Galois correspondence for difference fields (see Theorem 5.2.27 below). This result is due to R. Infante who developed the theory of strongly normal difference field extensions (see [70–74]). Under some natural assumptions, the class of such extensions of a difference field \(K\) includes, in particular, the class of solution fields of linear homogeneous difference equations over \(K\).

Let \(K\) be an ordinary inversive difference field of zero characteristic with basic set \(\sigma = \{\alpha\}\). Let \(M\) be a finitely generated \(\sigma^*\)-overfield of \(K\) such that \(K\) is algebraically closed in \(M\) and \(C_M = C_K = C\). As above, \(K_M\) will denote the algebraic closure of \(K(C_M)\) in \(M\). Furthermore, for any \(\sigma\)-isomorphism \(\phi\) of \(M/K\) into a \(\sigma^*\)-overfield of \(M\), \(C_\phi\) will denote the field of constants of \(M\langle \phi M \rangle^*\).

**Definition 5.2.24.** With the above conventions, \(M\) is said to be a strongly normal extension of \(K\) if for every \(\sigma\)-isomorphism \(\phi\) of \(M/K\) into a \(\sigma^*\)-overfield of \(M\), \(M(C_\phi)^* = M(\phi M)^* = \phi M(C_\phi)^*\).

**Proposition 5.2.25.** Let \(M\) be a solution field of a difference equation of the form (5.2.1) over \(K\). Then \(M\) is a strongly normal extension of \(K_M\).

**Proposition 5.2.26.** Let \(M\) be a strongly normal \(\sigma^*\)-field extension of \(K\). Then

(i) \(\text{Id } M/K = 1\).

(ii) If \(\phi\) is any \(\sigma\)-isomorphism of \(M/K\) into a \(\sigma^*\)-overfield of \(M\), then \(C_\phi\) is a finitely generated extension of \(C\) and \(\text{trdeg}_M M(\phi M)^* = \text{trdeg}_C C_\phi\).

**Theorem 5.2.27.** If \(M\) is a strongly normal \(\sigma^*\)-field extension of \(K\), then there is a connected algebraic group \(G\) defined over \(C_M\) such that the connected algebraic subgroups of \(G\) are in one-to-one correspondence with the intermediate \(\alpha^*\)-fields of \(M/K\) algebraically closed in \(M\). Furthermore, there is a field of constants \(C'\) such that \(C'\)-rational points of \(G\) are all the \(\sigma\)-isomorphisms of \(M/K\) into \(M(C')^*\) and this set is dense in \(G\).

### 5.3. Picard–Vessiot rings and the Galois theory of difference equations

In this section we discuss some basic results of the Galois theory of difference equations based on the study of simple difference rings associated with such equations. The complete theory is presented in [142] where one can find the proofs of all statements of this section.

All difference rings and fields considered below are supposed to be ordinary and inversive. The basic set of a difference ring will be always denoted by \(\sigma\) and the only element
of $\sigma$ will be denoted by $\phi$ (we follow the notation of [142]). As usual, $GL_n(R)$ will denote the set of all non-singular $n \times n$-matrices over a ring $R$.

Let $R$ be a difference ring, $A \in GL_n(R)$, and $Y$ a column vector $(y_1, \ldots, y_n)^T$ whose coordinates are $\sigma^*$-indeterminates over $R$. (The corresponding ring of $\sigma^*$-polynomials is still denoted by $R[y_1, \ldots, y_n]^*$. ) In what follows, we will study systems of difference equations of the form $\phi Y = AY$ where $\phi Y = (\phi y_1, \ldots, \phi y_n)^T$. Notice that an $n$-th order linear difference equation $\phi^n y + \cdots + a_1 \phi y + a_0 y = 0$ ($a_0, a_1, \ldots \in R$ and $y$ is a $\sigma^*$-indeterminate over $R$) is equivalent to such a system with $y_i = \phi^{i-1} y$ ($i = 1, \ldots, n$) and

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix}.$$ 

Clearly, $A \in GL_n(R)$ if and only if $a_0 \neq 0$.

With the above notation, a fundamental matrix with entries in $R$ for $\phi Y = AY$ is a matrix $U \in GL_n(R)$ such that $\phi U = AU$ ($\phi U$ is the matrix obtained by applying $\phi$ to every entry of $U$). If $U$ and $V$ are fundamental matrices for $\phi Y = AY$, then $V = UM$ for some $M \in GL_n(C_R)$ since $U^{-1} V$ is left fixed by $\phi$. (As in Section 5.2, $C_R$ denotes the ring of constants of $R$, that is, $C_R = \{ a \in R \mid \phi a = a \}$.)

**Definition 5.3.1.** Let $K$ be a difference field. A $K$-algebra $R$ is called a Picard–Vessiot ring (PVR) for an equation

$$\phi Y = AY \quad (A \in GL_n(K)) \quad (5.3.1)$$

if it satisfies the following conditions.

(i) $R$ is a $\sigma^*$-$K$-algebra (as usual, the automorphism of $R$ which extends $\phi$ is denoted by the same letter).

(ii) $R$ is a simple difference ring, that is, the only difference ideals of $R$ are (0) and $R$.

(iii) There exists a fundamental matrix for $\phi Y = AY$ with entries in $R$.

(iv) $R$ is minimal in the sense that no proper subalgebra $R$ satisfies (i)–(iii).

**Example 5.3.2** (see [142, Examples 1.3 and 1.6]). Let $C$ be an algebraically closed field, $\text{Char} C \neq 2$. Let us define an equivalence relation on the set of all sequences $a = (a_0, a_1, \ldots)$ of elements of $C$ by saying that $a$ is equivalent to $b = (b_0, b_1, \ldots)$ if there exists $N \in \mathbb{N}$ such that $a_n = b_n$ for all $n > N$. With coordinatewise addition and multiplication, the set of all equivalence classes forms a ring $S$. This ring can be treated as a difference ring with respect to its automorphism $\phi$ that maps each fundamental class of $(a_0, a_1, \ldots)$ to the equivalent class of $(a_1, a_2, \ldots)$. (It is easy to check that $\phi$ is well-defined.) To simplify notation we shall identify a sequence $a$ with its equivalence class.

Let $R$ be the difference subring of $S$ generated by $C$ and $j = (1, -1, 1, -1, \ldots)$, that is, $R = C[j]^*$. The $1 \times 1$-matrix whose only entry is $j$ is the fundamental matrix of the equation $\phi y = -y$. This ring is isomorphic to $C[X]/(X^2 - 1)$ ($C[X]$ is the polynomial ring in one indeterminate $X$ over $C$) whose only non-trivial ideals are generated by the cosets...
of $X - 1$ and $X + 1$. Since the ideals generated in $R$ by $j + 1$ and $j - 1$ are not difference ideals, $R$ is a simple difference ring. Therefore, $R$ is a PVR for $\phi y = -y$ over $C$. Note that $R$ is reduced but it is not an integral domain.

**Proposition 5.3.3.** Let $K$ be a difference $(\sigma^*-\tau)$ field with an algebraically closed field of constants $C_K$.

(i) If a $\sigma^*-K$-algebra $R$ is a simple difference ring finitely generated as a $K$-algebra, then $C_R = C_K$.

(ii) If $R_1$ and $R_2$ are two PVR’s for a difference equation (5.3.1), then there exists a $\sigma$-isomorphism between $R_1$ and $R_2$ that leaves the field $K$ fixed.

To form a PVR for a difference equation (5.3.1) one can use the following procedure suggested in [142, Chapter 1]. Let $(X_{ij})$ denote an $n \times n$-matrix of indeterminates over $K$ and $\det$ denote the determinant of this matrix. Then one can extend $\phi$ to an automorphism of the $K$-algebra $K[X_{ij}, \det^{-1}]$ by setting $(\phi X_{ij}) = A(X_{ij})$. If $I$ is a maximal difference ideal of $K[X_{ij}, \det^{-1}]$ then $K[X_{ij}, \det]/I$ is a PVR for (5.3.1), it satisfies all conditions of Definition 5.3.1. (It is easy to see that $I$ is a radical $\sigma^*$-ideal and $K[X_{ij}, \det]/I$ is a reduced prime difference ring.) Moreover, any PVR for difference equation (5.3.1) will be of this form.

Let $\overline{K}$ denote the algebraic closure of $K$ and let $D = \overline{K}[X_{ij}, \det^{-1}]$. Then the automorphism $\phi$ extends to an automorphism of $\overline{K}$ which, in turn, extends to an automorphism of $D$ such that $(\phi X_{ij}) = A(X_{ij})$ (the extensions of $\phi$ are also denoted by $\phi$). It is easy to see that every maximal ideal $M$ of $D$ has the form $(X_{11} - b_{11}, \ldots, X_{nn} - b_{nn})$ and corresponds to a matrix $B = (b_{ij}) \in GL_n(\overline{K})$. Then $\phi(M)$ is a maximal ideal of $D$ that corresponds to the matrix $A^{-1}\phi(B)$ where $\phi(B) = (\phi(b_{ij}))$. Thus, the action of $\phi$ on $D$ induces a map $\tau$ on $GL_n(\overline{K})$ such that $\tau(B) = A^{-1}\phi(B)$. The elements $f \in D$ are seen as functions on $GL_n(\overline{K})$. For any $f \in D, B \in GL_n(\overline{K})$, we have $(\phi f)(\tau(B)) = \phi(f(B))$. Furthermore, if $J$ is an ideal of $K[X_{ij}, \det^{-1}]$ such that $\phi(J) \subseteq J$, then $\phi(J) = J$. Also, for reduced algebraic subsets $Z$ of $GL_n(K)$, the condition $\tau(Z) \subseteq Z$ implies $\tau(Z) = Z$.

**Proposition 5.3.4** [142, Lemma 1.10]. The ideal $J$ of a reduced algebraic subset $Z$ of $GL_n(K)$ satisfies $\phi(J) = J$ if and only if $Z(\overline{K})$ satisfies $\tau Z(\overline{K}) = Z(\overline{K})$.

An ideal $I$ maximal among the $\phi$-invariant ideals corresponds to a minimal (reduced) algebraic subset $Z$ of $GL_n(K)$ such that $\tau Z(\overline{K}) = Z(\overline{K})$. Such a set is called a minimal $\tau$-invariant reduced set.

Let $Z$ be a minimal $\tau$-invariant reduced subset of $GL_n(K)$ with an ideal $I \subseteq K[X_{ij}, \det^{-1}]$ and let $O(Z) = K[X_{ij}, \det^{-1}]/I$. Let us denote the image of $X_{ij}$ in $O(Z)$ by $x_{ij}$ and consider the rings

$$K \left[ X_{ij}, \frac{1}{\det} \right] \subseteq O(Z) \otimes_K K \left[ X_{ij}, \frac{1}{\det(X_{ij})} \right]$$

$$= O(Z) \otimes_C C \left[ Y_{ij}, \frac{1}{\det(Y_{ij})} \right] \subseteq C \left[ Y_{ij}, \frac{1}{\det(Y_{ij})} \right].$$

(5.3.2)
Let \((I)\) denote the ideal of \(O(Z) \otimes K \{X_{ij}, \det^{-1}\}\) generated by \(I\) and let \(J = (I) \cap C\{Y_{ij}, \det^{-1}\}\). The ideal \((I)\) is \(\phi\)-invariant, the set of constants of \(O(Z)\) is \(C\), and \(J\) generates the ideal \((I)\) in \(O(Z) \otimes K \{X_{ij}, \det^{-1}\}\). Furthermore, one has natural mappings

\[
O(Z) \rightarrow O(Z) \otimes K O(Z) = O(Z) \otimes C \left( C \left[ Y_{ij}, \frac{1}{\det(Y_{ij})} \right] / J \right) \leftarrow C \left[ Y_{ij}, \frac{1}{\det(Y_{ij})} \right] / J. \tag{5.3.3}
\]

Suppose that \(O(Z)\) is a separable extension of \(K\) (for example, \(\text{Char } K = 0\) or \(K\) is perfect). One can show (see [142, Section 1.2]) that \(O(Z) \otimes K O(Z)\) is reduced. Therefore, \(C\{Y_{ij}, \det(Y_{ij})^{-1}\}/J\) is reduced and \(J\) is a radical ideal. Furthermore, the following considerations imply that \(J\) is the ideal of an algebraic subgroup of \(GL_n(C)\).

Let \(A \in GL_n(C)\) and let \(\delta_A\) denote the action on the terms of (5.3.2) defined by \((\delta_A X_{ij}) = X_{ij}A\) and \((\delta_A Y_{ij}) = Y_{ij}A\). Then the following eight properties are equivalent:

1. \(ZA = Z\);
2. \(ZA \cap Z \neq \emptyset\);
3. \(\delta_A I = I\);
4. \(I + \delta_A I\) is not the unit ideal of \(K\{X_{ij}, \det^{-1}\}\);
5. \(\delta_A (I) = (I)\);
6. \((I) + \delta_A (I)\) is not the unit ideal of \(O(Z) \otimes K \{X_{ij}, \det^{-1}\}\);
7. \(\delta_A J = J\);
8. \(J + \delta_A J\) is not the unit ideal of \(O(Z) \otimes C \{Y_{ij}, \det^{-1}\}\).

The set of all matrices \(A \in GL_n(C)\) satisfying the equivalent conditions (1)–(8) form a group.

**Proposition 5.3.5** (see [142, Lemma 1.12]). Let \(O(Z)\) be a separable extension of \(K\). With the above notation, \(A\) satisfies the equivalent conditions (1)–(8) if and only if \(A\) lies in the reduced subspace \(V\) of \(GL_n(C)\) defined by \(J\). Therefore, the set of such \(A\) is an algebraic group.

Let \(G\) denote the group of all automorphisms of \(O(Z)\) over \(K\) which commute with the action of \(\phi\). The group \(G\) is called the **difference Galois group** of the equation \(\phi(Y) = AY\) over the field \(K\).

If \(\delta \in G\), then \((\delta x_{ij}) = (x_{ij}) A\) where \(A \in GL_n(C)\) is such that \(\delta_A\) (as defined above) satisfies \(\delta_A I = I\). Therefore, one can identify \(G\) and the subspace \(V\) from the last proposition. Denoting the ring \(C\{Y_{ij}, \det^{-1}\}/J\) by \(O(G)\) and setting \(O(G_k) = O(G) \otimes C k, G_K = \text{spec}(O(G_k))\), one can use (5.3.3) to obtain the sequence

\[
O(Z) \rightarrow O(Z) \otimes K O(Z) = O(Z) \otimes C O(G) = O(Z) \otimes K O(G_K). \tag{5.3.4}
\]

The first embedding of rings corresponds to the morphism \(Z \times G_K \rightarrow Z\) given by \((z, g) \mapsto zg\). The identification \(O(Z) \otimes K O(Z) = O(Z) \otimes C O(G) = O(Z) \otimes K O(G_K)\)
corresponds to the fact that the morphism \( Z \times G_K \to Z \times Z \) given by \((z, g) \mapsto (zg, z)\) is an isomorphism. Thus, \( Z \) is a \( K \)-homogeneous space for \( G_K \), that is \( Z/K \) is a \( G \)-torsor.

The following result (proved in [142, Section 1.2]) shows that a PVR is the coordinate ring of a torsor for its difference Galois group.

**Theorem 5.3.6.** Let \( R \) be a separable PVR over \( K \), a difference field with an algebraically closed field of constants \( C \). Let \( G \) denote the group of the \( K \)-algebra automorphisms of \( R \) which commute with \( \phi \). Then

(i) \( G \) has a natural structure as reduced linear algebraic group over \( C \) and the affine scheme \( Z \) over \( C \) has the structure of a \( G \)-torsor over \( K \).

(ii) The set of \( G \)-invariant elements of \( R \) is \( K \) and \( R \) has no proper, non-trivial \( G \)-invariant ideals.

(iii) There exist idempotents \( e_0, \ldots, e_{t-1} \in R \) (\( t \geq 1 \)) such that

(a) \( R = R_0 \oplus \cdots \oplus R_{t-1} \) where \( R_i = e_i R \) for \( i = 0, \ldots, t - 1 \).

(b) \( \phi(e_i) = e_{i+1} \) (mod \( t \)) and so \( \phi \) maps \( R_i \) isomorphically onto \( R_{i+1} \) (mod \( t \)) and \( \phi^t \) leaves each \( R_i \) invariant.

(c) For each \( i \), \( R_i \) is a domain and is a Picard–Vessiot extension of \( e_i K \) with respect to \( \phi^t \).

Let \( K \) be a difference field with an algebraically closed field of constants \( C \) and \( R \) a PVR for an equation \( \phi(Y) = AY \) over \( K \). Let \( \delta = \delta_A \) and let \( R = R_0 \oplus \cdots \oplus R_{t-1} \) (\( R_i = e_i R \) for \( i = 0, \ldots, t - 1 \)) be as in the last theorem. Then \( \delta : R_i \to R_{i+1} \) is an isomorphism and \( R_0 \) is a PVR over \( K \) with respect to the automorphism \( \delta^t \). Let us define two mappings \( \Gamma : \text{Gal}(R_0/K) \to \text{Gal}(R/K) \) and \( \Delta : \text{Gal}(R/K) \to \mathbb{Z}/t\mathbb{Z} \) as follows. For any \( \psi \in \text{Gal}(R_0/K) \), we set \( \Gamma(\psi) = \chi \) where for \( r = (r_0, \ldots, r_{t-1}) \in R \), \( \chi(r_0, \ldots, r_{t-1}) = (\psi(r_0), \delta^t\psi^t(\delta r_1), \ldots, \delta^t\psi^t(\delta^{t-1}r_{t-1})) \). In order to define \( \Delta \), notice that if \( \chi \in \text{Gal}(R/K) \), then \( \chi \) permutes with each \( e_i \). If \( \chi(e_0) = e_j \), we define \( \Delta(\chi) = j \).

**Proposition 5.3.7.** Let \( R \) be a separable PVR over \( K \), a difference field with an algebraically closed field of constants \( C \).

(i) With the above notation, we have the exact sequence \( 0 \to \text{Gal}(R_0/K) \xrightarrow{\Gamma} \text{Gal}(R/K) \xrightarrow{\Delta} \mathbb{Z}/t\mathbb{Z} \to 0 \).

(ii) Let \( G \) denote the difference Galois group of \( R \) over \( K \). If \( H^1(\text{Gal}(\overline{K}/K), G(\overline{K})) = 0 \), then \( Z = \text{spec}(R) \) is \( G \)-isomorphic to the \( G \)-torsor \( G_K \) and so \( R = C[G] \otimes K \).

In what follows we present a characterization of the difference Galois group of a PVR over the field of rational functions \( C(z) \) in one variable \( z \) over an algebraically closed field \( C \) of zero characteristic (one can assume \( C = C \)). We fix \( a \in C(z) \), \( a \neq 0 \), and consider \( C(z) \) as an ordinary difference field with the basic automorphism \( \phi_a : z \mapsto z + a \) (\( \phi_a \) leaves the field \( C \) fixed). This difference field will be denoted by \( K \). Note that \( \phi_a \) does not extend to any proper finite field extension of \( K \) (see [142, Lemma 1.19]).

**Theorem 5.3.8.** Let \( K = C(z) \) be as above and let \( G \) be an algebraic subgroup of \( \text{GL}_n(C) \). Let \( \phi(Y) = AY \) be a system of difference equations with \( A \in G(K) \). Then
(i) The Galois group of \( \phi(Y) = AY \) over \( K \) is a subgroup of \( G_C \).

(ii) Any minimal element in the set of \( C \)-subgroups \( H \) of \( G \) for which there exists a \( B \in Gl_n(K) \) with \( B^{-1}A^{-1}\phi(B) \in H(K) \) is the difference Galois group of \( \phi(Y) = AY \) over \( K \).

(iii) The difference Galois group of \( \phi(Y) = AY \) over \( K \) is \( G \) if and only if for any \( B \in G(K) \) and any proper \( C \)-subgroup \( H \) of \( G \), one has \( B^{-1}A^{-1}\phi(B) \notin H(K) \).

**Definition 5.3.9.** Let \( K \) be an ordinary difference field with a basic set \( \sigma = \{ \phi \} \) and let \( A \in Gl_n(K) \). A difference overring \( L \) of \( K \) is said to be the total Picard–Vessiot ring (TPVR) of the equation \( \phi(Y) = AY \) over \( K \) if \( L \) is the total ring of fractions of the PVR \( R \) of the equation.

As we have seen, a PVR \( R \) is a direct sum of domains: \( R = R_0 \oplus \cdots \oplus R_{t-1} \) where each \( R_i \) is invariant under the action of \( \phi^i \). The automorphism \( \phi \) of \( R \) permutes \( R_0, \ldots, R_{t-1} \) in a cyclic way (that is, \( \phi(R_i) = R_{i+1} \) for \( i = 1, \ldots, t-2 \) and \( \phi(R_{t-1}) = R_0 \)). It follows that the TPVR \( L \) is the direct sum of fields: \( L = L_0 \oplus \cdots \oplus L_{t-1} \) where each \( L_i \) is the field of fractions of \( R_i \), and \( \phi \) permutes \( L_0, \ldots, L_{t-1} \) in a cyclic way.

**Proposition 5.3.10.** With the above notation, let \( K \) be a perfect difference field with an algebraically closed field of constants \( C \) and let \( \phi(Y) = BY \) be a difference equation over \( K (B \in Gl_n(K)) \). Let a difference ring extension \( K' \supseteq K \) have the following properties:

(i) \( K' \) has no nilpotent elements and every non-zero divisor of \( K' \) is invertible.

(ii) The set of constants of \( K' \) is \( C \).

(iii) There is a fundamental matrix \( F \) for the equation with entries in \( K' \).

(iv) \( K' \) is minimal with respect to (i), (ii), and (iii).

Then \( K' \) is \( K \)-isomorphic as a difference ring to the TPVR of the equation.

**Corollary 5.3.11.** Let \( K \) be as in the last proposition and let \( \phi(Y) = AY \) be a difference equation over \( K (A \in Gl_n(K)) \). Let a difference overring \( R \supseteq K \) have the following properties:

(i) \( R \) has no nilpotent elements.

(ii) The set of constants of the total quotient ring of \( R \) is \( C \).

(iii) There is a fundamental matrix \( F \) for the equation with entries in \( R \).

(iv) \( R \) is minimal with respect to (i), (ii), and (iii).

Then \( R \) is a PVR of the equation.

With the above notation, let \( R = R_0 \oplus \cdots \oplus R_{t-1} \) be the PVR of the equation \( \phi(Y) = AY \) \( (A \in Gl_n(K)) \). Let us consider the difference field \( (K, \phi^t) \) (that is, the field \( K \) treated as a difference field with the basic set \( \sigma_t = \{ \phi^t \} \)) and the difference equation \( \phi^t(Y) = A_t Y \) with \( A_t = \phi^{t-1}(A) \cdots \phi^2(A) \phi(A) A \).

**Proposition 5.3.12.**

(i) Each component \( R_i \) of \( R \) is a PVR for the equation \( \phi^i(Y) = A_t Y \) over the difference field \( (K, \phi^i) \).
(ii) Let $d \geq 1$ be a divisor of $t$. Using cyclic notation for the indices $\{0, \ldots, t - 1\}$, we consider the subrings $\bigoplus_{m=0}^{(t/d)-1} R_{t+md}$ of $R = R_0 \oplus \cdots \oplus R_{t-1}$. Then each of these subrings is a PVR for the equation $φ^d(Y) = A_dY$ over the difference field $(K, φ^d)$.

**Proposition 5.3.13.** Let $L$ be the TPVR of an equation $φ(Y) = AY$ ($A ∈ GL_n(K)$) over a perfect difference field $K$ whose field of constants $C = C_K$ is algebraically closed. Let $G$ denote the difference Galois group of the equation and let $H$ be an algebraic subgroup of $G$. Then $G$ acts on $L$ and moreover:

(i) $L^G$, the set of $G$-invariant elements of $L$, is equal to $K$.

(ii) If $L^H = K$, then $H = G$.

The following result describes the Galois correspondence for total Picard–Vessiot rings. As is noticed in [142, Section 1.3], one cannot expect a similar theorem for Picard–Vessiot rings. Indeed, let $K$ be as in the last proposition and $R = K ⊗_C C[G]$ where $C[G]$ is the ring of regular functions on an algebraic group $G$ defined over $C$. For an algebraic subgroup $H$ of $G$, the ring of invariants $R^H$ is the ring of regular functions on $(G/H)_K$.

In some cases, e.g., $G = GL_n(C)$ and $H$ a Borel subgroup, the space $G/H$ is a connected projective variety and so the ring of regular functions on $(G/H)_K$ is just $K$.

**Theorem 5.3.14.** Let $K$ be a difference field of zero characteristic with a basic set $σ = \{φ\}$. Let $A ∈ GL_n(K)$ and let $L$ be a TVPR of the equation $φ(Y) = AY$ over $K$. Let $F$ denote the set of intermediate difference rings $F$ such that $K ≤ F ≤ L$ and every non-zero divisor of $F$ is a unit of $F$. Furthermore, let $G$ denote the set of algebraic subgroups of $G$.

(i) For any $F ∈ F$, the subgroup $G(L/F) ⊆ G$ of the elements of $G$ which fix $F$ point-wise, is an algebraic subgroup of $G$.

(ii) For any algebraic subgroup $H$ of $G$, the ring $L^H$ belongs to $F$.

(iii) Let $α : F → G$ and $β : G → F$ denote the maps $F → G(L/F)$ and $H → L^H$, respectively. Then $α$ and $β$ are each others inverses.

**Corollary 5.3.15.** With the notation of Theorem 5.3.14, a group $H ∈ G$ is a normal subgroup of $G$ if and only if the difference ring $F = L^H$ has the property that for every $z ∈ F \setminus K$, there is an automorphism $δ$ of $F/K$ which commutes with $φ$ and satisfies $δz ≠ z$. If $H ∈ G$ is normal, then the group of all automorphisms $δ$ of $F/K$ which commute with $φ$ is isomorphic to $G/H$.

**Corollary 5.3.16.** With the above notation, suppose that an algebraic group $H ⊆ G$ contains $G^0$, the component of the identity of $G$. Then the difference ring $R^H$ ($R$ is a PVR for the equation $φ(Y) = AY$ over $K$) is a finite dimension vector space over $K$ with dimension equal to $G : H$.

A number of applications of the above-mentioned results on ring-theoretical difference Galois theory to various types of algebraic difference equations can be found in [64–67] and [142]. Using the technique of difference Galois groups, the monograph [142] also develops the analytic theory of ordinary difference equations over the fields $C(z)$ and $C(\{z^{-1}\})$. 
We conclude this section with a fundamental result on the inverse problem of ring-theoretical difference Galois theory.

**Theorem 5.3.17** [142, Theorem 3.1]. Let $K = C(z)$ be the field of fractions of one variable $z$ over an algebraically closed field $C$ of zero characteristic. Consider $K$ as an ordinary difference field with respect to the automorphism $\phi$ that leaves the field $C$ fixed and maps $z$ to $z + 1$. Then any connected algebraic subgroup $G$ of $\text{Gl}_n(C)$ is the difference Galois group of a difference equation $\phi(Y) = AY, A \in \text{Gl}_n(K)$.

**References**


Difference algebra


Section 5A
Groups and Semigroups
Reflection Groups

Meinolf Geck
Department of Mathematical Sciences, King’s College, Aberdeen University, Aberdeen AB24 3UE, Scotland, UK
E-mail: geck@maths.abdn.ac.uk

Gunter Malle
FB Mathematik, Universität Kaiserslautern, Postfach 3049, D-67653 Kaiserslautern, Germany
E-mail: malle@mathematik.uni-kl.de

Contents
1. Finite groups generated by reflections .............................. 340
  1.2. Invariants .................................................. 340
  1.5. Parabolic subgroups ........................................ 342
  1.7. Exponents, coexponents and fake degrees ...................... 342
  1.8. Reflection data .............................................. 343
  1.9. Regular elements ............................................ 344
  1.13. The Shephard–Todd classification ........................... 345
2. Real reflection groups ........................................... 349
  2.1. Coxeter groups ............................................... 349
  2.3. Cartan matrices ............................................... 351
  2.5. Classification of finite Coxeter groups ....................... 352
  2.8. Conjugacy classes and the length function ................... 353
  2.10. The crystallographic condition ................................ 354
  2.11. Root systems ................................................ 355
  2.13. Torsion primes ............................................... 357
  2.14. Affine Weyl groups ......................................... 358
  2.15. Kac–Moody algebras ........................................ 359
  2.16. Groups with a BN-pair ....................................... 360
  2.17. Connected reductive algebraic groups ....................... 361
3. Braid groups ..................................................... 362
  3.1. The braid group of a complex reflection group ............... 362
  3.4. The centre and regular elements ................................ 364
  3.7. Knots and links, Alexander and Markov theorem ............... 366
  3.9. The HOMFLY-PT polynomial .................................. 367
  3.10. Further aspects of braid groups .............................. 368
4. Representation theory ............................................. 368
   4.1. Fields of definition .......................................... 368
   4.3. Macdonald–Lusztig–Spaltenstein induction .................. 369
   4.5. Irreducible characters ....................................... 369
   4.8. Fake degrees .................................................. 372
   4.11. Modular representations of $S_n$ ........................... 373
5. Hints for further reading ......................................... 373
   5.1. Crystallographic reflection groups .......................... 374
   5.2. Quaternionic reflection groups ................................ 374
   5.3. Reflection groups over finite fields .......................... 374
   5.6. $p$-adic reflection groups .................................... 375
   5.8. $p$-compact groups ........................................... 376
References ............................................................... 377
This chapter is concerned with the theory of finite reflection groups, that is, finite groups generated by reflections in a real or complex vector space. This is a rich theory, both for intrinsic reasons and as far as applications in other mathematical areas or mathematical physics are concerned. The origin of the theory can be traced back to the ancient study of symmetries of regular polyhedra. Another extremely important impetus comes from the theory of semisimple Lie algebras and Lie groups, where finite reflection groups occur as “Weyl groups”. In the last decade, Broué’s “Abelian defect group conjecture” (a conjecture concerning the representations of finite groups over fields of positive characteristic) has lead to a vast research program, in which complex reflection groups, corresponding braid groups and Hecke algebras play a prominent role. Thus, the theory of reflection groups is at the same time a well-established classical piece of mathematics and still a very active research area. The aim of this chapter (and a subsequent one on Hecke algebras) is to give an overview of both these aspects.

As far as the study of reflection groups as such is concerned, there are (at least) three reasons why this leads to an interesting and rich theory:

Classification. Given a suitable notion of “irreducible” reflection groups, it is possible to give a complete classification, with typically several infinite families of groups and a certain number of exceptional cases. In fact, this classification can be seen as the simplest possible model for much more complex classification results concerning related algebraic structures, such as complex semisimple Lie algebras, simple algebraic groups and, eventually, finite simple groups. Besides the independent interest of such a classification, we mention that there is a certain number of results on finite reflection groups which can be stated in general terms but whose proof requires a case-by-case analysis according to the classification. (For example, the fact that every element in a finite real reflection group is conjugate to its inverse.)

Presentations. Reflection groups have a highly symmetric “Coxeter type presentation” with generators and defining relations (visualised by “Dynkin diagrams” or generalisations thereof), which makes it possible to study them by purely combinatorial methods (length function, reduced expressions and so on). From this point of view, the associated Hecke algebras can be seen as “deformations” of the group algebras of finite reflection groups, where one or several formal parameters are introduced into the set of defining relations. One of the most important developments in this direction is the discovery of the Kazhdan–Lusztig polynomials and the whole theory coming with them. (This is discussed in more detail in the chapter on Hecke algebras.)

Topology and geometry. The action of a reflection group on the underlying vector space opens the possibility of using geometric methods. First of all, the ring of invariant symmetric functions on that vector space always is a polynomial ring (and this characterises finite reflection groups). Furthermore, we have a corresponding hyperplane arrangement which gives rise to the definition of an associated braid group as the fundamental group of a certain topological space. For the symmetric group, we obtain in this way the classical Artin braid group, with applications in the theory of knots and links.
Furthermore, all these aspects are related to each other which – despite being quite elementary taken individually – eventually leads to a highly sophisticated theory.

We have divided our survey into four major parts. The first part deals with finite complex or real reflection groups in general. The second part deals with finite real reflection groups and the relations with the theory of Coxeter groups. The third part is concerned with the associated braid groups. Finally, in the fourth part, we consider complex irreducible characters of finite reflection groups.

We certainly do not pretend to give a complete picture of all aspects of the theory of reflection groups and Coxeter groups. Our references cover the period up until 2003. Especially, we will not say so much about areas that we do not feel competent in; to our best knowledge, we try to give at least some references for further reading in such cases. This concerns, in particular, all aspects of infinite (affine, hyperbolic, ...) Coxeter groups.

1. Finite groups generated by reflections

1.1. Definitions. Let $V$ be a finite-dimensional vector space over a field $K$. A reflection on $V$ is a non-trivial element $g \in \text{GL}(V)$ of finite order which fixes a hyperplane in $V$ pointwise. There are two types of reflections, according to whether $g$ is semisimple (hence diagonalisable) or unipotent. Often, the term reflection is reserved for the first type of elements, while the second are called transvections. They can only occur in positive characteristic. Here, we will almost exclusively be concerned with ground fields $K$ of characteristic 0, which we may and will then assume to be subfields of the field $\mathbb{C}$ of complex numbers. Then, by our definition, reflections are always semisimple and (thus) diagonalisable. Over fields $K$ contained in the field $\mathbb{R}$ of real numbers, reflections necessarily have order 2, which is the case motivating their name. Some authors reserve the term reflection for this case, and speak of pseudo-reflections in the case of arbitrary (finite) order.

Let $g \in \text{GL}(V)$ be a reflection. The hyperplane $C_V(g)$ fixed point-wise by $g$ is called the reflecting hyperplane of $g$. Then $V = C_V(g) \oplus V_g$ for a unique $g$-invariant subspace $V_g$ of $V$ of dimension 1. Any non-zero vector $v \in V_g$ is called a root for $g$. Thus, a root for a reflection is an eigenvector with eigenvalue different from 1. Now assume in addition that $V$ is Hermitean. Then conversely, given a vector $v \neq 0$ in $V$ and a natural number $n \geq 2$ we may define a reflection in $V$ with root $v$ and of order $n$ by $g.v := \exp(2\pi i/n)v$, and $g|_{V^\perp} = \text{id}$.

A reflection group on $V$ is now a finite subgroup $W \leq \text{GL}(V)$ generated by reflections. Note that any finite subgroup of $\text{GL}(V)$ leaves invariant a non-degenerate Hermitean form. Thus, there is no loss in assuming that a reflection group $W$ leaves such a form invariant.

1.2. Invariants

Let $V$ be a finite-dimensional vector space over $K \subseteq \mathbb{C}$. Let $K[V]$ denote the algebra of symmetric functions on $V$, i.e., the symmetric algebra $S(V^*)$ of the dual space $V^*$ of $V$. So $K[V]$ is a commutative algebra over $K$ with a grading $K[V] = \bigoplus_{d \geq 0} K[V]^d$, where, for any $d \geq 0$, $K[V]^d$ denotes the $d$-th symmetric power of $V^*$. If $W \leq \text{GL}(V)$ then $W$
Reflection groups

acts naturally on $K[V]$, respecting the grading. Now reflection groups are characterised by the structure of their invariant ring $K[V]^W$:

1.3. THEOREM (Shephard and Todd, [167], Chevalley, [46]). Let $V$ be a finite-dimensional vector space over a field of characteristic 0 and $W \leq \text{GL}(V)$ a finite group. Then the following are equivalent:

(i) the ring of invariants $K[V]^W$ is a polynomial ring,
(ii) $W$ is generated by reflections.

The implication from (i) to (ii) is an easy consequence of Molien’s formula

$$P(K[V]^W, x) = \frac{1}{|W|} \sum_{g \in W} \frac{1}{\det_V(1 - gx)}$$

for the Hilbert series $P(K[V]^W, x)$ of the ring of invariants $K[V]^W$, see Shephard and Todd, [167, p. 289]. It follows from Auslander’s purity of the branch locus that this implication remains true in arbitrary characteristic (see Benson, [7, Theorem 7.2.1], for example). The other direction was proved by Shephard and Todd as an application of their classification of complex reflection groups (see Section 1.13). Chevalley gave a general proof avoiding the classification which uses the combinatorics of differential operators.

Let $W$ be an $n$-dimensional reflection group. By Theorem 1.3 the ring of invariants is generated by $n$ algebraically independent polynomials (so-called basic invariants), which may be taken to be homogeneous. Although these polynomials are not uniquely determined in general, their degrees $d_1 \leq \cdots \leq d_n$ are. They are called the degrees of $W$. Then $|W| = d_1 \cdots d_n$, and the Molien formula shows that $N(W) := \sum_{i=1}^n m_i$ is the number of reflections in $W$, where $m_i := d_i - 1$ are the exponents of $W$.

The quotient $K[V]_W$ of $K[V]$ by the ideal generated by the invariants of strictly positive degree is called the coinvariant algebra of $(V, W)$. This is again a naturally graded $W$-module, whose structure is described by:

1.4. THEOREM (Chevalley, [46]). Let $V$ be a finite-dimensional vector space over a field $K$ of characteristic 0 and $W \leq \text{GL}(V)$ a reflection group. Then $K[V]^W$ carries a graded version of the regular representation of $W$. The grading is such that

$$\sum_{i \geq 0} \dim K[V]^W i x^i = \prod_{i=1}^n x^{d_i} - 1, \quad \frac{x-1}{x-1},$$

where $K[V]^W_i$ denotes the homogeneous component of degree $i$.

(See also Bourbaki, [25, V.5.2, Theorem 2].) The polynomial

$$P_W := \sum_{i \geq 0} \dim K[V]^W_i x^i$$

is called the Poincaré-polynomial of $W$. 
1.5. Parabolic subgroups

Let $W$ be a reflection group on $V$. The parabolic subgroups of $W$ are by definition the pointwise stabilisers

$$W_{V'} := \{ g \in W \mid g \cdot v = v \text{ for all } v \in V' \}.$$

of subspaces $V' \subseteq V$. The following result is of big importance in the theory of reflection groups:

1.6. Theorem (Steinberg, [175]). Let $W \leq \text{GL}(V)$ be a complex reflection group. For any subspace $V' \subseteq V$ the parabolic subgroup $W_{V'}$ is generated by the reflections it contains, that is, by the reflections whose reflecting hyperplane contains $V'$. In particular, parabolic subgroups are themselves reflection groups.

For the proof, Steinberg characterises reflection groups via eigenfunctions of differential operators with constant coefficients that are invariant under finite linear groups. Lehrer [129] has recently found an elementary proof. For a generalisation to positive characteristic see Theorem 5.4.

1.7. Exponents, coexponents and fake degrees

Let $W$ be a complex reflection group. For $w \in W$ define $k(w) := \dim V^{(w)}$, the dimension of the fixed space of $w$ on $V$. Solomon, [169], proved the following remarkable formula for the generating function of $k$

$$\sum_{w \in W} x^{k(w)} = \prod_{i=1}^{n} (x + m_i),$$

by showing that the algebra of $W$-invariant differential forms with polynomial coefficients is an exterior algebra of rank $n$ over the algebra $K[V]^W$, generated by the differentials of a set of basic invariants (see also Flatto, [79], Benson, [7, Theorem 7.3.1]). The formula was first observed by Shephard and Todd, [167, 5.3] using their classification of irreducible complex reflection groups. Dually, Orlik and Solomon, [154], showed

$$\sum_{w \in W} \det_V(x)^{k(w)} = \prod_{i=1}^{n} (x - m_i^*)$$

for some non-negative integers $m_1^* \leq \cdots \leq m_n^*$, the coexponents of $W$ (see Lehrer and Michel, [130], for a generalisation, and Kusuoka, [123], Orlik and Solomon, [155,156], for versions over finite fields).

Let $\chi$ be an irreducible character of $W$. The fake degree of $\chi$ is the polynomial

$$R_\chi := \sum_{d \geq 0} [K[V_d^W], \chi] \chi_d = \frac{1}{|W|} \sum_{w \in W} \frac{\chi(w)}{\det_V(x w - 1)^*} \prod_{i=1}^{n} (x_{d_i} - 1) \in \mathbb{Z}[x],$$
that is, the graded multiplicity of $\chi^*$ in the $W$-module $K[V]_W$. Thus, in particular, $R_\chi$ specialises to the degree $\chi(1)$ at $x = 1$. The exponents $(e_i(\chi) \mid 1 \leq i \leq \chi(1))$ of an irreducible character $\chi$ of $W$ are defined by the formula $R_\chi = \sum_{i=1}^{\chi(1)} x^{e_i(\chi)}$. The exponents $m_i$ of $W$ are now just the exponents of the contragradient of the reflection representation $\rho^* := \text{tr} V^*$, that is, $R_{\rho^*} = \sum_{i=1}^{n} x^{m_i}$. Dually, the coexponents are the exponents of $\rho$. In particular, for real reflection groups exponents and coexponents coincide. In general $N^*(W) := \sum_{i=1}^{n} m_i^*$ equals the number of reflecting hyperplanes of $W$. The $d_i^* := m_i^* - 1$ are sometimes called the codegrees of $W$.

If $W$ is a Weyl group (see Section 2.10), the fake degrees constitute a first approximation to the degrees of principal series unipotent characters of finite groups of Lie type with Weyl group $W$. See also Section 4.8 for further properties.

1.8. Reflection data

In the general theory of finite groups of Lie type (where an algebraic group comes with an action of a Frobenius map) as well as in the study of Levi subgroups it is natural to consider reflection groups together with an automorphism $\phi$ normalising the reflection representation; see the survey article Broué and Malle, [36]. This leads to the following abstract definition.

A pair $(V, W\phi)$ is called a reflection datum if $V$ is a vector space over a subfield $K \subseteq \mathbb{C}$ and $W\phi$ is a coset in $\text{GL}(V)$ of a reflection group $W \subseteq \text{GL}(V)$, where $\phi \in \text{GL}(V)$ normalises $W$.

A sub-reflection datum of a reflection datum $(V, W\phi)$ is a reflection datum of the form $(V', W'(w\phi)|_{V'})$, where $V'$ is a subspace of $V$, $W'$ is a reflection subgroup of $N_W(V')|_{V'}$ stabilising $V'$ (hence, a reflection subgroup of $N_W(V')/W_{V'}$), and $w\phi$ is an element of $W\phi$ stabilising $V'$ and normalising $W'$. A Levi sub-reflection datum of $(V, W\phi)$ is a sub-reflection datum of the form $(V, W_{V'}(w\phi))$ for some subspace $V' \subseteq V$ (note that, by Theorem 1.6, $W_{V'}$ is indeed a reflection subgroup of $W$). A torus of $G$ is a sub-reflection datum with trivial reflection group.

Let $G = (V, W\phi)$ be a reflection datum. Then $\phi$ acts naturally on the symmetric algebra $K[V]$. It is possible to choose basic invariants $f_1, \ldots, f_n \in K[V]_W$, such that $f_i^\phi = e_i f_i$ for roots of unity $e_1, \ldots, e_n$. The multiset $\{(d_i, \epsilon_i)\}$ of generalised degrees of $G$ then only depends on $W$ and $\phi$ (see, for example, Springer, [170, Lemma 6.1]). The polynomial order of the reflection datum $G = (V, W\phi)$ is by definition the polynomial

$$|G| := \frac{\epsilon_G x^{N(W)}}{|W| \sum_{w \in W} \det_V(1-x^{-1}w\phi)^2} = x^{N(W)} \prod_{i=1}^{n} (x^{d_i} - \epsilon_i),$$

where $\epsilon_G := (-1)^n \epsilon_1 \cdots \epsilon_n$. Let $\Phi(x)$ be a cyclotomic polynomial over $K$. A torus $\mathbb{T} = (V', (w\phi)|_{V'})$ of $G$ is called a $\Phi$-torus if the polynomial order of $\mathbb{T}$ is a power of $\Phi$.

Reflection data can be thought of as the skeletons of finite reductive groups.
1.9. Regular elements

In this section we present results which show that certain subgroups respectively subquotients of reflection groups are again reflection groups. Let \((V, w\phi)\) be a reflection datum over \(K = \mathbb{C}\). For \(w\phi \in W\) and a root of unity \(\zeta \in \mathbb{C}^\times\) write
\[
V(w\phi, \zeta) := \{v \in V \mid w\phi \cdot v = \zeta v\}
\]
for the \(\zeta\)-eigenspace of \(w\phi\). Note that \((V(w\phi, \zeta), w\phi)\) is an \((x - \zeta)\)-torus of \((V, w\phi)\) in the sense defined above. These \((x - \zeta)\)-tori for fixed \(\zeta\) satisfy a kind of Sylow theory.

Let \(f_1, \ldots, f_n\) be a set of basic invariants for \(W\) and \(H_i\) the surface defined by \(f_i = 0\). Springer, [170], proves that
\[
\bigcup_{w\phi \in W} V(w\phi, \zeta) = \bigcap_{i : \epsilon_i \zeta^{d_i} = 1} H_i,
\]
the irreducible components of this algebraic set are just the maximal \(V(w\phi, \zeta)\), \(W\) acts transitively on these components, and their common dimension is just the number \(a(d, \phi)\) of indices \(i\) such that \(\epsilon_i \zeta^{d_i} = 1\), where \(d\) denotes the order of \(\zeta\). (Note that \(a(d, \phi)\) only depends on \(d\), not on \(\zeta\) itself.) From this he obtains:

1.10. THEOREM (Springer, [170, Theorems 3.4 and 6.2]). Let \((V, W\phi)\) be a reflection datum over \(\mathbb{C}\), \(\zeta\) a primitive \(d\)-th root of unity. Then:

(i) \(\max\{\dim V(w\phi, \zeta) \mid w \in W\} = a(d, \phi)\).

(ii) For any \(w \in W\) there exists a \(w' \in W\) such that \(V(w\phi, \zeta) \subseteq V(w'\phi, \zeta)\) and \(V(w'\phi, \zeta)\) has maximal dimension.

(iii) If \(\dim V(w\phi, \zeta) = \dim V(w'\phi, \zeta) = a(d, \phi)\) then there exists a \(u \in W\) with \(u \cdot V(w\phi, \zeta) = V(w'\phi, \zeta)\).

This can be rephrased as follows: Let \(K\) be a subfield of \(\mathbb{C}\), \(\Phi\) a cyclotomic polynomial over \(K\). A torus \(T\) of \(G\) is called a \(\Phi\)-Sylow torus, if its order equals the full \(\Phi\)-part of the order of \(G\). Then \(\Phi\)-tori of \(G\) satisfy the three statements of Sylow’s theorem. (For an analogue of the statement on the number of Sylow subgroups see Broué, Malle and Michel, [37, Theorem 5.1(4)].)

This can in turn be used to deduce a Sylow theory for tori in finite groups of Lie type (see Broué and Malle, [34]).

A vector \(v \in V\) is called regular (for \(W\)) if it is not contained in any reflecting hyperplane, i.e. (by Theorem 1.6), if its stabiliser \(W_v\) is trivial. Let \(\zeta \in \mathbb{C}\) be a root of unity. An element \(w\phi \in W\) is \(\zeta\)-regular if \(V(w\phi, \zeta)\) contains a regular vector. By definition, if \(w\phi\) is regular for some root of unity, then so is any power of \(w\phi\). If \(\phi = 1\), theorem 1.6 of Steinberg implies that the orders of \(w\) and \(\zeta\) coincide. An integer \(d\) is a regular number for \(W\) if it is the order of a regular element of \(W\).

1.11. THEOREM (Springer, [170, Theorem 6.4 and Proposition 4.5]). Let \(w\phi \in W\phi\) be \(\zeta\)-regular of order \(d\). Then:
(i) \( \dim V(w\phi, \zeta) = a(d, \phi) \).

(ii) The centraliser of \( w\phi \) in \( W \) is isomorphic to a reflection group in \( V(w\phi, \zeta) \) whose degrees are the \( d_i \) with \( \varepsilon_i \zeta^{d_i} = 1 \).

(iii) The elements of \( W\phi \) with property (i) form a single conjugacy class under \( W \).

(iv) Let \( \phi = 1 \) and let \( \chi \) be an irreducible character of \( W \). Then the eigenvalues of \( w \) in a representation with character \( \chi \) are \( (\zeta^{m_i} \chi | 1 \leq i \leq \chi(1)) \).

In particular, it follows from (iv) that the eigenvalues of a \( \zeta \)-regular element \( w \) on \( V \) are \( (\zeta^{m_i} | 1 \leq i \leq n) \).

Interestingly enough, the normaliser modulo centraliser of arbitrary Sylow tori of reflection data are naturally reflection groups, as the following generalisation of the previous result shows:

1.12. Theorem (Lehrer and Springer, [131, 132]). Let \( w \in W \) and \( \tilde{V} := V(w\phi, \zeta) \) be such that \( (\tilde{V}, w\phi) \) is a \( \Phi \)-Sylow torus. Let \( N := \{ w' \in W | w' \cdot \tilde{V} = \tilde{V} \} \) be the normaliser, \( C := \{ w' \in W | w' \cdot v = v \text{ for all } v \in \tilde{V} \} \) the centraliser of \( \tilde{V} \).

(i) Then \( N/C \) acts as a reflection group on \( \tilde{V} \), with reflecting hyperplanes the intersections with \( \tilde{V} \) of those of \( W \).

(ii) A set of basic invariants of \( N/C \) is given by the restrictions to \( \tilde{V} \) of those \( f_i \) with \( \varepsilon_i \zeta^{d_i} = 1 \).

(iii) If \( W \) is irreducible on \( V \), then so is \( N/C \) on \( \tilde{V} \).

In the case of regular elements, the second assertion of (i) goes back to Lehrer, [128, 5.8], Denef and Loeser, [64]; see also Broué and Michel, [39, Proposition 3.2].

1.13. The Shephard–Todd classification

Let \( V \) be a finite-dimensional complex vector space and \( W \leq \text{GL}(V) \) a reflection group. Since \( W \) is finite, the representation on \( V \) is completely reducible, and \( W \) is the direct product of irreducible reflection subgroups. Thus, in order to determine all reflection groups over \( \mathbb{C} \), it is sufficient to classify the irreducible ones. This was achieved by Shephard and Todd, [167].

To describe this classification, first recall that a subgroup \( W \leq \text{GL}(V) \) is called imprimitive if there exists a direct sum decomposition \( V = V_1 \oplus \cdots \oplus V_k \) with \( k > 1 \) stabilised by \( W \) (that is, \( W \) permutes the summands). The bulk of irreducible complex reflection groups consists of imprimitive ones. For any \( d, e, n \geq 1 \) let \( G(de, e, n) \) denote the group of monomial \( n \times n \)-matrices (that is, matrices with precisely one non-zero entry in each row and column) with non-zero entries in the set of \( de \)-th roots of unity, such that the product over these entries is a \( d \)-th root of unity.

Explicit generators may be chosen as follows: \( G(d, 1, n) \) is generated on \( \mathbb{C}^n \) with standard Hermitian form by the reflection \( t_1 \) of order \( d \) with root the first standard basis vector \( b_1 \) and by the permutation matrices \( t_2, \ldots, t_n \) for the transpositions \( (1, 2), (2, 3), \ldots, (n-1, n) \). For \( d > 1 \) this is an irreducible reflection group, isomorphic to the wreath product \( C_d \wr \Sigma_n \) of the cyclic group of order \( d \) with the symmetric group \( \Sigma_n \).
where the base group is generated by the reflections of order $d$ with roots the standard basis vectors, and a complement consists of all permutation matrices.

Let $\gamma_d: G(d, 1, n) \to \mathbb{C}^\times$ be the linear character of $G(d, 1, n)$ obtained by tensoring the determinant on $V$ with the sign character on the quotient $\mathfrak{S}_n$. Then for any $e > 1$ we have

$$G(de, e, n) := \ker(\gamma_{de}) \leq G(de, 1, n).$$

This is an irreducible reflection subgroup of $G(de, 1, n)$ for all $n \geq 2$, $d \geq 1$, $e \geq 2$, except for $(d, e, n) = (2, 2, 2)$. It is generated by the reflections

$$t_1^e, t_1^{-1}, t_2 t_1, t_2, t_3, \ldots, t_n,$$

where the first generator is redundant if $d = 1$. Clearly, $G(de, e, n)$ stabilises the decomposition $V = \mathbb{C}b_1 \oplus \cdots \oplus \mathbb{C}b_n$ of $V$, so it is imprimitive for $n > 1$. The order of $G(de, e, n)$ is given by $d^n e^{n-1} n!$. Using the wreath product structure it is easy to show that the only isomorphisms among groups in this series are $G(2, 1, 2) \cong G(4, 4, 2)$, and $G(de, e, 1) \cong G(d, 1, 1)$ for all $d, e$, while $G(2, 2, 2)$ is reducible. All these are isomorphisms of reflection groups.

In its natural action on $\mathbb{Q}^n$ the symmetric group $\mathfrak{S}_n$ stabilises the 1-dimensional subspace consisting of vectors with all coordinates equal and the $(n - 1)$-dimensional subspace consisting of those vectors whose coordinates add up to 0. In its action on the latter, $\mathfrak{S}_n$ is an irreducible and primitive reflection group. The classification result may now be stated as follows (see also Cohen, [50]):

1.14. **Theorem** (Shephard and Todd, [167]). The irreducible complex reflection groups are the groups $G(de, e, n)$, for $de \geq 2$, $n \geq 1$, $(de, e, n) \neq (2, 2, 2)$, the groups $\mathfrak{S}_n$ ($n \geq 2$) in their $(n - 1)$-dimensional natural representation, and 34 further primitive groups.

Moreover, any irreducible $n$-dimensional complex reflection group has a generating set of at most $n + 1$ reflections.

The primitive groups are usually denoted by $G_4, \ldots, G_{37}$, as in the original article [167] (where the first three indices were reserved for the families of imprimitive groups $G(de, e, n)$ ($de, n \geq 2$), cyclic groups $G(d, 1, 1)$ and symmetric groups $\mathfrak{S}_{n+1}$). An $n$-dimensional irreducible reflection groups generated by $n$ of its reflections is called well-generated. The groups for which this fails are the imprimitive groups $G(de, e, n)$, $d, e, n \geq 1$, and the primitive groups

$$G_i \quad \text{with} \quad i \in \{7, 11, 12, 13, 15, 19, 22, 31\}.$$

By construction $G(de, e, n)$ contains the well-generated group $G(de, de, n)$. More generally, the classification implies the following, for which no a priori proof is known:

1.15. **Corollary.** Any irreducible complex reflection group $W \leq \text{GL}(V)$ contains a well-generated reflection subgroup $W' \leq W$ which is still irreducible on $V$. 

The primitive groups $G_4, \ldots, G_{37}$ occur in dimensions 2 up to 8. In Table 1 we collect some data on the irreducible complex reflection groups. (These and many more data for complex reflection groups have been implemented by Jean Michel into the CHEVIE-system, [89].) In the first part, the dimension is always equal to $n$, in the second it can be read of from the number of degrees. We give the degrees, the codegrees in case they are not described by the following Theorem 1.16 (that is, if $W$ is not well-generated), and the character field $K_W$ of the reflection representation. For the exceptional groups we also give the structure of $W/Z(W)$ and we indicate the regular degrees by boldface (that is, those degrees which are regular numbers for $W$, see Cohen, [50, p. 395 and p. 412], and Springer, [170, Tables 1–6]). The regular degrees for the infinite series are: $n, n + 1$ for $S_{n+1}$, $dn$ for $G(de, e, n)$ with $d > 1$, $(n-1)e$ for $G(e, e, n)$ with $n|e$, and $(n-1)e, n$ for $G(e, e, n)$ with $n|e$. Lehrer and Michel, [130, Theorem 3.1], have shown that an integer is a regular number if and only if it divides as many degrees as codegrees.

Fundamental invariants for most types are given in Shephard and Todd, [167], as well as defining relations and further information on parabolic subgroups (see also Coxeter, [56], Shephard, [166], and Broué, Malle and Rouquier, [38], for presentations, and the tables in Cohen, [50], and Broué, Malle and Rouquier, [38, Appendix 2]).

If the irreducible reflection group $W$ has an invariant of degree 2, then it leaves invariant a non-degenerate quadratic form, so the representation may be realised over the reals. Conversely, if $W$ is a real reflection group, then it leaves invariant a quadratic form. Thus the real irreducible reflection groups are precisely those with $d_1 = 2$, that is, the infinite series $G(2, 1, n)$, $G(2, 2, n)$, $G(e, e, 2)$ and $S_{n+1}$, and the six exceptional groups $G_{23}, G_{28}, G_{30}, G_{35}, G_{36}, G_{37}$ (see Section 2.5).

The degrees and codegrees of a finite complex reflection group satisfy some remarkable identities. As an example, let us quote the following result, for which at present only a case-by-case proof is known:

**1.16. Theorem (Orlik and Solomon, [154]).** Let $W$ be an irreducible complex reflection group in dimension $n$. Then the following are equivalent:

(i) $d_i + d_{n-i+1}^* = d_n$ for $i = 1, \ldots, n$,

(ii) $N + N^* = nd_n$,

(iii) $d_i^* < d_n$ for $i = 1, \ldots, n$,

(iv) $W$ is well-generated.

See also Terao and Yano, [180], for a partial explanation.

From the Shephard–Todd classification, it is straightforward to obtain a classification of reflection data. An easy argument allows to reduce to the case where $W$ acts irreducibly on $V$. Then either up to scalars $\phi$ can be chosen to be a reflection, or $W = G_{28}$ is the real reflection group of type $F_4$, and $\phi$ induces the graph automorphism on the $F_4$-diagram (see Broué, Malle and Michel, [37, Proposition 3.13]). An infinite series of examples is obtained from the embedding of $G(de, e, n)$ into $G(de, 1, n)$ (which is the full projective normaliser in all but finitely many cases). Apart from this, there are only six further cases, which we list in Table 2.
Table 1
Irreducible complex reflection groups

<table>
<thead>
<tr>
<th>W</th>
<th>Degrees</th>
<th>Codegrees</th>
<th>$K_W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(d, 1, n)$ $(d \geq 2, n \geq 1)$</td>
<td>$d, 2d, \ldots, nd$</td>
<td>*</td>
<td>$\mathbb{Q}(\zeta_d)$</td>
</tr>
<tr>
<td>$G(d, e, n)$ $(d, e, n \geq 2)$</td>
<td>$ed, 2ed, \ldots, (n - 1)ed, nd$</td>
<td>0, $ed, \ldots, (n - 1)ed$</td>
<td>$\mathbb{Q}(\zeta_{de})$</td>
</tr>
<tr>
<td>$G(e, e, n)$ $(e \geq 2, n \geq 3)$</td>
<td>$e, 2e, \ldots, (n - 1)e, n$</td>
<td>*</td>
<td>$\mathbb{Q}(\zeta_e)$</td>
</tr>
<tr>
<td>$G(e, e, 2)$ $(e \geq 3)$</td>
<td>$2, e, \ldots, (n - 1)e, 2, e, \ldots, (n - 1)e, 2$</td>
<td>*</td>
<td>$\mathbb{Q}(\zeta_e + \zeta_e^{-1})$</td>
</tr>
<tr>
<td>$G_{n+1}$ $(n \geq 1)$</td>
<td>$2, 3, \ldots, n + 1$</td>
<td>*</td>
<td>$\mathbb{Q}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>W</th>
<th>Degrees</th>
<th>Codegrees</th>
<th>$K_W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_4$</td>
<td>4, 6</td>
<td>*</td>
<td>$\mathbb{Q}(\zeta_3)$ $\times \mathbb{A}_4$</td>
</tr>
<tr>
<td>$G_5$</td>
<td>6, 12</td>
<td>*</td>
<td>$\mathbb{A}_4$</td>
</tr>
<tr>
<td>$G_6$</td>
<td>4, 12</td>
<td>*</td>
<td>$\mathbb{A}_4$</td>
</tr>
<tr>
<td>$G_7$</td>
<td>12, 12</td>
<td>0, 12</td>
<td>$\mathbb{A}_4$</td>
</tr>
<tr>
<td>$G_8$</td>
<td>8, 12</td>
<td>*</td>
<td>$\mathbb{E}_3$</td>
</tr>
<tr>
<td>$G_9$</td>
<td>8, 24</td>
<td>*</td>
<td>$\mathbb{E}_4$</td>
</tr>
<tr>
<td>$G_{10}$</td>
<td>12, 24</td>
<td>*</td>
<td>$\mathbb{E}_4$</td>
</tr>
<tr>
<td>$G_{11}$</td>
<td>24, 24</td>
<td>0, 24</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{12}$</td>
<td>6, 8</td>
<td>0, 10</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{13}$</td>
<td>8, 12</td>
<td>0, 16</td>
<td>$\mathbb{A}_4$</td>
</tr>
<tr>
<td>$G_{14}$</td>
<td>6, 24</td>
<td>*</td>
<td>$\mathbb{A}_4$</td>
</tr>
<tr>
<td>$G_{15}$</td>
<td>12, 24</td>
<td>0, 24</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{16}$</td>
<td>20, 30</td>
<td>*</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{17}$</td>
<td>20, 60</td>
<td>*</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{18}$</td>
<td>30, 60</td>
<td>*</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{19}$</td>
<td>60, 60</td>
<td>0, 60</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{20}$</td>
<td>12, 30</td>
<td>*</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{21}$</td>
<td>12, 60</td>
<td>*</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{22}$</td>
<td>12, 20</td>
<td>0, 28</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{23}$</td>
<td>2, 6, 10</td>
<td>*</td>
<td>$\mathbb{A}_5$</td>
</tr>
<tr>
<td>$G_{24}$</td>
<td>4, 6, 14</td>
<td>*</td>
<td>$GL_2(2)$</td>
</tr>
<tr>
<td>$G_{25}$</td>
<td>6, 9, 12</td>
<td>*</td>
<td>$SL_2(3)$</td>
</tr>
<tr>
<td>$G_{26}$</td>
<td>6, 12, 18</td>
<td>*</td>
<td>$SL_2(3)$</td>
</tr>
<tr>
<td>$G_{27}$</td>
<td>6, 12, 30</td>
<td>*</td>
<td>$SL_2(3)$</td>
</tr>
<tr>
<td>$G_{28}$</td>
<td>2, 6, 8, 12</td>
<td>*</td>
<td>$2^4 : G_3 \times \mathbb{G}_3$</td>
</tr>
<tr>
<td>$G_{29}$</td>
<td>4, 8, 12, 20</td>
<td>*</td>
<td>$2^4 : \mathbb{G}_3$</td>
</tr>
<tr>
<td>$G_{30}$</td>
<td>2, 12, 20, 30</td>
<td>*</td>
<td>$\mathbb{A}_5 \times 2$</td>
</tr>
<tr>
<td>$G_{31}$</td>
<td>8, 12, 20, 24</td>
<td>0, 12, 16, 28</td>
<td>$2^4 : \mathbb{G}_6$</td>
</tr>
<tr>
<td>$G_{32}$</td>
<td>12, 18, 24, 30</td>
<td>*</td>
<td>$U_4(2)$</td>
</tr>
<tr>
<td>$G_{33}$</td>
<td>4, 6, 10, 12, 18</td>
<td>*</td>
<td>$O_3(3)$</td>
</tr>
<tr>
<td>$G_{34}$</td>
<td>6, 12, 18, 24, 42</td>
<td>*</td>
<td>$O_3^+(3).2$</td>
</tr>
<tr>
<td>$G_{35}$</td>
<td>2, 5, 6, 8, 9, 12</td>
<td>*</td>
<td>$O_5^+(2)$</td>
</tr>
<tr>
<td>$G_{36}$</td>
<td>2, 6, 8, 10, 12, 14, 18</td>
<td>*</td>
<td>$O_7(2)$</td>
</tr>
<tr>
<td>$G_{37}$</td>
<td>2, 8, 12, 14, 18, 20, 24, 30</td>
<td>*</td>
<td>$O_8^+(2).2$</td>
</tr>
</tbody>
</table>
2. Real reflection groups

In this section we discuss in more detail the special case where $W$ is a **real** reflection group. This is a well-developed theory, and there are several good places to learn about real reflection groups: the classical Bourbaki volume, [25], the very elementary text by Benson and Grove, [8], the relevant chapters in Curtis and Reiner, [58], Hiller, [99], and Humphreys, [105]. Various pieces of the theory have also been recollected in a concise way in articles by Steinberg, [178]. The exposition here partly follows Geck and Pfeiffer, [95, Chapter 1]. We shall only present the most basic results and refer to the above textbooks and our bibliography for further reading.

2.1. Coxeter groups

Let $S$ be a finite non-empty index set and $M = (m_{st})_{s,t \in S}$ be a symmetric matrix such that $m_{ss} = 1$ for all $s \in S$ and $m_{st} \in \{2, 3, 4, \ldots\} \cup \{\infty\}$ for all $s \neq t$ in $S$. Such a matrix is called a **Coxeter matrix**. Now let $W$ be a group containing $S$ as a subset. ($W$ may be finite or infinite.) Then the pair $(W, S)$ is called a **Coxeter system**, and $W$ is called a **Coxeter group**, if $W$ has a presentation with generators $S$ and defining relations of the form

$$(st)^{m_{st}} = 1 \quad \text{for all } s, t \in S \text{ with } m_{st} < \infty;$$

in particular, this means that $s^2 = 1$ for all $s \in S$. Therefore, the above relations (for $s \neq t$) can also be expressed in the form

$$sts \cdots = tst \cdots \quad \text{for all } s, t \in S \text{ with } 2 \leq m_{st} < \infty.$$ 

We say that $C$ is of finite type and that $(W, S)$ is a finite Coxeter system if $W$ is a finite group. The information contained in $M$ can be visualised by a corresponding **Coxeter graph**, which is defined as follows. It has vertices labelled by the elements of $S$, and two vertices labelled by $s \neq t$ are joined by an edge if $m_{st} \geq 3$. Moreover, if $m_{st} \geq 4$, we label the edge by $m_{st}$. The standard example of a finite Coxeter system is the pair $(\mathfrak{S}_n, \{s_1, \ldots, s_{n-1}\})$ where $s_i = (i, i+1)$ for $1 \leq i \leq n - 1$. The corresponding graph is
Coxeter groups have a rich combinatorial structure. A basic tool is the length function \( l : W \to \mathbb{N}_0 \), which is defined as follows. Let \( w \in W \). Then \( l(w) \) is the length of a shortest possible expression \( w = s_1 \cdots s_k \) where \( s_i \in S \). An expression of \( w \) of length \( l(w) \) is called a reduced expression for \( w \). We have \( l(1) = 0 \) and \( l(s) = 1 \) for \( s \in S \). Here is a key result about Coxeter groups.

2.2. Theorem (Matsumoto, [141]; see also Bourbaki, [25]). Let \((W,S)\) be a Coxeter system and \( \mathcal{M} \) be a monoid, with multiplication \( \ast : \mathcal{M} \times \mathcal{M} \to \mathcal{M} \). Let \( f : S \to \mathcal{M} \) be a map such that

\[
\underbrace{f(s) \ast f(t) \ast f(s) \ast \cdots}_{m_{st} \text{ times}} = \underbrace{f(t) \ast f(s) \ast f(t) \ast \cdots}_{m_{st} \text{ times}}
\]

for all \( s \neq t \) in \( S \) such that \( m_{st} < \infty \). Then there exists a unique map \( F : W \to \mathcal{M} \) such that

\[
F(w) = \underbrace{f(s_1) \ast \cdots \ast f(s_k)}_{m_{st} \text{ times}}
\]

whenever \( w = s_1 \cdots s_k \) (\( s_i \in S \)) is reduced.

Typically, Matsumoto’s theorem can be used to show that certain constructions with reduced expressions of elements of \( W \) actually do not depend on the choice of the reduced expressions. We give two examples.

(1) Let \( w \in W \) and take a reduced expression \( w = s_1 \cdots s_k \) with \( s_i \in S \). Then the set \( \{s_1, \ldots, s_k\} \) does not depend on the choice of the reduced expression.

(Indeed, let \( \mathcal{M} \) be the monoid whose elements are the subsets of \( S \) and product given by \( A \ast B := A \cup B \). Then the assumptions of Matsumoto’s theorem are satisfied for the map \( f : S \to \mathcal{M} \), \( s \mapsto \{s\} \), and this yields the required assertion.)

(2) Let \( w \in W \) and fix a reduced expression \( w = s_1 \cdots s_k \) (\( s_i \in S \)). Consider the set of all subexpressions:

\[
\mathcal{S}(w) := \{y \in W \mid y = s_{i_1} \cdots s_{i_l} \text{ where } l \geq 0 \text{ and } 1 \leq i_1 < \cdots < i_l \leq k\}.
\]

Then \( \mathcal{S}(w) \) does not depend on the choice of the reduced expression for \( w \).

(Indeed, let \( \mathcal{M} \) be the monoid whose elements are the subsets of \( W \) and product given by \( A \ast B := \{ab \mid a \in A, b \in B\} \) (for \( A, B \subseteq W \)). Then the assumptions of Matsumoto’s theorem are satisfied for the map \( f : S \to \mathcal{M} \), \( s \mapsto \{s\} \), and this yields the required assertion.)

We also note that the so-called exchange condition and the cancellation law are further consequences of the above results. The “cancellation law” states that, given \( w \in W \) and an expression \( w = s_1 \cdots s_k \) (\( s_i \in S \)) which is not reduced, one can obtain a reduced expression of \( w \) by simply cancelling some of the factors in the given expression. This law together with (2) yields that the relation

\[
y \leq w \iff y \in \mathcal{S}(w)
\]

is a partial order on \( W \), called the Bruhat–Chevalley order. This ordering has been extensively studied; see, for example, Verma, [184], Deodhar, [65], Björner, [19], Lascoux and
Schützenberger, [126], and Geck and Kim, [90]. By Chevalley, [48], it is related to the Bruhat decomposition in algebraic groups; we will explain this result in 2.16 below.

2.3. Cartan matrices

Let $M = (m_{st})_{s,t \in S}$ be a Coxeter matrix as above. We can also associate with $M$ a group generated by reflections. This is done as follows. Choose a matrix $C = (c_{st})_{s,t \in S}$ with entries in $\mathbb{R}$ such that the following conditions are satisfied:

(C1) For $s \neq t$ we have $c_{st} \leq 0$; furthermore, $c_{st} \neq 0$ if and only if $c_{ts} \neq 0$.

(C2) We have $c_{ss} = 2$ and, for $s \neq t$, we have $c_{st}c_{ts} = 4 \cos^2(\pi/m_{st})$.

Such a matrix $C$ will be called a Cartan matrix associated with $M$. For example, we could simply take $c_{st} := -2 \cos(\pi/m_{st})$ for all $s,t \in S$; this may be called the standard Cartan matrix associated with $M$. We always have $0 \leq c_{st}c_{ts} \leq 4$. Here are some values for the product $c_{st}c_{ts}$:

<table>
<thead>
<tr>
<th>$m_{st}$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>$\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{st}c_{ts}$</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>$(3+\sqrt{5})/2$</td>
<td>3</td>
<td>2</td>
<td>$\sqrt{2}$</td>
</tr>
</tbody>
</table>

Now let $V$ be an $\mathbb{R}$-vector space of dimension $|S|$, with a fixed basis $\{\alpha_s \mid s \in S\}$. We define a linear action of the elements in $S$ on $V$ by the rule:

$s : V \to V, \quad \alpha_t \mapsto \alpha_t - c_{st}\alpha_s \ (t \in S)$.

It is easily checked that $s \in \text{GL}(V)$ has order 2 and precisely one eigenvalue $-1$ (with eigenvector $\alpha_s$). Thus, $s$ is a reflection with root $\alpha_s$. We then define

$W = W(C) := \langle S \rangle \subseteq \text{GL}(V)$;

thus, if $|W| < \infty$ is finite, then $W$ will be a real reflection group. Now we can state the following basic result.

2.4. THEOREM (Coxeter, [54,55]). Let $M = (m_{st})_{s,t \in S}$ be a Coxeter matrix and $C$ be a Cartan matrix associated with $M$. Let $W(C) = \langle S \rangle \subseteq \text{GL}(V)$ be the group constructed as in 2.3. Then the pair $(W(C), S)$ is a Coxeter system. Furthermore, the group $W(C)$ is finite if and only if

the matrix $(-\cos(\pi/m_{st}))_{s,t \in S}$ is positive-definite, \hspace{1cm} (*)

i.e., we have $\det(-\cos(\pi/m_{st}))_{s,t \in J} > 0$ for every subset $J \subseteq S$. All finite real reflection groups arise in this way.

The fact that $(W(C), S)$ is a Coxeter system is proved in [95, 1.2.7]. The finiteness condition can be found in Bourbaki, [25, Chapter V, §4, no. 8]. Finally, the fact that all finite real reflection groups arise in this way is established in [25, Chapter V, §3, no. 2]. The “note historique” in [25] contains a detailed account of the history of the above result.
2.5. Classification of finite Coxeter groups

(See also Section 1.13.) Let $M = (m_{st})_{s,t \in S}$ be a Coxeter matrix. We say that $M$ is decomposable if there is a partition $S = S_1 \sqcup S_2$ with $S_1$, $S_2 \neq \emptyset$ and such that $m_{st} = 2$ whenever $s \in S_1$, $t \in S_2$. If $C = (c_{st})$ is any Cartan matrix associated with $M$, then this condition translates to: $c_{st} = c_{ts} = 0$ whenever $s \in S_1$, $t \in S_2$. Correspondingly, we also have a direct sum decomposition $V = V_1 \oplus V_2$ where $V_1$ has basis $\{\alpha_s \mid s \in S_1\}$ and $V_2$ has basis $\{\alpha_s \mid s \in S_2\}$. Then it easily follows that we have an isomorphism

$$W(C) \cong W(C_1) \times W(C_2), \quad w \mapsto (w|_{V_1}, w|_{V_2}).$$

In this way, the study of the groups $W(C)$ is reduced to the case where $C$ is indecomposable (i.e., there is no partition $S = S_1 \sqcup S_1$ as above). If this holds, we call the corresponding Coxeter system $(W, S)$ an irreducible Coxeter system.

2.6. Theorem. The Coxeter graphs of the indecomposable Coxeter matrices $M$ such that condition $(\ast)$ in Theorem 2.4 holds are precisely the graphs in Table 3.

For the proof of this classification, see [25, Chapter VI, no. 4.1]. The identification with the groups occurring in the Shephard–Todd classification (see Theorem 1.14) is given in the following table.

<table>
<thead>
<tr>
<th>Coxeter graph</th>
<th>Shephard–Todd</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_n-1$</td>
<td>$E_n$</td>
</tr>
<tr>
<td>$B_n$</td>
<td>$G(2, 1, n)$</td>
</tr>
<tr>
<td>$D_n$</td>
<td>$G(2, 2, n)$</td>
</tr>
<tr>
<td>$I_2(m)$</td>
<td>$G(m, m, 2)$</td>
</tr>
<tr>
<td>$H_4$</td>
<td>$G_{23}$</td>
</tr>
<tr>
<td>$H_4$</td>
<td>$G_{30}$</td>
</tr>
<tr>
<td>$F_4$</td>
<td>$G_{28}$</td>
</tr>
<tr>
<td>$E_6$</td>
<td>$G_{35}$</td>
</tr>
<tr>
<td>$E_7$</td>
<td>$G_{36}$</td>
</tr>
<tr>
<td>$E_8$</td>
<td>$G_{37}$</td>
</tr>
</tbody>
</table>

Thus, any finite irreducible real reflection group is the reflection group arising from a Cartan matrix associated with one of the graphs in Table 3.

Now, there are a number of results on finite Coxeter groups which can be formulated in general terms but whose proof requires a case-by-case verification using the above classification. We mention two such results, concerning conjugacy classes.

2.7. Theorem (Carter, [41]). Let $(W, S)$ be a finite Coxeter system. Then every element in $W$ is conjugate to its inverse. More precisely, given $w \in W$, there exist $x, y \in W$ such that $w = xy$ and $x^2 = y^2 = 1$.

Every element $x \in W$ such that $x^2 = 1$ is a product of pairwise commuting reflections in $W$. Given $w \in W$ and an expression $w = xy$ as above, the geometry of the roots involved in the reflections determining $x, y$ yields a diagram which can be used to label the conjugacy class of $w$. Complete lists of these diagrams can be found in [41].
2.8. Conjugacy classes and the length function

Let $(W, S)$ be a Coxeter system and $C$ be a conjugacy class in $W$. We will be interested in studying how conjugation inside $C$ relates to the length function on $W$. For this purpose, we introduce two relations, following Geck and Pfeiffer, [94].

Given $x, y \in W$ and $s \in S$, we write $x \overset{s}{\to} y$ if $y = sxs$ and $l(y) \leq l(x)$. We shall write $x \to y$ if there are sequences $x_0, x_1, \ldots, x_n \in W$ and $s_1, \ldots, s_n \in S$ (for some $n \geq 0$) such that

$$x = x_0 \xrightarrow{s_1} x_1 \xrightarrow{s_2} x_2 \xrightarrow{s_3} \cdots \xrightarrow{s_n} x_n = y.$$ 

Thus, we have $x \to y$ if we can go from $x$ to $y$ by a chain of conjugations by generators in $S$ such that, at each step, the length of the elements either remains the same or decreases.

In a slightly different direction, let us now consider two elements $x, y \in W$ such that $l(x) = l(y)$. We write $x \sim y$ (where $w \in W$) if $wx = yw$ and $l(wx) = l(w) + l(x)$ or $xw = wy$ and $l(wy) = l(w) + l(y)$. We write $x \sim y$ if there are sequences $x_0, x_1, \ldots, x_n \in W$ and $w_1, \ldots, w_n \in W$ (for some $n \geq 0$) such that

$$x = x_0 \overset{w_1}{\sim} x_1 \overset{w_2}{\sim} x_2 \overset{w_3}{\sim} \cdots \overset{w_n}{\sim} x_n = y.$$ 

Thus, we have $x \sim y$ if we can go from $x$ to $y$ by a chain of conjugations with elements of $W$ such that, at each step, the length of the elements remains the same and an additional length condition involving the conjugating elements is satisfied. This additional condition has the following significance. Consider a group $\mathcal{M}$ with multiplication $\ast$ and assume that we have a map $f : S \to \mathcal{M}$ which satisfies the requirements in Matsumoto’s Theorem 2.2.
Then we have a canonical extension of $f$ to a map $F:W \to \mathcal{M}$ such that $F(ww') = F(w) \ast F(w')$ whenever $l(ww') = l(w) + l(w')$. Hence, in this setting, we have

$$x \sim y \Rightarrow F(x), F(y) \text{ are conjugate in } \mathcal{M}.$$ 

Thus, we can think of the relation “$\sim$” as “universal conjugacy”.

2.9. Theorem (Geck and Pfeiffer [94,95]). Let $(W, S)$ be a finite Coxeter system and $C$ be a conjugacy class in $W$. We set $l_{\min}(C) := \min\{l(w) \mid w \in C\}$ and

$$C_{\min} := \{w \in C \mid l(w) = l_{\min}(C)\}.$$ 

Then the following hold:

(a) For every $x \in C$, there exists some $y \in C_{\min}$ such that $x \rightarrow y$.

(b) For any two elements $x, y \in C_{\min}$, we have $x \sim y$.

Precursors of the above result for type $A$ have been found much earlier by Starkey, [173]; see also Ram, [162]. The above result allows to define the character table of the Iwahori–Hecke algebra associated with $(W, S)$. This is discussed in more detail in the chapter on Hecke algebras. See Richardson, [165], Geck and Michel, [93], Geck and Pfeiffer, [95, Chapter 3], Geck, Kim and Pfeiffer, [91] and Shi, [168], for further results on conjugacy classes. Krammer, [121], studies the conjugacy problem for arbitrary (infinite) Coxeter groups.

2.10. The crystallographic condition

Let $M = (m_{st})_{s,t \in S}$ be a Coxeter matrix such that the connected components of the corresponding Coxeter graph occur in Table 3. We say that $M$ satisfies the crystallographic condition if there exists a Cartan matrix $C$ associated with $M$ which has integral coefficients. In this case, the corresponding reflection group $W = W(C)$ is called a Weyl group. The significance of this notion is that there exists a corresponding semisimple Lie algebra over $\mathbb{C}$; see 2.15.

Now assume that $M$ is crystallographic. By condition (C2) this implies $m_{st} \in \{2, 3, 4, 6\}$ for all $s \neq t$. Conversely, if $m_{st}$ satisfies this condition, then we have

$$
\begin{align*}
    c_{st}c_{ts} &= 0 \quad \text{if } m_{st} = 2, \\
    c_{st}c_{ts} &= 1 \quad \text{if } m_{st} = 3, \\
    c_{st}c_{ts} &= 2 \quad \text{if } m_{st} = 4, \\
    c_{st}c_{ts} &= 3 \quad \text{if } m_{st} = 6.
\end{align*}
$$

Thus, in each of these cases, we see that there are only two choices for $c_{st}$ and $c_{ts}$: we must have $c_{st} = -1$ or $c_{ts} = -1$ (and then the other value is determined). We encode this additional information in the Coxeter graph, by putting an arrow on the edge between the nodes labelled by $s, t$ according to the following scheme:
**Reflection groups**

<table>
<thead>
<tr>
<th>Edge between $s \neq t$</th>
<th>Values for $c_{st}, c_{ts}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{st} = 3$</td>
<td>no arrow</td>
</tr>
<tr>
<td></td>
<td>$c_{st} = c_{ts} = -1$</td>
</tr>
<tr>
<td>$m_{st} = 4$</td>
<td>$s$ $t$</td>
</tr>
<tr>
<td></td>
<td>$c_{st} = -1$ $c_{ts} = -2$</td>
</tr>
<tr>
<td>$m_{st} = 6$</td>
<td>$s$ $t$</td>
</tr>
<tr>
<td></td>
<td>$c_{st} = -1$ $c_{ts} = -3$</td>
</tr>
</tbody>
</table>

Table 4  
Dynkin diagrams of Cartan matrices of finite type

The Coxeter graph of $M$ equipped with this additional information will be called a Dynkin diagram; it uniquely determines an integral Cartan matrix $C$ associated with $M$ (if such a Cartan matrix exists). The complete list of connected components of Dynkin diagrams is given in table 4. We see that all irreducible finite Coxeter groups are Weyl groups, except for those of type $H_3, H_4$ and $I_2(m)$ where $m = 5$ or $m \geq 7$. Note that type $B_n$ is the only case where we have two different Dynkin diagrams associated with the same Coxeter graph.

The following discussion of root systems associated with finite reflection groups follows the appendix on finite reflection groups in Steinberg, [176].

**2.11. Root systems**

Let $V$ be a finite-dimensional real vector space and let $( , )$ be a positive-definite scalar product on $V$. Given a non-zero vector $\alpha \in V$, the corresponding reflection $w_\alpha \in \text{GL}(V)$ is defined by

$$w_\alpha(v) = v - 2\frac{(v, \alpha)}{(\alpha, \alpha)} \alpha \quad \text{for all } v \in V.$$
Note that, for any \( w \in \text{GL}(V) \), we have \( ww_a w^{-1} = w_{w(a)} \). A finite subset \( \phi \subseteq V \setminus \{0\} \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a simple system if the following conditions are satisfied:

(R) For any \( \alpha \in \Phi \), we have \( \Phi \cap R_{\alpha} = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).

Let \( W(\Phi) \subseteq \text{GL}(V) \) be the subgroup generated by the reflections \( w_{\alpha} \) \( (\alpha \in \Phi) \). A subset \( \Pi \subseteq \Phi \) is called a root system if the following conditions are satisfied:

(R1) For any \( \alpha \in \Phi \), we have \( \Phi \cap \mathbb{R} \alpha = \{ \pm \alpha \} \);

(R2) For every \( \alpha, \beta \in \Phi \), we have \( w_{\alpha}(\beta) \in \Phi \).
2.12. Proposition. Given \( \alpha \in \Pi \) and \( w \in W \), we have
\[
\begin{align*}
  w^{-1}(\alpha) &\in \Phi^+ \iff l(w_{\alpha}w) = l(w) + 1, \\
  w^{-1}(\alpha) &\in \Phi^- \iff l(w_{\alpha}w) = l(w) - 1.
\end{align*}
\]
Furthermore, for any \( w \in W \), we have \( l(w) = |\{ \alpha \in \Phi^+ \mid w(\alpha) \in \Phi^- \}|. \)

The root systems associated with finite Weyl groups are explicitly described in Bourbaki, [25, pp. 251–276]. For type \( H_3 \) and \( H_4 \), see Humphreys, [105, 2.13].

Bremke and Malle, [26,27], have studied suitable generalisations of root systems and length functions for the infinite series \( G(d,1,n) \) and \( G(e,e,n) \), which have subsequently been extended in weaker form by Rampetas and Shoji, [163], to arbitrary imprimitive reflection groups. For investigations of root systems see also Nebe, [147], and Hughes and Morris, [103]. But there is no general theory of root systems and length functions for complex reflection groups (yet).

2.13. Torsion primes

Assume that \( \Phi \subseteq V \) is a root system as above, with a set of simple roots \( \Pi \subseteq \Phi \). Assume that the corresponding Cartan matrix \( C \) is indecomposable and has integral coefficients. Thus, its Dynkin diagram is one of the graphs in Table 4. Following Springer and Steinberg, [172, §1.4], we shall now discuss “bad primes” and “torsion primes” with respect to \( \Phi \).

For every \( \alpha \in \Phi \), the corresponding coroot is defined by \( \alpha^* := \frac{2\alpha}{(\alpha,\alpha)} \). Then \( \Phi^* := \{ \alpha^* \mid \alpha \in \Phi \} \) also is a root system, the dual of \( \Phi \). The Dynkin diagram of \( \Phi^* \) is obtained from that of \( \Phi \) by reversing the arrows. (For example, the dual of a root system of type \( B_n \) is of type \( C_n \).)

Let \( L(\Phi) \) denote the lattice spanned by \( \Phi \) in \( V \). A prime number \( p > 0 \) is called bad for \( \Phi \) if \( L(\Phi)/L(\Phi_1) \) has \( p \)-torsion for some (integrally) closed subsystem \( \Phi_1 \) of \( \Phi \). The prime \( p \) is called a torsion prime if \( L(\Phi^*)/L(\Phi^*_1) \) has \( p \)-torsion for some closed subsystem \( \Phi_1 \) of \( \Phi \). Note that \( \Phi^*_1 \) need not be closed in \( \Phi^* \), and so the torsion primes for \( \Phi \) and the bad primes for \( \Phi^* \) need not be the same. The bad primes can be characterised as follows. Let \( \alpha_0 = \sum_{\alpha \in \Pi} m_{\alpha} \alpha \) be the unique positive root of maximal height. (The height of a root is the sum of the coefficients in the expression of that root as a linear combination of simple roots.) Then we have:

\[
\begin{align*}
p \text{ bad} &\iff p = m_{\alpha} \text{ for some } \alpha \iff p \text{ divides } m_{\alpha} \text{ for some } \alpha \iff p \leq m_{\alpha} \text{ for some } \alpha.'
\end{align*}
\]

Now let \( \alpha_0^* = \sum_{\alpha \in \Pi} m_{\alpha}^* \alpha^*. \) Then \( p \) is a torsion prime if and only if \( p \) satisfies one of the above conditions, with \( m_{\alpha} \) replaced by \( m_{\alpha}^* \). For the various root systems, the bad primes and the torsion primes are given as follows.

<table>
<thead>
<tr>
<th>Type</th>
<th>( A_n ) ( B_n ) ( C_n ) ( D_n ) ( G_2 ) ( F_4 ) ( E_6 ) ( E_7 ) ( E_8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad</td>
<td>none 2 2 2 2,3,3,3,3,2,3,2,3,2,3,5</td>
</tr>
<tr>
<td>Torsion</td>
<td>none 2 none 2 2 2,3,3,3,2,3,2,3,2,3,5</td>
</tr>
</tbody>
</table>
The bad primes and torsion primes play a role in various questions related to sub-root systems, centralisers of semisimple elements in algebraic groups, the classification of unipotent classes in simple algebraic groups and so on; see [172] and also the survey in [43, §§1.14–1.15].

2.14. Affine Weyl groups

Let \( \Phi \subseteq V \) be a root system as above, with Weyl group \( W \).

Let \( L(\Phi) := \sum_{\alpha \in \Pi} \mathbb{Z} \alpha \subseteq V \) be the lattice spanned by the roots in \( V \). Then \( W \) leaves \( L(\Phi) \) invariant and we have a natural group homomorphism \( W \to \text{Aut}(L(\Phi)) \). The semidirect product

\[
W_\alpha(\Phi) := L(\Phi) \rtimes W
\]

is called the affine Weyl group associated with the root system \( W \); see Bourbaki, [25, Chapter VI, §2]. The group \( W_\alpha(\Phi) \) itself is a Coxeter group. The corresponding presentation can also be encoded in a graph, as follows. Let \( \alpha_0 \) be the unique positive root of maximal height in \( \Phi \). We define an extended Cartan matrix \( \tilde{C} \) by similar rules as before:

\[
\tilde{c}_{\alpha\beta} := \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \quad \text{for } \alpha, \beta \in \Pi \cup \{-\alpha_0\}.
\]

The extended Dynkin diagrams encoding these matrices for irreducible \( W \) are given in Table 5. They are obtained from the diagrams in Table 4 by adjoining an additional node (corresponding to \(-\alpha_0\)) and putting edges according to the same rules as before.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Extended Dynkin diagrams</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tilde{A}_n )</td>
<td>( n \geq 2 )</td>
</tr>
<tr>
<td>( \tilde{A}_1 )</td>
<td></td>
</tr>
<tr>
<td>( \tilde{B}_n )</td>
<td>( n \geq 3 )</td>
</tr>
<tr>
<td>( \tilde{F}_4 )</td>
<td></td>
</tr>
<tr>
<td>( \tilde{G}_2 )</td>
<td></td>
</tr>
<tr>
<td>( \tilde{E}_6 )</td>
<td></td>
</tr>
<tr>
<td>( \tilde{E}_7 )</td>
<td></td>
</tr>
<tr>
<td>( \tilde{E}_8 )</td>
<td></td>
</tr>
</tbody>
</table>
In the following subsections, we describe some situations where Coxeter groups and root systems arise “in nature”.

### 2.15. Kac–Moody algebras

Here we briefly discuss how Coxeter groups and root systems arise in the theory of Lie algebras or, more generally, Kac–Moody algebras. We follow the exposition in Kac, [113]. Let $C = (c_{st})_{s,t \in S}$ be a Cartan matrix all of whose coefficients are integers. We also assume that $C$ is symmetrisable, i.e., there exists a diagonal invertible matrix $D$ and a symmetric matrix $B$ such that $C = DB$. A realisation of $C$ is a triple $(h, \Pi, \Pi^\vee)$ where $h$ is a complex vector space, $\Pi = \{\alpha_s \mid s \in S\} \subseteq h^* := \text{Hom}(h, \mathbb{C})$ and $\Pi^\vee = \{\alpha_s^\vee \mid s \in S\}$ are subsets of $h^*$ and $h$, respectively, such that the following conditions hold.

(a) Both sets $\Pi$ and $\Pi^\vee$ are linearly independent;
(b) we have $\langle \alpha_s^\vee, \alpha_t \rangle := \alpha_t(\alpha_s^\vee) = c_{st}$ for all $s,t \in S$;
(c) $|S| - \text{rank}(C) = \dim h - |S|$.

Let $g(C)$ be the corresponding Kac–Moody algebra. Then $g(C)$ is a Lie algebra which is generated by $h$ together with two collections of elements $\{e_s \mid s \in S\}$ and $\{f_s \mid s \in S\}$, where the following relations hold:

\[
\begin{align*}
[e_s, f_t] &= \delta_{st}\alpha_s^\vee, \\
[h, h'] &= 0, \\
[h, e_s] &= (h, \alpha_s)e_s, \\
[h, f_s] &= -\langle h, \alpha_s \rangle f_s, \\
(\text{ad } e_s)^{1-c_{st}} e_t &= 0, \\
(\text{ad } f_s)^{1-c_{st}} f_t &= 0.
\end{align*}
\]

(By [113, 9.11], this is a set of defining relations for $g(C)$.) We have a direct sum decomposition

\[ g(C) = h \oplus \bigoplus_{0 \neq \alpha \in Q} g_\alpha(C), \quad \text{where } Q := \sum_{s \in S} \mathbb{Z}\alpha_s \subseteq h^* \]

and $g_\alpha(C) := \{x \in g(C) \mid [h, x] = \alpha(h)x\} \text{ for all } h \in h$; here, $h = g_0$. The set of all $0 \neq \alpha \in Q$ such that $g_\alpha(C) \neq \{0\}$ will be denoted by $\Phi$ and called the root system of $g(C)$.

For each $s \in S$, we define a linear map $\sigma_s : h^* \to h^*$ by the formula

\[ \sigma_s(\lambda) = \lambda - \langle \lambda, \alpha_s^\vee \rangle \alpha_s \quad \text{for } \lambda \in h^*. \]

Then it is easily checked that $\sigma_s$ is a reflection where $\sigma_s(\alpha_s) = -\alpha_s$. We set

\[ W = W(C) = \langle \sigma_s \mid s \in S\rangle \subseteq \text{GL}(h^*). \]

Now we can state (see [113, 3.7, 3.11 and 3.13]):
(a) The pair \((W, \{\sigma_s \mid s \in S\})\) is a Coxeter system; the corresponding Coxeter matrix is the one associated to \(C\).

(b) The root system \(\Phi\) is invariant under the action of \(W\) and we have \(l(\sigma_s w) = l(w) + 1\) if and only if \(w^{-1}(\alpha_s) \in \Phi^+\), where \(\Phi^+\) is defined as in (2.11).

Thus, \(W\) and \(\Phi\) have similar properties as before. Note, however, that here we did not make any assumption on \(C\) (except that it is symmetrisable with integer entries) and so \(W\) and \(\Phi\) may be infinite. The finite case is characterised as follows:

\[
|W| < \infty \iff |
\Phi| < \infty \iff \text{dim } g(C) < \infty
\]

Thus, all connected components of \(C\) occur in table 4.

(This follows from [113, 3.12] and the characterisation of finite Coxeter groups in Theorem 2.4.) In fact, the finite-dimensional Kac–Moody algebras are precisely the “classical” semisimple complex Lie algebras (see, for example, Humphreys, [104]). The Kac–Moody algebras and the root systems associated to so-called Cartan matrices of affine type have an extremely rich structure and many applications in other branches of mathematics and mathematical physics; see Kac, [113].

2.16. Groups with a \(BN\)-pair

Let \(G\) be an abstract group. We say that \(G\) is a group with a \(BN\)-pair or that \(G\) admits a Tits system if there are subgroups \(B, N \subseteq G\) such that the following conditions are satisfied.

\((BN1)\) \(G\) is generated by \(B\) and \(N\).

\((BN2)\) \(T := B \cap N\) is normal in \(N\) and the quotient \(W := N/T\) is a finite group generated by a set \(S\) of elements of order 2.

\((BN3)\) \(n_s B n_s \neq B\) if \(s \in S\) and \(n_s\) is a representative of \(s\) in \(N\).

\((BN4)\) \(n_s B n \subseteq B n_s B \cup B n B\) for any \(s \in S\) and \(n \in N\).

The group \(W\) is called the Weyl group of \(G\). In fact, it is a consequence of the above axioms that the pair \((W, S)\) is a Coxeter system; see [25, Chapter IV, §2, Théorème 2]. The notion of groups with a \(BN\)-pair was invented by Tits; see [181]. The standard example of a group with a \(BN\)-pair is the general linear group \(G = \text{GL}_n(K)\), where \(K\) is any field and

\[
B := \text{subgroup of all upper triangular matrices in } G, \\
N := \text{subgroup of all monomial matrices in } G.
\]

(A matrix is called monomial if it has exactly one non-zero entry in each row and each column.) We have

\[
T := B \cap N = \text{subgroup of all diagonal matrices in } G
\]

and \(W = N/T \cong \mathfrak{S}_n\). Thus, \(\mathfrak{S}_n\) is the Weyl group of \(G\). More generally, the Chevalley groups (and their twisted analogues) associated with the semisimple complex Lie algebras all have \(BN\)-pairs; see Chevalley, [45], Carter, [42], and Steinberg, [176].
The above set of axioms imposes very strong conditions on the structure of a group $G$ with a $BN$-pair. For example, we have the following Bruhat decomposition, which gives the decomposition of $G$ into double cosets with respect to $B$:

$$G = \bigsqcup_{w \in W} B w B.$$  

(More accurately, we should write $B n_w B$ where $n_w$ is a representative of $w \in W$ in $N$. But, since any two representatives of $w$ lie in the same coset of $T \subseteq B$, the double coset $B n_w B$ does not depend on the choice of the representative.)

Furthermore, the proof of the simplicity of the Chevalley groups and their twisted analogues is most economically performed using the simplicity criterion for abstract groups with a $BN$-pair in Bourbaki, [25, Chapter IV, §2, no. 7].

Groups with a $BN$-pair play an important rôle in finite group theory. In fact, it is known that every finite simple group possesses a $BN$-pair, except for the cyclic groups of prime order, the alternating groups of degree $\geq 5$, and the 26 sporadic simple groups; see Gorenstein et al., [96]. Given a finite group $G$ with a $BN$-pair, the irreducible factors of the Weyl group $W$ are of type $A_n$, $B_n$, $D_n$, $G_2$, $F_4$, $E_6$, $E_7$, $E_8$ or $I_2(8)$. This follows from the classification by Tits, [181] (rank $\geq 3$), Hering, Kantor, Seitz, [98,117] (rank 1) and Fong and Seitz, [81] (rank 2). Note that there is only one case where $W$ is not crystallographic: this is the case where $W$ has a component of type $I_2(8)$ (the dihedral group of order 16), which corresponds to the twisted groups of type $F_4$ discovered by Ree (see Carter, [42], or Steinberg, [176]).

In another direction, $BN$-pairs with infinite Weyl groups arise naturally in the theory of $p$-adic groups; see Iwahori and Matsumoto, [109].

### 2.17. Connected reductive algebraic groups

Here, we assume that the reader has some familiarity with the theory of linear algebraic groups; see Borel, [23], Humphreys, [106], or Springer, [171]. Let $G$ be a connected reductive algebraic group over an algebraically closed field $K$. Let $B \subseteq G$ be a Borel subgroup. Then we have a semidirect product decomposition $B = UT$ where $U$ is the unipotent radical of $B$ and $T$ is a maximal torus. Let $N = N_G(T)$, the normaliser of $T$ in $G$. Then the groups $B, N$ form a $BN$-pair in $G$; furthermore, $W$ must be a finite Weyl group (and not just a Coxeter group as for general groups with a $BN$-pair). This is a deep, important result whose proof goes back to Chevalley, [47]; detailed expositions can be found in the monographs by Borel, [23, Chapter IV, 14.15], Humphreys, [106, §29.1], or Springer, [171, Chapter 8].

For example, in $G = GL_n(K)$, the subgroup $B$ of all upper triangular matrices is a Borel subgroup by the Lie–Kolchin theorem (see, for example, Humphreys, [106, 17.6]). Furthermore, we have a semidirect product decomposition $B = UT$ where $U \subseteq B$ is the normal subgroup consisting of all upper triangular matrices with 1 on the diagonal and $T$ is the group of all diagonal matrices in $B$. Since $K$ is infinite, it is easily checked that $N = N_G(T)$, the group of all monomial matrices.
Returning to the general case, let us consider the Bruhat cells $BwB$ ($w \in W$). These are locally closed subsets of $G$ since they are orbits of $B \times B$ on $G$ under left and right multiplication. The Zariski closure of $BwB$ is given by

$$BwB = \bigcup_{y \in S(w)} ByB.$$  

This yields the promised geometric description of the Bruhat–Chevalley order $\leq$ on $W$ (as defined in the remarks following Theorem 2.2.) The proof (see, for example, Springer, [171, §8.5]) relies in an essential way on the fact that $G/B$ is a projective variety.

3. Braid groups

3.1. The braid group of a complex reflection group

For a complex reflection group $W \leq \text{GL}(V)$, $V = \mathbb{C}^n$, denote by $\mathcal{A}$ the set of its reflecting hyperplanes in $V$. The topological space $V_{\text{reg}} := V \setminus \bigcup_{H \in \mathcal{A}} H$ is (pathwise) connected in its inherited complex topology. For a fixed base point $x_0 \in V_{\text{reg}}$ we define the pure braid group of $W$ as the fundamental group $P(W) := \pi_1(V_{\text{reg}}, x_0)$.

Now $W$ acts on $V_{\text{reg}}$, and by the theorem of Steinberg (Theorem 1.6) the covering $\bar{\gamma} : V_{\text{reg}} \rightarrow V_{\text{reg}}/W$ is Galois, with group $W$. This induces a short exact sequence

$$1 \rightarrow P(W) \rightarrow B(W) \rightarrow W \rightarrow 1$$

for the braid group $B(W) := \pi_1(V_{\text{reg}}/W, \bar{x_0})$ of $W$.

If $W = S_n$ in its natural permutation representation, the group $B(W)$ is just the classical Artin braid group on $n$ strings, [4].

We next describe some natural generators of $B(W)$. Let $H \in \mathcal{A}$ be a reflecting hyperplane. Let $x_H \in H$ and $r > 0$ such that the open ball $B(x_H, 2r)$ around $x_H$ does not intersect any other reflecting hyperplane and $x_0 \notin B(x_H, 2r)$. Choose a path $\gamma : [0, 1] \rightarrow V$ from the base point $x_0$ to $x_H$, with $\gamma(t) \in V_{\text{reg}}$ for $t < 1$. Let $t_0$ be minimal subject to $\gamma(t) \in B(x_H, r)$ for all $t > t_0$. Then $\gamma' := \gamma(t/t_0)$ is a path from $x_0$ to $\gamma(t_0)$. Then

$$\lambda : [0, 1] \rightarrow B(x_H, 2r), \quad t \mapsto \gamma(t_0) \exp(2\pi it/e_H),$$

where $e_H = |W_H|$ is the order of the fixator of $H$ in $W$, defines a closed path in the quotient $V_{\text{reg}}/W$. The homotopy class in $B(W)$ of the composition $\gamma' \circ \lambda \circ \gamma'^{-1}$ is then called a braid reflection (see Broué, [33]) or generator of the monodromy around $H$. Its image in $W$ is a reflection $s_H$ generating $W_H$, with non-trivial eigenvalue $\exp(2\pi i/e_H)$. It can be
shown that $B(W)$ is generated by all braid reflection, when $H$ varies over the reflecting hyperplanes of $W$ (Broué, Malle and Rouquier, [38, Theorem 2.17]).

Assume from now on that $W$ is irreducible. Recall the definition of $N$, $N^*$ in Sections 1.2 and 1.7 as the number of reflections respectively of reflecting hyperplanes. The following can be shown without recourse to the classification of irreducible complex reflection groups:

3.2. THEOREM (Bessis, [11]). Let $W \leq GL_n(\mathbb{C})$ be an irreducible complex reflection group with braid group $B(W)$. Let $d$ be a degree of $W$ which is a regular number for $W$ and let $r := (N + N^*)/d$. Then $r \in \mathbb{N}$, and there exists a subset $S = \{s_1, \ldots, s_r\} \subset B(W)$ with:

(i) $s_1, \ldots, s_r$ are braid reflections, so their images $s_1, \ldots, s_r \in W$ are reflections.

(ii) $S$ generates $B(W)$, and hence $S := \{s_1, \ldots, s_r\}$ generates $W$.

(iii) There exists a finite set $\mathcal{R}$ of relations of the form $w_1 = w_2$, where $w_1, w_2$ are words of equal length in $s_1, \ldots, s_r$, such that $(s_1, \ldots, s_r \mathcal{R})$ is a presentation for $B(W)$.

(iv) Let $e_s$ denote the order of $s \in S$. Then $(s_1, \ldots, s_r \mathcal{R} ; s^{e_s} = 1 \forall s \in S)$ is a presentation for $W$, where now $\mathcal{R}$ is viewed as a set of relations on $S$.

(v) $(s_1 \cdots s_r)^d$ is central in $B(W)$ and lies in $P(W)$.

(vi) The product $c := s_1 \cdots s_r$ is a $\zeta := \exp(2\pi i/d)$-regular element of $W$ (hence has eigenvalues $\zeta^{-m_1}, \ldots, \zeta^{-m_r}$).

It follows from the classification (see table 1) that there always exists a regular degree. In many cases, for example if $W$ is well-generated, the number $(N + N^*)/r$ is regular, when $r$ is chosen as the minimal number of generating reflections for $W$ (so $n \leq r \leq n + 1$). Thus, in those cases $B(W)$ is finitely presented on the same minimal number of generators as $W$. Under the assumptions (i) or (ii) of Theorem 1.16, the largest degree $d_n$ is regular, whence Theorem 1.16(iv) is a consequence of the previous theorem.

At present, presentations of the type described in Theorem 3.2 have been found for all but six irreducible types, by case-by-case considerations, see Bannai, [5], Naruki, [146], Broué, Malle and Rouquier, [38]. For the remaining six groups, conjectural presentations have been found by Bessis and Michel using computer calculations.

For the case of real reflection groups, Brieskorn, [28], and Deligne, [61], determined the structure of $B(W)$ by a nice geometric argument. They show that the generators in Theorem 3.2 (with $r = n$) can be taken as suitable preimages of the Coxeter generators, and the relations $\mathcal{R}$ as the Coxeter relations. For the case of $W(A_n) = \mathfrak{S}_{n+1}$ of the classical braid group, this was first shown by Artin, [4].

A topological space $X$ is called $K(\pi, 1)$ if all homotopy groups $\pi_i(X)$ for $i \neq 1$ vanish. The following is conjectured by Arnol’d to be true for all irreducible complex reflection groups:

3.3. THEOREM. Assume that $W$ is not of type $G_i$, $i \in \{24, 27, 29, 31, 33, 34\}$. Then $V^{\text{reg}}$ and $V^{\text{reg}}/W$ are $K(\pi, 1)$-spaces.

This was proved by a general argument for Coxeter groups by Deligne, [61], after Fox and Neuwirth, [82], showed it for type $A_n$ and Brieskorn, [29], for those of type different
from $H_3$, $H_4$, $E_6$, $E_7$, $E_8$. For the non-real Shephard groups (non-real groups with Coxeter braid diagrams), it was proved by Orlik and Solomon, [158]. The case of the infinite series $G(de, e, r)$ has been solved by Nakamura, [145]. In that case, there exists a locally trivial fibration

$$V^{\text{reg}}(G(de, e, n)) \to V^{\text{reg}}(G(de, e, n - 1)),$$

with fiber isomorphic to $\mathbb{C}$ minus $m(de, e, n)$ points, where

$$m(de, e, n) := \begin{cases} (n - 1)de + 1 & \text{for } d \neq 1, \\ (n - 1)(e - 1) & \text{for } d = 1. \end{cases}$$

This induces a split exact sequence

$$1 \to F_m \to P(G(de, e, n)) \to P(G(de, 1, n - 1)) \to 1$$

for the pure braid group, with a free group $F_m$ of rank $m = m(de, e, n)$. In particular, the pure braid group has the structure of an iterated semidirect product of free groups (see Broué, Malle and Rouquier, [38, Proposition 3.37]).

### 3.4. The centre and regular elements

Denote by $\pi$ the class in $P(W)$ of the loop

$$[0, 1] \to V^{\text{reg}}, \quad t \mapsto x_0 \exp(2\pi it).$$

Then $\pi$ lies in the centre $Z(P(W))$ of the pure braid group. Furthermore,

$$[0, 1] \to V^{\text{reg}}, \quad t \mapsto x_0 \exp(2\pi it/|Z(W)|),$$

defines a closed path in $V^{\text{reg}}/W$, so an element $\beta$ of $B(W)$, which is again central. Clearly $\pi = \beta|Z(W)|$.

The following was shown independently by Brieskorn and Saito, [30], and Deligne, [61], for Coxeter groups, and by Broué, Malle and Rouquier, [38, Theorem 2.24], for the other groups:

### 3.5. Theorem. Assume that $W$ is not of type $G_i$, $i \in \{24, 27, 29, 31, 33, 34\}$. Then the centre of $B(W)$ is infinite cyclic generated by $\beta$, the centre of $P(W)$ is infinite cyclic generated by $\pi$, and the exact sequence (1) induces an exact sequence

$$1 \to Z(P(W)) \to Z(B(W)) \to Z(W) \to 1.$$
In their papers, Brieskorn and Saito, [30], and Deligne, [61], also solve the word problem and the conjugation problem for braid groups attached to real reflection groups.

For each $H \in A$ choose a linear form $\alpha_H : V \to \mathbb{C}$ with kernel $H$. Let $e_H := |W_H|$, the order of the minimal parabolic subgroup fixing $H$. The discriminant of $W$, defined as

$$\delta := \delta(W) := \prod_{H \in A} \alpha^e_H,$$

is then a $W$-invariant element of the symmetric algebra $S(V^*)$ of $V^*$, well-defined up to non-zero scalars (Cohen, [50, 1.8]). It thus induces a continuous function $\delta : V_{\text{reg}}/W \to \mathbb{C}^\times$, hence by functoriality a group homomorphism $\pi_1(\delta) : B(W) \to \pi_1(\mathbb{C}^\times, 1) \cong \mathbb{Z}$. For $b \in B(W)$ let $l(b) := \pi_1(\delta)(b)$ denote the length of $b$. For example, every braid reflection $s$ has length $l(s) = 1$, and we have

$$l(b) = (N + N^*)/|Z(W)|$$

and hence

$$l(\pi) = N + N^*$$

by Broué, Malle and Rouquier, [38, Corollary 2.21].

The elements $b \in B(W)$ with $l(b) \geq 0$ form the braid monoid $B^+(W)$. A $d$-th root of $\pi$ is by definition an element $w \in B^+$ with $w^d = \pi$.

Let $d$ be a regular number for $W$, and $w$ a $d$-th root of $\pi$. Assume that the image $w$ of $w$ in $W$ is $\xi$-regular for some $d$-th root of unity $\xi$ (in the sense of 1.9). (This is, for example, the case if $W$ is a Coxeter group by Broué and Michel, [39, Theorem 3.12].) By Theorem 1.11(ii) the centraliser $W(w) := C_W(w)$ is a reflection group on $V(w, \xi)$, with reflecting hyperplanes the intersections of $V(w, \xi)$ with the hyperplanes in $A$ by Theorem 1.12(i). Thus the hyperplane complement of $W(w)$ on $V(w, \xi)$ is just $V^\text{reg}(w)$ := $V_{\text{reg}} \cap V(w, \xi)$. Assuming that the base point $x_0$ has been chosen in $V^\text{reg}(w)$, this defines natural maps

$$P(W(w)) \to P(W) \quad \text{and} \quad \psi_w : B(W(w)) \to B(W).$$

By Broué and Michel, [39, 3.4], the image of $B(W(w))$ in $B(W)$ centralises $w$. It is conjectured (see Bessis, Digne and Michel, [12, Conjecture 0.1]) that $\psi_w$ defines an isomorphism $B(W(w)) \cong C_{B(W)}(w)$. The following partial answer is known:

3.6. Theorem (Bessis, Digne and Michel, [12, Theorem 0.2]). Let $W$ be an irreducible reflection group of type $\mathfrak{S}_n$, $G(d, 1, n)$ or $G_i$, $i \in \{4, 5, 8, 10, 16, 18, 25, 26, 32\}$, and let $w \in W$ be regular. Then $\psi_w$ induces an isomorphism $B(W(w)) \cong C_{B(W)}(w)$.

This has also been proved by Michel, [142, Corollary 4.4], in the case that $W$ is a Coxeter group and $w$ acts on $W$ by a diagram automorphism. The injectivity of $\psi_w$ was shown for all but finitely many types of $W$ by Bessis, [10, Theorem 1.3].

The origin of Artin’s work on the braid group associated with the symmetric group lies in the theory of knots and links. We shall now briefly discuss this connection and explain the construction of the “HOMFLY-PT” invariant of knots and links (which includes the famous Jones polynomial as a special case). We follow the exposition in Geck and Pfeiffer, [95, §4.5].
3.7. Knots and links, Alexander and Markov theorem

If \( n \) is a positive integer, an oriented \( n \)-link is an embedding of \( n \) copies of the interval \([0, 1] \subset \mathbb{R}\) into \( \mathbb{R}^3 \) such that 0 and 1 are mapped to the same point (the orientation is induced by the natural ordering of \([0, 1]\)); a 1-link is also called a knot. We are only interested in knots and links modulo isotopy, i.e., homeomorphic transformations which preserve the orientation. We refer to Birman, [16], Crowell and Fox, [57], or Burde and Zieschang, [40], for precise versions of the above definitions.

By Artin’s classical interpretation of \( B(\mathfrak{S}_n) \) as the braid group on \( n \) strings, each generator of \( B(\mathfrak{S}_n) \) can be represented by oriented diagrams as indicated below; writing any \( g \in B(\mathfrak{S}_n) \) as a product of the generators and their inverses, we also obtain a diagram for \( g \), by concatenating the diagrams for the generators. “Closing” such a diagram by joining the end points, we obtain the plane projection of an oriented link in \( \mathbb{R}^3 \):

\[
\begin{align*}
\text{s}_i &\quad \begin{array}{cccccccc}
1 & 2 & i & i+1 & \cdots & \cdots & n-1 & n \\
1 & 2 & i & i+1 & \cdots & \cdots & n-1 & n
\end{array} \\
\text{s}_i^{-1} &\quad \begin{array}{cccccccc}
1 & 2 & i & i+1 & \cdots & \cdots & n-1 & n \\
1 & 2 & i & i+1 & \cdots & \cdots & n-1 & n
\end{array}
\end{align*}
\]

By Alexander’s theorem (see Birman, [16], or, for a more recent proof, Vogel, [186]), every oriented link in \( \mathbb{R}^3 \) is isotopic to the closure of an element in \( B(\mathfrak{S}_n) \), for some \( n \geq 1 \). The question of when two links in \( \mathbb{R}^3 \) are isotopic can also be expressed algebraically. For this purpose, we consider the infinite disjoint union

\[
B_{\infty} := \bigsqcup_{n \geq 1} B(\mathfrak{S}_n).
\]

Given \( g, g' \in B_{\infty} \), we write \( g \sim g' \) if one of the following relations is satisfied:

(I) We have \( g, g' \in B(\mathfrak{S}_n) \) and \( g' = x^{-1}gx \) for some \( x \in B(\mathfrak{S}_n) \).

(II) We have \( g \in B(\mathfrak{S}_n) \), \( g' \in B(\mathfrak{S}_{n+1}) \) and \( g' = g \mathfrak{s}_n \) or \( g' = g \mathfrak{s}_n^{-1} \).

The above two relations are called Markov relations. By a classical result due to Markov (see Birman, [16], or, for a more recent proof, Traczyk, [183]), two elements of \( B_{\infty} \) are equivalent under the equivalence relation generated by \( \sim \) if and only if the corresponding links obtained by closure are isotopic. Thus, to define an invariant of oriented links is the same as to define a map on \( B_{\infty} \) which takes equal values on elements \( g, g' \in B_{\infty} \) satisfying (I) or (II).

We now consider the Iwahori–Hecke algebra \( H_\mathbb{C}(\mathfrak{S}_n) \) of the symmetric group \( \mathfrak{S}_n \) over \( \mathbb{C} \). By definition, \( H_\mathbb{C}(\mathfrak{S}_n) \) is a quotient of the group algebra of \( B(\mathfrak{S}_n) \), where we factor...
Reflection groups

by an ideal generated by certain quadratic relations depending on two parameters $u, v \in \mathbb{C}$. This is done such that

$$T_{s_i}^2 = uT_1 + vT_{s_i} \quad \text{for } 1 \leq i \leq n - 1,$$

where $T_{s_i}$ denotes the image of the generator $s_i$ of $B(\mathcal{S}_n)$ and $T_1$ denotes the identity element. For each $w \in \mathcal{S}_n$, we have a well-defined element $T_w$ such that

$$T_w T_{w'} = T_{ww'} \quad \text{whenever } l(ww') = l(w) + l(w').$$

This follows easily from Matsumoto’s Theorem 2.2. In fact, one can show that the elements $\{T_w \mid w \in \mathcal{S}_n\}$ form a $\mathbb{C}$-basis of $H_\mathbb{C}(\mathcal{S}_n)$. (For more details, see the chapter on Hecke algebras.) The map $w \mapsto T_w$ ($w \in \mathcal{S}_n$) extends to a well-defined algebra homomorphism from the group algebra of $B(\mathcal{S}_n)$ over $\mathbb{C}$ onto $H_\mathbb{C}(\mathcal{S}_n)$. Furthermore, the inclusion $\mathcal{S}_{n-1} \subseteq \mathcal{S}_n$ also defines an inclusion of algebras $H_\mathbb{C}(\mathcal{S}_{n-1}) \subseteq H_\mathbb{C}(\mathcal{S}_n)$.

3.8. Theorem (Jones, Ocneanu, [112]). There is a unique family of $\mathbb{C}$-linear maps $\tau_n : H_\mathbb{C}(\mathcal{S}_n) \to \mathbb{C}$ ($n \geq 1$) such that the following conditions hold:

(M1) $\tau_1(T_1) = 1$;

(M2) $\tau_{n+1}(hT_{s_n}) = \tau_{n+1}(hT_{s_n}^{-1}) = \tau_n(h)$ for all $n \geq 1$ and $h \in H_\mathbb{C}(\mathcal{S}_n)$;

(M3) $\tau_n(hh') = \tau_n(h'h)$ for all $n \geq 1$ and $h, h' \in H_\mathbb{C}(\mathcal{S}_n)$.

Moreover, we have $\tau_{n+1}(h) = v^{-1}(1-u)\tau_n(h)$ for all $n \geq 1$ and $h \in H_\mathbb{C}(\mathcal{S}_n)$.

In [112], Jones works with an Iwahori–Hecke algebra of $\mathcal{S}_n$ where the parameters are related by $v = u - 1$. The different formulation above follows a suggestion by J. Michel. It results in a simplification of the construction of the link invariants below. (The simplification arises from the fact that, due to the presence of two different parameters in the quadratic relations, the “singularities” mentioned in [112, p. 349, notes (1)] simply disappear.) Generalisations of Theorem 3.8 to types $G(d, 1, n)$ and $D_n$ have been found in Geck and Lambropoulou, [92], Lambropoulou, [125], and Geck, [87].

3.9. The HOMFLY-PT polynomial

We can now construct a two-variable invariant of oriented knots and links as follows. Consider an oriented link $L$ and assume that it is isotopic to the closure of $g \in B(\mathcal{S}_n)$ for $n \geq 1$. Then we set

$$X_L(u, v) := \tau_n(\bar{g}) \in \mathbb{C} \quad \text{with } \tau_n \text{ as in Theorem 3.8}.$$

Here, $\bar{g}$ denotes the image of $g$ under the natural map $\mathbb{C}[B(\mathcal{S}_n)] \to H_\mathbb{C}(\mathcal{S}_n)$, $w \mapsto T_w$ ($w \in \mathcal{S}_n$). It is easily checked that $X_L(u, v)$ can be expressed as a Laurent polynomial in $u$ and $v$; the properties (M2) and (M3) make sure that $\tau_n(\bar{g})$ does not depend on the choice of $g$. If we make the change of variables $u = t^2$ and $v = tx$, we can identify the above invariant with the HOMFLY-PT polynomial $P_L(t, x)$ discovered by Freyd et al., [83], and
Przytycki and Traczyk, [161]; see also Jones, [112, (6.2)]. Furthermore, the *Jones polynomial* $J_L(t)$ is obtained by setting $u = t^2$, $v = \sqrt{t(t-1)}$ (see [112, §11]). Finally, setting $u = 1$ and $v = \sqrt{t - 1}/\sqrt{t}$, we obtain the classical *Alexander polynomial* $A_L(t)$ whose definition can be found in Crowell and Fox, [57].

For a survey about recent developments in the theory of knots and links, especially since the discovery of the Jones polynomial, see Birman, [17].

### 3.10. Further aspects of braid groups

One of the old problems concerning braid groups is the question whether or not they are linear, i.e., whether there exists a faithful linear representation on a finite-dimensional vector space. Significant progress has been made recently on this problem. Krammer, [122], and Bigelow, [15], proved that the classical Artin braid group is linear. Then Digne, [69], and Cohen and Wales, [53], extended this result and showed that all Artin groups of crystallographic type have a faithful representation of dimension equal to the number of reflections of the associated Coxeter group.

On the other hand, there is one particular representation of the braid group associated with $S_n$, the so-called Burau representation (see Birman, [16], for which it has been a longstanding problem to determine for which values of $n$ it is faithful. Moody, [143], showed that it is not faithful for $n \geq 10$; this bound was improved by Long and Paton, [133], to 6. Recently, Bigelow, [14], showed that the Burau representation is not faithful already for $n = 5$. (It is an old result of Magnus and Peluso that the Burau representation is faithful for $n = 3$.)

In a different direction, Deligne’s and Brieskorn–Saito’s solution of the word and conjugacy problem in braid groups led to new developments in combinatorial group and monoid theory; see, for example, Dehornoy and Paris, [60], and Dehornoy, [59].

### 4. Representation theory

In this section we report about the representation theory of finite complex reflection groups.

#### 4.1. Fields of definition

Let $W$ be a finite complex reflection group on $V$. Let $K_W$ denote the character field of the reflection representation of $W$, that is, the field generated by the traces $tr_V(w)$, $w \in W$. It is easy to see that the reflection representation can be realised over $K_W$ (see, for example, [7, Proposition 7.1.1]). But we have a much stronger statement:

#### 4.2. Theorem (Benard, [6], Bessis, [9]). Let $W$ be a complex reflection group. Then the field $K_W$ is a splitting field for $W$.

The only known proof for this result is case-by-case, treating the reflection groups according to the Shephard–Todd classification.
The field $K_W$ has a nice description at least in the case of well-generated groups, that is, irreducible groups generated by $\dim(V)$ reflections. In this case, the largest degree $d_n$ of $W$ is regular, so there exists an element $c := s_1 \cdots s_n$ as in theorem 3.2(vi) of Bessis, called a Coxeter element of $W$, with eigenvalues $\zeta^{-m_1}, \ldots, \zeta^{-m_n}$ in the reflection representation, where $\zeta := \exp(2\pi i / d_n)$ and the $m_i$ are the exponents of $W$.

If $W$ is a real reflection group, then $W$ is a Coxeter group associated with some Coxeter matrix $M$ (see Theorem 2.4) and we have

$$K_W = \mathbb{Q}(\cos(2\pi / m_s) \mid s, t \in S) \subset \mathbb{R}.$$ 

In particular, this shows that $K_W = \mathbb{Q}$ if $W$ is a finite Weyl group.

For well-generated irreducible complex reflection groups $W \leq \text{GL}(V)$, the field of definition $K_W$ is generated over $\mathbb{Q}$ by the coefficients of the characteristic polynomial on $V$ of a Coxeter element, see Malle, [140, Theorem 7.1]. This characterisation is no longer true for non-well generated reflection groups.

### 4.3. Macdonald–Lusztig–Spaltenstein induction

Let $W$ be a complex reflection group on $V$. Recall from Section 1.7 the definition of the fake degree $R_\chi$ of an irreducible character $\chi \in \text{Irr}(W)$. The $b$-invariant $b_\chi$ of $\chi$ is defined as the order of vanishing of $R_\chi$ at $x = 0$, that is, as the minimum of the exponents $e_i(\chi)$ of $\chi$. The coefficient of $x^{b_\chi}$ in $R_\chi$ is denoted by $\gamma_\chi$.

**THEOREM (Macdonald, [136], Lusztig and Spaltenstein, [135]).** Let $W$ be a complex reflection group on $V$, $W'$ a reflection subgroup (on $V'_1 := V/C_V(W')$). Let $\psi$ be an irreducible character of $W'$ such that $\gamma_\psi = 1$. Then $\text{Ind}_{W'}^W(\psi)$ has a unique irreducible constituent $\chi \in \text{Irr}(W)$ with $b_\chi = b_\psi$. This satisfies $\gamma_\chi = 1$. All other constituents have $b$-invariant bigger than $b_\psi$.

The character $\chi \in \text{Irr}(W)$ in Theorem 4.4 is called the $j$-induction $j_{W'}^W(\psi)$ of the character $\psi \in \text{Irr}(W')$. Clearly, $j$-induction is transitive; it is also compatible with direct products.

An important example of characters $\psi$ with $\gamma_\psi = 1$ is given by the determinant character $\det_V : W \to \mathbb{C}^\times$ of a complex reflection group (see Geck and Pfeiffer, [95, Theorem 5.2.10]).

### 4.5. Irreducible characters

There is no general construction of all irreducible representations of a complex reflection group known. Still, we have the following partial result:

**THEOREM (Steinberg).** Let $W \leq \text{GL}(V)$ be an irreducible complex reflection group. Then the exterior powers $\Lambda^i(V)$, $1 \leq i \leq \dim V$, are irreducible, pairwise non-equivalent representations of $W$. 
A proof in the case of well-generated groups can be found in Bourbaki, [25, V, §2, Example 3(d)], and Kane, [114, Theorem 24-3 A], for example. The general case then follows with corollary 1.15.

We now give some information on the characters of individual reflection groups. The irreducible characters of the symmetric group $S_n$ were determined by Frobenius, [84], see also Macdonald, [137], and Fulton, [85]. Here we follow the exposition in Geck and Pfeiffer, [95, 5.4].

Let $\lambda = (\lambda_1, \ldots, \lambda_r) \vdash n$ be a partition of $n$. The corresponding Young subgroup $S_\lambda$ of $S_n$ is the common setwise stabiliser $\{1, \ldots, \lambda_1\}, \{\lambda_1 + 1, \ldots, \lambda_1 + \lambda_2\}, \ldots$, abstractly isomorphic to $S_\lambda = S_{\lambda_1} \times \cdots \times S_{\lambda_r}$. This is a parabolic subgroup of $S_n$ in the sense of Section 1.5. For each partition we have the two induced characters $\pi_\lambda := \text{Ind}_{S_n}^{S_\lambda} (1_{\lambda_1} \# \cdots \# 1_{\lambda_r})$, $\theta_\lambda := \text{Ind}_{S_n}^{S_\lambda} (e_{\lambda_1} \# \cdots \# e_{\lambda_r})$.

Then $\pi_\lambda$ and $\theta_\lambda^*$ have a unique irreducible constituent $\chi_\lambda \in \text{Irr}(S_n)$ in common, where $\lambda^*$ denotes the partition dual to $\lambda$. This constituent can also be characterised in terms of $j$-induction as $\chi_\lambda = \text{j}_{S_n}^{S_\lambda} (1_{\lambda_1} \# \cdots \# 1_{\lambda_r})$.

Then the $\chi_\lambda$ are mutually distinct and exhaust the irreducible characters of $S_n$, so $\text{Irr}(S_n) = \{\chi_\lambda \mid \lambda \vdash n\}$. From the above construction it is easy to see that all $\chi_\lambda$ are afforded by rational representations.

The construction of the irreducible characters of the imprimitive group $G(d, 1, n)$ goes back at least to Osima, [159] (see also Read, [164], Hughes, [102], Bessis, [9]) via their abstract structure as wreath product $C_d \wr S_n$. Let us fix $d \geq 2$ and write $W_n := G(d, 1, n)$. A $d$-tuple $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_{d-1})$ of partitions $\alpha_i \vdash n_i$ with $\sum n_i = n$ is called a $d$-partition of $n$. We denote by $W_\alpha$ the natural subgroup $W_n \times \cdots \times W_{n_{d-1}}$ of $W_n$, where $\alpha_j \vdash n_j$, corresponding to the Young subgroup $S_{\alpha_0} \times \cdots \times S_{\alpha_{d-1}}$ of $S_n$. Via the natural projection $W_n_j \to S_{\alpha_j}$ the characters of $S_{\alpha_j}$ may be regarded as characters of $W_n_j$. Thus each $\alpha_j$ defines an irreducible character $\chi_{\alpha_j}$ of $W_n$. For any $m$, let $\zeta_d : W_m \to \mathbb{C}^\times$ be the linear character defined by $\zeta_d(t_i) = \exp(2\pi i/d)$, $\zeta_d(t_i) = 1$ for $i > 1$, with the standard generators $t_i$ from 1.13. Then for any $d$-partition $\alpha$ of $n$ we can define a character $\chi_\alpha$ of $W_n$ as the induction of the exterior product $\chi_\alpha := \text{Ind}_{W_n}^{W_\alpha} (\chi_{\alpha_0} \# (\chi_{\alpha_1} \otimes \zeta_d) \# \cdots \# (\chi_{\alpha_{d-1}} \otimes \zeta_d^{d-1}))$.

By Clifford-theory $\chi_\alpha$ is irreducible, $\chi_\alpha \neq \chi_\beta$ if $\alpha \neq \beta$, and all irreducible characters of $W_n$ arise in this way, so $\text{Irr}(W_n) = \{\chi_\alpha \mid \alpha = (\alpha_0, \ldots, \alpha_{d-1}) \vdash_d n\}$.

We describe the irreducible characters of $G(de, e, n)$ in terms of those of $W_n := G(de, 1, n)$. Recall that the imprimitive reflection group $G(de, e, n)$ is generated by the
Reflection groups

371

reflections $t_2, \tilde{t}_2 := t_1^{-1}t_2t_1, t_3, \ldots, t_n$, and $t_1^e$. Denote by $\pi$ the cyclic shift on $de$-partitions of $n$, i.e.,

$$\pi(\alpha_0, \ldots, \alpha_{de-1}) = (\alpha_1, \ldots, \alpha_{de-1}, \alpha_0).$$

By definition we then have $\chi_{\pi(\alpha)} \otimes \zeta_{de} = \chi_\alpha$. Let $s_e(\alpha)$ denote the order of the stabiliser of $\alpha$ in the cyclic group $\langle \pi_d \rangle$. Then upon restriction to $G(de, e, n)$ the irreducible character $\chi_\alpha$ of $W_n$ splits into $s_e(\alpha)$ different irreducible constituents, and this exhausts the set of irreducible characters of $G(de, e, n)$. More precisely, let $\alpha$ be a $de$-partition of $n$ with $\tilde{e} := s_e(\alpha)$, $W_{\alpha, e} := W_\alpha \cap G(de, e, n)$, and $\psi_\alpha$ the restriction of $\chi_\alpha$ to $W_{\alpha, e}$. Then $\psi_\alpha$ is invariant under the element $\sigma := (t_2 \cdots t_n)^{n/\tilde{e}}$ (note that $\tilde{e} = s_e(\alpha)$ divides $n$), and it extends to the semidirect product $W_{\alpha, e} \langle \sigma \rangle$. The different extensions of $\psi_\alpha$ induced to $G(de, e, n)$ then exhaust the irreducible constituents of the restriction of $\chi_\alpha$ to $G(de, e, n)$. Thus, we may parametrise $\text{Irr}(G(de, e, n))$ by $de$-partitions of $n$ up to cyclic shift by $\pi_d$ in such a way that any $\alpha$ stands for $s_e(\alpha)$ different characters.

In order to describe the values of the irreducible characters we need the following definitions. We identify partitions with their Young diagrams. A $d$-partition $\alpha$ is called a hook if it has just one non-empty part, which is a hook (i.e., does not contain a $2 \times 2$-block). The position of the non-empty part is then denoted by $\tau(\alpha)$. If $\alpha, \beta$ are $d$-partitions such that $\beta_i$ is contained in $\alpha_i$ for all $0 \leq i \leq d - 1$, then $\alpha \setminus \beta$ denotes the $d$-partition $(\alpha_i \setminus \beta_i | 0 \leq i \leq d - 1)$, where $\alpha_i \setminus \beta_i$ is the set theoretic difference of $\alpha_i$ and $\beta_i$. If $\alpha \setminus \beta$ is a hook, we denote by $l_{\alpha \setminus \beta}$ the number of rows of the hook $(\alpha \setminus \beta)_{\tau(\alpha \setminus \beta)}$ minus 1. With these notations the values of the irreducible characters of $G(d, 1, n)$ can be computed recursively with a generalised Murnaghan–Nakayama rule (see also Stembridge, [179]):

4.7. Theorem (Osima, [159]). Let $\alpha$ and $\gamma$ be $d$-partitions of $n$, let $m$ be a part of $\gamma_t$ for some $1 \leq t \leq d$, and denote by $\gamma'$ the $d$-partition of $n - m$ obtained from $\gamma$ by deleting the part $m$ from $\gamma_t$. Then the value of the irreducible character $\chi_\alpha$ on an element of $G(d, 1, n)$ with cycle structure $\gamma$ is given by

$$\chi_\alpha(\gamma) = \sum_{\alpha \setminus \beta \vdash dm} \zeta_d^{st} (-1)^{l_{\beta}} \chi_\beta(\gamma')$$

where the sum ranges over all $d$-partitions $\beta$ of $n - m$ such that $\alpha \setminus \beta$ is a hook and where $s = \tau(\alpha \setminus \beta)$.

An overview of the irreducible characters of the exceptional groups (and of their projective characters) is given in Humphreys, [107]. All character tables of irreducible complex reflection groups are also available in the computer algebra system CHEVIE, [89].
4.8. Fake degrees

The fake degrees (introduced in Section 1.7) of all complex reflection groups are known. For the symmetric groups, they were first determined by Steinberg, [174], as generic degrees of the unipotent characters of the general linear groups over a finite field. From that result, the fake degrees of arbitrary imprimitive reflection groups can easily be derived (see Malle, [138, Bem. 2.10 and 5.6]).

4.9. Theorem (Steinberg, [174], Lusztig, [134]). The fake degrees of the irreducible complex reflection groups $G(d,e,n)$ are given as follows:

(i) Let $\chi \in \text{Irr}(G(d,1,n))$ be parameterised by the $d$-partition $(\alpha_0, \ldots, \alpha_{d-1})$, where $\alpha_i = (\alpha_{i1} \geq \cdots \geq \alpha_{imi}) \vdash n_i$, and let $(S_0, \ldots, S_{d-1})$, where $S_i = (\alpha_{i1} + m_i - 1, \ldots, \alpha_{imi})$, denote the corresponding tuple of $\beta$-numbers. Then

$$R_\chi = \prod_{i=1}^n (x^{id} - 1) \prod_{i=0}^{d-1} \frac{\Delta(S_i, x^d) x^{n_i}}{\Theta(S_i, x^d) x^{d(n_i^2) + d(n_i^2) + \cdots}},$$

where, for a finite subset $S \subset \mathbb{N}$,

$$\Delta(S, x) := \prod_{\lambda, \lambda' \in S} (x^{i \lambda} - x^{i \lambda'}), \quad \Theta(S, x) := \prod_{\lambda \in S} \prod_{h=1}^\lambda (x^h - 1).$$

(ii) The fake degree of $\chi \in \text{Irr}(G(de,e,n))$ is obtained from the fake degrees in $G(de,1,n)$ as

$$R_\chi = \frac{x^{nd} - 1}{x^{nde} - 1} \sum_{\psi \in \text{Irr}(G(de,1,n))} \langle \chi, \psi \rangle_{G(de,e,n)} R_\psi.$$

For exceptional complex reflection groups, the fake degrees can easily be computed, for example in the computer algebra system CHEVIE, [89]. In the case of exceptional Weyl groups, they were first studied by Beynon and Lusztig, [13].

The fake degrees of reflection groups satisfy a remarkable palindromicity property:

4.10. Theorem (Opdam, [151,152], Malle, [140]). Let $W$ be a complex reflection group. There exists a permutation $\delta$ on $\text{Irr}(W)$ such that for every $\chi \in \text{Irr}(W)$ we have

$$R_\chi(x) = x^c R_{\delta(\chi)}(x^{-1}),$$

where $c = \sum_r (1 - \chi(r)/\chi(1))$, the sum running over all reflections $r \in W$.

The interest of this result also lies in the fact that the permutation $\delta$ is strongly related to the irrationalities of characters of the associated Hecke algebra. Theorem 4.10 was first observed empirically by Beynon and Lusztig, [13], in the case of Weyl groups. Here $\delta$
is non-trivial only for characters such that the corresponding character of the associated Iwahori–Hecke algebra is non-rational. An a priori proof in this case was later given by Opdam, [151]. If $W$ is complex, Theorem 4.10 was verified by Malle, [140, Theorem 6.5], in a case-by-case analysis. Again in all but possibly finitely many cases, $\delta$ comes from the irrationalities of characters of the associated Hecke algebra. Opdam, [152, Theorem 4.2 and Corollary 6.8], gives a general argument which proves theorem 4.10 under a suitable assumption on the braid group $B(W)$.

The above discussion is exclusively concerned with representations over a field of characteristic 0 (the "semisimple case"). We close this chapter with some remarks concerning the modular case.

4.11. Modular representations of $\mathfrak{S}_n$

Frobenius’ theory (as developed further by Specht, James and others) yields a parametrisation of $\text{Irr}(\mathfrak{S}_n)$ and explicit formulas for the degrees and the values of all irreducible characters. As soon as we consider representations over a field of characteristic $p > 0$, the situation changes drastically. James, [110], showed that the irreducible representations of $\mathfrak{S}_n$ still have a natural parametrisation, by so-called $p$-regular partitions. Furthermore, the decomposition matrix relating representations in characteristic 0 and in characteristic $p$ has a lower triangular shape with 1s on the diagonal. This result shows that, in principle, a knowledge of the irreducible representations of $\mathfrak{S}_n$ in characteristic $p$ is equivalent to the knowledge of the decomposition matrix.

There are a number of results known about certain entries of that decomposition matrix, but a general solution to this problem is completely open; see James, [110], for a survey. Via the classical Schur algebras (see Green, [97]) it is known that the decomposition numbers of $\mathfrak{S}_n$ in characteristic $p$ can be obtained from those of the finite general linear group $\text{GL}_n(\mathbb{F}_q)$ (where $q$ is a power of $p$). Now, at first sight, the problem of computing the decomposition numbers for $\text{GL}_n(\mathbb{F}_q)$ seems to be much harder than for $\mathfrak{S}_n$. However, Erdmann, [77], has shown that, if one knows the decomposition numbers of $\mathfrak{S}_m$ (for sufficiently large values of $m$), then one will also know the decomposition numbers of $\text{GL}_m(\mathbb{F}_q)$. Thus, the problem of determining the decomposition numbers for symmetric groups appears to be as difficult as the corresponding problem for general linear groups.

In a completely different direction, Dipper and James, [70], showed that the decomposition numbers of $\mathfrak{S}_n$ can also be obtained from the so-called $q$-Schur algebra, which is defined in terms of the Hecke algebra of $\mathfrak{S}_n$. Thus, the problem of determining the representations of $\mathfrak{S}_n$ in characteristic $p$ is seen to be a special case of the more general problem of studying the representations of Hecke algebras associated with finite Coxeter groups. This is discussed in more detail in the chapter on Hecke algebras. In this context, we just mention here that James’ result on the triangularity of the decomposition matrix of $\mathfrak{S}_n$ is generalised to all finite Weyl groups in Geck, [88].

5. Hints for further reading

Here, we give some hints on topics which were not touched in the previous sections.
5.1. Crystallographic reflection groups

Let $W$ be a discrete subgroup of the group of all affine transformations of a finite-dimensional affine space $E$ over $K = \mathbb{C}$ or $K = \mathbb{R}$, generated by affine reflections. If $W$ is finite, it necessarily fixes a point and $W$ is a finite complex reflection group. If $W$ is infinite and $K = \mathbb{R}$, the irreducible examples are precisely the affine Weyl groups (see Section 2.14). In the complex case $K = \mathbb{C}$ there are two essentially different cases. If $E/W$ is compact, the group $W$ is called crystallographic. The non-crystallographic groups are now just the complexifications of affine Weyl groups. The crystallographic reflection groups have been classified by Popov, [160] (see also Kaneko, Tokunaga and Yoshida, [182,115], for related results). As in the real case they are extensions of a finite complex reflection group $W_0$ by an invariant lattice which is generated by roots for $W_0$.

It turns out that presentations for these groups can be obtained as in the case of affine Weyl groups by adding a further generating reflection corresponding to a highest root in a root system for $W_0$ (see Malle, [139]). As for the finite complex reflection groups in Section 3.1, the braid group $B(W)$ of $W$ is defined as the fundamental group of the hyperplane complement. For many of the irreducible crystallographic complex reflection groups it is known that a presentation for $B(W)$ can be obtained by omitting the order relations from the presentation of $W$ described above (see Dung, [72], for affine groups, Malle, [139], for the complex case).

5.2. Quaternionic reflection groups

Cohen, [51], has obtained the classification of finite reflection groups over the quaternions. This is closely related to finite linear groups over $\mathbb{C}$ generated by bireflections, that is, elements of finite order which fix pointwise a subspace of codimension 2. Indeed, using the identification of the quaternions as a certain ring of $2 \times 2$-matrices over the complex numbers, reflections over the quaternions become complex bireflections. The primitive bireflection groups had been classified previously by Huffman and Wales, [101,189]. One of the examples is a 3-dimensional quaternionic representation of the double cover of the sporadic Hall–Janko group $J_2$, see Wilson, [191].

Presentations for these groups resembling the Coxeter presentations for Weyl groups are given by Cohen, [52].

In recent work on the McKay correspondence, quaternionic reflection groups play an important rôle under the name of symplectic reflection groups in the construction of so-called symplectic reflection algebras, see, for example, Etingof and Ginzburg, [78].

5.3. Reflection groups over finite fields

Many of the general results for complex reflection groups presented in section 1 are no longer true for reflection groups over fields of positive characteristic. Most importantly, the ring of invariants of such a reflection group is not necessarily a polynomial ring. Nevertheless we have the following criterion due to Serre, [25, V.6, Example 8], and Nakajima, [144], generalising Theorem 1.6:
5.4. THEOREM (Serre, Nakajima, [144]). Let $V$ be a finite-dimensional vector space over a field $K$ and $W \leq \text{GL}(V)$ a finite group such that $K[V]^W$ is polynomial. Then the pointwise stabilizer of any subspace $U \subseteq V$ has polynomial ring of invariants (and thus is generated by reflections).

The irreducible reflection groups over finite fields were classified by Wagner, [187,188], and Zalesskiï and Serežkin, [192], the determination of transvection groups was completed by Kantor, [116]. In addition to the modular reductions of complex reflection groups, there arise the infinite families of classical linear, symplectic, unitary and orthogonal groups, as well as some further exceptional examples. For a complete list see, for example, Kemper and Malle, [118, Section 1].

The results of Wagner, [188], and Kantor, [116], are actually somewhat stronger, giving a classification of all indecomposable reflection groups $W$ over finite fields of characteristic $p$ for which the maximal normal $p$-subgroup is contained in the intersection $W' \cap Z(W)$ of the centre with the derived group.

Using this classification, the irreducible reflection groups over finite fields with polynomial ring of invariants could be determined, leading to the following criterion:

5.5. THEOREM (Kemper and Malle, [118]). Let $V$ be a finite-dimensional vector space over $K$, $W \leq \text{GL}(V)$ a finite irreducible linear group. Then $K[V]^W$ is a polynomial ring if and only if $W$ is generated by reflections and the pointwise stabilizer in $W$ of any nontrivial subspace of $V$ has a polynomial ring of invariants.

The list of groups satisfying this criterion can be found in [118, Theorem 7.2]. That paper also contains some information on indecomposable groups.

It is an open question whether at least the field of invariants of a reflection group in positive characteristic is purely transcendental (by Kemper and Malle, [119], the answer is positive in the irreducible case).

For further discussions of modular invariant theory of reflection groups see also Derksen and Kemper, [68, 3.7.4].

5.6. $p$-adic reflection groups

Let $R$ be an integral domain, $L$ an $R$-lattice of finite rank, i.e., a torsion-free finitely generated $R$-module, and $W$ a finite subgroup of $\text{GL}(L)$ generated by reflections. Again one can ask under which conditions the invariants of $W$ on the symmetric algebra $R[L]$ of the dual $L^*$ are a graded polynomial ring. In the case of Weyl groups Demazure shows the following extension of Theorem 1.3:

5.7. THEOREM (Demazure, [63]). Let $W$ be a Weyl group, $L$ the root lattice of $W$, and $R$ a ring in which all torsion primes of $W$ are invertible. Then the invariants of $W$ on $R[L]$ are a graded polynomial algebra, and $R[L]$ is a free graded module over $R[L]^W$.

In the case of general lattices for reflection groups, the following example may be instructive: Let $W = \mathcal{S}_3$ the symmetric group of degree 3. Then the weight lattice $L$ of $\mathcal{S}_3$,
considered as $\mathbb{Z}_3$-lattice, yields a faithful reflection representation of $S_3$ with the following property: $\mathbb{Z}_3[L]^{S_3}$ is not polynomial, while both the reflection representations over the quotient field $\mathbb{Q}_3$ and over the residue field $\mathbb{F}_3$ have polynomial invariants, the first with generators in degrees 2 and 3, the second with generators in degrees 1 and 6.

The list of all irreducible $p$-adic reflection groups, that is, reflection groups over the field of $p$-adic numbers $\mathbb{Q}_p$, was obtained by Clark and Ewing, [49], building on the Shephard–Todd theorem. We reproduce it in Table 6.

Using a case-by-case argument based on the Clark–Ewing classification and his own classification of $p$-adic lattices for reflection groups, Notbohm, [150], was able to determine all finite reflection groups $W$ over the ring of $p$-adic integers $\mathbb{Z}_p$, $p > 2$, with polynomial ring of invariants. This was subsequently extended by Andersen, Grodal, Møller and Viruel, [3], to include the case $p = 2$.

### 5.8. $p$-compact groups

The $p$-adic reflection groups play an important rôle in the theory of so-called $p$-compact groups, which constitute a homotopy theoretic analogue of compact Lie groups. By definition, a $p$-compact group is a $p$-complete topological space $BX$ such that the homology $H_*(X; \mathbb{F}_p)$ of the loop space $X = \Omega BX$ is finite. Examples for $p$-compact groups are $p$-completions of classifying spaces of compact Lie groups. Further examples were
Reflection groups

constructed by Clark and Ewing, [49], Aguadé, [2], Dwyer and Wilkerson, [73], and Notbohm, [149]. To each $p$-compact group $X$ Dwyer and Wilkerson, [74], associate a maximal torus (unique up to conjugacy) together with a `Weyl group', which comes equipped with a representation as a reflection group over the $p$-adic integers $\mathbb{Z}_p$, which is faithful if $X$ is connected. Conversely, by a theorem of Andersen et al., [3], a connected $p$-compact group, for $p > 2$, is determined up to isomorphism by its Weyl group data, that is, by its Weyl group in a reflection representation on a $\mathbb{Z}_p$-lattice.

It has been shown that at least for $p > 2$ all $p$-adic reflection groups (as classified by Clark and Ewing) and all their $\mathbb{Z}_p$-reflection representations arise in that way (see Andersen et al., [3], Notbohm, [149], and also Adams and Wilkerson, [1], Aguadé, [2]).

References


N. Bourbaki, Groupes et algèbres de Lie, chap. 4, 5 et 6, Hermann (1968).

K. Bremke and G. Malle, Reduced words and a length function for \( G(e, 1, n) \), Indag. Math. 8 (1997), 453–469.


K. Bremke and G. Malle, Reduced words and a length function for \( G(e, 1, n) \), Indag. Math. 8 (1997), 453–469.


H.S.M. Coxeter, Discrete groups generated by reflections, Ann. of Math. 35 (1934), 588–621.

H.S.M. Coxeter, The complete enumeration of finite groups of the form \( R_1^2 = (R_1 R_2)^{2^i} = 1 \), J. London Math. Soc. 10 (1955), 21–28.

Reflection groups


[99] W. Huffman and D. Wales, Linear groups containing an involution with two eigenvalues $-1$, J. Algebra 45 (1977), 465–515.
Reflection groups


[176] R. Steinberg, Lectures on Chevalley groups, Mimeographed notes, Department of Mathematics, Yale University (1967).


This page intentionally left blank
Hurwitz Groups and Hurwitz Generation

M.C. Tamburini

Dipartimento di Matematica e Fisica, Università Cattolica del Sacro Cuore, Brescia, Italy
E-mail: c.tamburini@dmf.unicatt.it

M. Vsemirnov

St. Petersburg Division of Steklov Institute of Mathematics, St. Petersburg, Russia
E-mail: vsemir@pdmi.ras.ru

Contents
1. Introduction ........................................ 387
2. Triangle groups ........................................ 388
3. Finite simple and quasi-simple groups which are Hurwitz  ........................................ 390
   3.1. Alternating groups ........................................ 391
   3.2. Classical groups ........................................ 394
   3.3. Exceptional groups of Lie type and sporadic simple groups ........................................ 396
4. Low-dimensional representations of Hurwitz groups ........................................ 397
   4.1. Scott’s formula and negative results ........................................ 398
   4.2. Rigidity .................................................. 401
   4.3. Classical groups of small rank which are Hurwitz ........................................ 403
5. Related results ........................................ 410
   5.1. Number-theoretic aspects ........................................ 410
   5.2. Other groups which are (2, 3, 7)-generated ........................................ 412
Acknowledgements ........................................ 415
Appendix .................................................. 415
References .................................................. 424
1. Introduction

A \((2, 3, 7)\)-generated group is a group generated by two elements of order 2 and 3 respectively such that their product has order 7. Such a group is called Hurwitz if it is finite. In other words, Hurwitz groups are the non-trivial finite homomorphic images of the abstract triangle group \(T(2, 3, 7)\) defined by the presentation

\[ T(2, 3, 7) = \langle X, Y \mid X^2 = Y^3 = (XY)^7 = 1 \rangle. \]

In particular, the Hurwitz groups form a wide and remarkable class of the so-called \((2, 3)\)-generated groups, i.e. the non-trivial epimorphic images of the free product

\[ C_2 \ast C_3 = \langle X, Y \mid X^2 = Y^3 = 1 \rangle. \]

It is well known that \(C_2 \ast C_3\) is isomorphic to \(\text{PSL}_2(\mathbb{Z})\) (R. Fricke and F. Klein, [32]). The structure of normal subgroups of \(\text{PSL}(2, \mathbb{Z})\) and the corresponding factor groups were the subject of intensive study; for instance, see [26,27,77,81,80] and, especially, the remarkable paper of M.W. Liebeck and A. Shalev, [50].

The study of Hurwitz groups goes back to the late XIX century and shows an important connection with the theory of Riemann surfaces. In 1893, A. Hurwitz, [37], proved that the automorphism group of an algebraic curve of genus \(g \geq 2\) always has order at most \(84(g - 1)\) and that this upper bound is attained precisely when the group is an image of \(T(2, 3, 7)\). Hurwitz’s discovery originated from the example (due to F. Klein, [45]), of \(\text{PSL}_2(7)\), the smallest Hurwitz group, acting as the automorphism group of the quartic \(x^3y + y^3z + z^3x = 0\) of genus 3.

Since then, examples of Hurwitz groups were rather fragmentary until the pioneering paper of A.M. Macbeath, [57], appeared in 1969. In this paper he describes all prime powers \(q\) such that the group \(\text{PSL}_2(q)\) is Hurwitz. On the other hand, a result of J. Cohen, [6], asserts that the Hurwitz subgroups of \(\text{PSL}_3(q)\) are just those which arise from representations of the above groups discovered by A.M. Macbeath. And this fact may have erroneously discouraged, for a long time, the search for (projective) linear groups which are Hurwitz.

The next significant step in the positive direction was done by G. Higman and M.D.E. Conder who developed a very powerful method of building new permutational representations of \(T(2, 3, 7)\) via combinatorial diagrams. As a result, Conder, [8], proved that almost all alternating groups are Hurwitz. Later in the papers of A. Lucchini, M.C. Tamburini and J.S. Wilson, [55,56], these constructive ideas were generalized to a linear context, providing a new bunch of Hurwitz groups, which include most finite classical simple groups of sufficiently large rank. Actually several authors considered the problem of determining which finite simple groups are Hurwitz. Among them, G. Malle, [60,61], gave precise answers for many classes of exceptional simple groups of Lie type. And, by the contributions of A. Woldar, [95], R.A. Wilson, [92–94], and others, it is now known exactly which of the 26 sporadic simple groups are Hurwitz.

It can be shown that there are \(2^{\aleph_0}\) non-isomorphic \((2, 3, 7)\)-generated groups, [56]. So any attempt to classify all of them is not realistic. But, as mentioned above, there have
been significant achievements in studying specific classes of groups (e.g., finite simple groups) with respect to the property of being Hurwitz. And there have been achievements in classifying the low-dimensional linear and projective representations of $T(2, 3, 7)$ over an algebraically closed field $\mathbb{F}$ of characteristic $p \geq 0$, [82,78].

In this connection there is a crucial formula, due to L.L. Scott, [68]. Given a group $H = \langle a_1, \ldots, a_m \rangle$ and a representation $f : H \to \text{GL}_n(\mathbb{F})$, this formula restricts the similarity invariants of $f(a_1), \ldots, f(a_m)$ and of the product $f(a_1 \cdots a_m)$. Using Scott’s result, L. Di Martino, M.C. Tamburini and A.E. Zalesskii, [25], excluded most of the linear classical groups in dimensions up to 19 from being Hurwitz. For small $n$, further combination of Scott’s formula with results of K. Strambach and H. Völklein, [74], on linearly rigid triples allows to classify the irreducible Hurwitz subgroups of $\text{SL}_n(\mathbb{F})$ and $\text{PSL}_n(\mathbb{F})$, for $n \leq 5$. We refer to Section 4.3, which is devoted to this classification, and to [82].

There are other aspects of Hurwitz groups which are interesting in themselves, and also shed more light on the understanding of the groups. For example, using number theory, M. Vsemirnov, V. Mysovskikh and M.C. Tamburini, [88], gave an alternative definition of $T(2, 3, 7)$ as a unitary group over an appropriate ring. This result also has a strict relation to Macbeath’s theorem.

The aim of this chapter is to survey the main achievements and ideas in this field as well as bring together recent results widely dispersed in the literature. Some of the results or proofs appear here for the first time. However, we do not touch some specific matters already covered in previous survey articles. For further reading we recommend the survey articles [13,22,23,40], and [91].

2. Triangle groups

DEFINITION 2.1. Let $G$ be a group and $k, l, m$ be integers $\geq 2$. If $x, y \in G$ have orders $k$ and $l$, respectively, and $z = xy$ has order $m$, we say that the triple $(x, y, z)$ is a $(k, l, m)$-triple. A group $G$ is called $(k, l, m)$-generated if it can be generated by two elements $x$ and $y$ such that $(x, y, xy)$ is a $(k, l, m)$-triple. In this case we also say that $(x, y, xy)$ is a $(k, l, m)$-generating triple.

In particular, any $(k, l, m)$-generated group is a homomorphic image of the abstract triangle group $T(k, l, m)$ defined by the presentation

$$T(k, l, m) = \langle X, Y \mid X^k = Y^l = (XY)^m = 1 \rangle.$$ 

The groups $T(k, l, m)$ have a nice geometric description.¹ Let $\Delta = \Delta(k, l, m)$ be a triangle having angles of size $\frac{\pi}{k}$, $\frac{\pi}{l}$, $\frac{\pi}{m}$, that is $\Delta$ is a spherical, Euclidean or hyperbolic triangle depending on whether

$$\frac{1}{k} + \frac{1}{l} + \frac{1}{m}$$

¹This geometric interpretation also explains the terminology ‘triangle group’.
Hurwitz groups and Hurwitz generation

is greater than, equal to or less than 1. Then $T(k, l, m)$ can be defined as a group of motions of the two-dimensional space (sphere, Euclidean plane or hyperbolic plane, respectively), namely, as the group generated by rotations of angles $\frac{2\pi}{k}$, $\frac{2\pi}{l}$, $\frac{2\pi}{m}$ around the corresponding vertices of $\triangle$. We just mention that hyperbolic triangle groups are special cases of Fuchsian groups, i.e., finitely generated discontinuous groups of orientation-preserving non-Euclidean motions (for instance, see [59]).

Let $T^*(k, l, m)$ be the group of motions, generated by reflections around the sides of $\triangle$. It can be shown that the images of $\triangle$ under $T^*(k, l, m)$ tessellate the corresponding space without overlapping. In addition, $T^*(k, l, m)$ admits the presentation

$$T^*(k, l, m) = \langle X, Y, T \mid X^k = Y^l = (XY)^m = T^2 = (XT)^2 = (YT)^2 = 1 \rangle$$

and $T(k, l, m)$ is a subgroup of index 2 in $T^*(k, l, m)$. In particular, $T(k, l, m)$ is finite precisely when

$$\frac{1}{k} + \frac{1}{l} + \frac{1}{m} > 1,$$

see, e.g., [18, Section 6.4], where these groups appear under the name polyhedral groups.

The above geometric description of triangle groups $T(k, l, m)$ allows to embed them into $\text{PSL}(2, \mathbb{C})$. We indicate an explicit embedding only when

$$\frac{1}{k} + \frac{1}{l} + \frac{1}{m} < 1.$$

The following construction is taken from [59, Chapter II, Exercises 5, 6]. Set

$$\kappa = e^{-i\pi}, \quad \lambda = e^{i\pi}, \quad \mu = e^{i\alpha}, \quad r = \rho^{-1} - \rho,$$

where $\rho$ is the positive root of

$$t^2(\mu + \mu^{-1} + \lambda\kappa^{-1} + \kappa\lambda^{-1}) = \mu + \mu^{-1} + \lambda\kappa + (\lambda\kappa)^{-1}.$$

Let $X$, $Y$ be the Möbius transformations with matrices

$$\begin{pmatrix} \kappa & 0 \\ 0 & \kappa^{-1} \end{pmatrix} \quad \text{and} \quad r^{-1} \begin{pmatrix} \rho\lambda^{-1} - \lambda\rho^{-1} & \lambda - \lambda^{-1} \\ \lambda^{-1} - \lambda & \lambda\rho - (\lambda\rho)^{-1} \end{pmatrix},$$

respectively. Then $X$, $Y$ map the interior of the unit disc $|z| < 1$ into itself and $k, l, m$ are respectively the exact orders of $X$, $Y$, $XY$. In addition, the fixed points of $X$, $Y$, $XY$ within the unit disc are the vertices of a non-Euclidean triangle $\triangle$ with angles $\frac{\pi}{k}$, $\frac{\pi}{l}$, $\frac{\pi}{m}$ and $X$, $Y$ actually generate the triangle group $T(k, l, m)$.

The above interpretation relates triangle groups and in particular $T(2, 3, 7)$ to hyperbolic geometry and the theory of Riemann surfaces. The following theorem explains the importance of $T(2, 3, 7)$ and Hurwitz groups in this context.
THEOREM 2.2 (Hurwitz, [37]). Let $S$ be a compact Riemann surface of genus $g \geq 2$ and $H$ be its automorphism group. Then $|H| \leq 84(g - 1)$. Moreover, a finite group $H$ of order $84(g - 1)$ is the automorphism group of a compact Riemann surface of genus $g$ if and only if $H$ is $(2, 3, 7)$-generated.

The proof of this result is based on the fact that, among all Fuchsian groups, $T(2, 3, 7)$ has the fundamental domain of the smallest volume. Details can be found in many standard textbooks like [59, Section II.7] or [42, Section 5.11]. For example, the smallest Hurwitz group $\text{PSL}_2(7)$ of order 168 is the group of automorphisms of Klein’s quartic $x^3y + y^3z + z^3x = 0$ of genus 3. As we will see later, there are infinitely many non-isomorphic Hurwitz groups. In other words there are infinitely many other values of the genus $g$ for which the Hurwitz upper bound is attained. However, it is not attained when $g = 2$. Moreover, it can be shown that there are infinitely many values of $g$ for which it is not attained. The precise values of $g$ for which the Hurwitz bound is attained are still unknown, see, e.g., [42, Section 5.11] where this problem was posed. M.D.E. Conder, [11,12], showed that in the range $1 < g < 11905$ there are just 32 integers $g$ such that there exists a compact Riemann surface of genus $g$ with the automorphism group of the maximal possible order $84(g - 1)$. Moreover, Conder also determined all the 92 normal subgroups of $T(2, 3, 7)$ of index less than $10^6$. In particular, there are exactly 14 simple Hurwitz groups of order less than one million, [12, Table 1].

Finally, the above geometric interpretation also gives some information about indices of subgroups of $T(2, 3, 7)$. Let $G$ be a subgroup of $T(2, 3, 7)$ of index $n$. It has a fundamental domain consisting of $n$ translates of the hyperbolic triangle $\Delta(2, 3, 7)$. The domain has, say, $r$ (respectively, $s$, $t$) elliptic vertices of order 2 (respectively, 3, 7) and the corresponding Riemann surface has genus $g$. The numbers $n$, $g$, $r$, $s$, $t$ are not independent: they are related via the genus formula

$$n = 84(g - 1) + 21r + 28s + 36t.$$  \hfill (1)

As an easy consequence, we have that $n$ must satisfy

$$\left\lceil \frac{n}{2} \right\rceil + 2\left\lceil \frac{n}{3} \right\rceil + 6\left\lceil \frac{n}{7} \right\rceil \geq 2n - 2.$$  \hfill (2)

W.W. Stothers, [73], showed that, with the exception of $(16,0,0,1,2)$, $(21,1,1,0,0)$ and $(31,1,0,0,1)$, any quintuple $(n,g,r,s,t)$ satisfying (1) corresponds to a subgroup of $T(2,3,7)$.

3. Finite simple and quasi-simple groups which are Hurwitz

Throughout this chapter $\mathbb{F}$ always denotes an algebraically closed field of characteristic $p \geq 0$. The first significant and well-known result is the following:

THEOREM 3.1 (Macbeath, [57]). Let $p > 0$. The group $\text{PSL}_2(\mathbb{F})$ contains exactly one conjugacy class of Hurwitz subgroups. Namely,
(1) $\text{PSL}_2(p)$, if $p \equiv 0, \pm 1 \pmod{7}$;
(2) $\text{PSL}_2(p^3)$, if $p \equiv \pm 2, \pm 3 \pmod{7}$.

We will give a short proof of a more general statement in Theorem 4.9. In fact Macbeath himself showed more, because he classified all the subgroups of $\text{PSL}_2(\mathbb{F})$ which are finite epimorphic images of triangle groups. His original proof is based on a detailed analysis of conjugacy classes and knowledge of the subgroups of $\text{PSL}_2(q)$.

This subject is investigated further in [47] and [49], where criteria are established to determine for which finite fields $\text{GF}(q)$ a given triangle group has $\text{PSL}_2(q)$ or $\text{PGL}_2(q)$ as factor group with torsion free kernel. Finally, permutational representations of the triangle group $T(2, 3, 7)$, which arise from the action of $\text{PSL}_2(q)$ on the corresponding finite projective line, are studied in [63].

### 3.1. Alternating groups

As noted at the end of the previous section, the above result provides many transitive permutational representations of $T(2, 3, 7)$. For example those arising from the action of $\text{PSL}_2(q)$, when Hurwitz, on the $q + 1$ points of the projective line. Permutational representations of $T(2, 3, 7)$ of small degrees, and a method of joining them via handles developed by G. Higman, were the starting point for constructive methods, and culminated in the famous theorem that almost all the alternating groups are Hurwitz, which appeared in the paper of M.D.E. Conder, [8]. The same methods, later generalized to a linear context by A. Lucchini, M.C. Tamburini and J.S. Wilson, led to the discovery that most finite classical groups are Hurwitz. In this section we attempt to describe the above results in a uniform way, which is close to the approach used in [55].

Let $V$ be a free module over a ring $R$, with basis $\Omega$ of cardinality $n$, and let $\text{GL}_n(R)$ act on $V$. It is natural to identify $\text{Sym}(\Omega)$ with the subgroup of $\text{GL}_n(R)$ consisting of permutation matrices, and $\text{Alt}(\Omega)$ with the subgroup of $\text{SL}_n(R)$ of even permutation matrices. Assume that $(X, Y, Z)$ is a $(2, 3, 7)$-generating triple of the triangle group $T(2, 3, 7)$, and that

$$\psi : T(2, 3, 7) \to \text{GL}_n(R)$$

is a representation.

**Definition 3.2.** An ordered pair $(v_1, v_2)$ of distinct elements from $\Omega$ is called a **handle** for $\psi$ if the following conditions hold:

1. $\psi(X)$ fixes $v_1$ and $v_2$ and leaves invariant the submodule $\langle \Omega \setminus \{v_1, v_2\} \rangle$;
2. $\psi(Y)$ acts as the permutation $(v_1, v_2, v_3)$ for some $v_3 \in \Omega$, and leaves invariant $\langle \Omega \setminus \{v_1, v_2, v_3\} \rangle$.

The role of handles, in the process of joining representations of the triangle group, is made clear by the following lemma. Here we assume that $\{e_1, \ldots, e_n\}$ is the canonical basis of the free $R$-module $R^n$ and, similarly, that $\{e'_1, \ldots, e'_n\}$ is the canonical basis of $R'^n$. 

LEMMA 3.3. Given two representations
\[ \psi : T(2, 3, 7) \to \text{GL}_n(R), \quad \psi' : T(2, 3, 7) \to \text{GL}_{n'}(R) \]
assume that \( \psi \) has handles \( \{ e_1, e_2 \} \) and \( \psi' \) has handles \( \{ e'_1, e'_2 \} \).
Let \( X_i \) to be one of the following involutions of \( \text{GL}_n + n'(R) \):
\[
X_1 := \begin{pmatrix}
I_{n-2} & I_2 \\
I_2 & I_{n'-2}
\end{pmatrix}, \quad X_2 := \begin{pmatrix}
I_2 & I_{n-2} \\
-I_2 & I_{n'-2}
\end{pmatrix},
\]
where \( t \in R \). Then, for \( i = 1, 2 \), the map
\[
X \mapsto X_1 \begin{pmatrix}
\psi_1(X) \\
\psi_2(X)
\end{pmatrix}, \quad Y \mapsto \begin{pmatrix}
\psi_1(Y) \\
\psi_2(Y)
\end{pmatrix}
\]
defines a representation \( T(2, 3, 7) \to \text{GL}_n + n'(R) \).

This elementary lemma rests on the following two facts.
(1) For \( i = 1, 2 \), the involution \( X_i \) and the involution
\[
\begin{pmatrix}
\psi_1(X) \\
\psi_2(X)
\end{pmatrix}
\]
commute, having disjoint supports. Hence their product is again an involution.
(2) The product
\[
X_i \begin{pmatrix}
\psi_1(XY) \\
\psi_2(XY)
\end{pmatrix}
\]
has order 7, being conjugate to
\[
\begin{pmatrix}
\psi_1(XY) \\
\psi_2(XY)
\end{pmatrix}.
\]
In particular, if \( \psi \) and \( \psi' \) are transitive permutational representations of \( T(2, 3, 7) \), of respective degrees \( n \) and \( n' \), the representation described in Lemma 3.3, relative to \( X_1 \), is a transitive permutational representation of degree \( n + n' \).

In [8], to define a \( (2, 3, 7) \)-generating triple \( (x, y, z) \) of \( \text{Alt}(\Omega) \) when \( n > 167 \) and \( n \neq 173, 174, 181, 188, 202 \), Conder uses \( 3 + 14 \) transitive permutational representations of \( T(2, 3, 7) \), each of which is depicted by a diagram, whose vertices are permuted by \( T(2, 3, 7) \). The first three diagrams, denoted \( A, E \) and \( G \), have respectively 14, 28 and 42 vertices. (\( A \), for example, corresponds to the action of \( \text{PSL}_2(13) \) on the 14 points of the
The remaining fourteen diagrams can be labeled $H_d$, $d = 0, \ldots, 13$. Each $H_d$ has $d'$ vertices, where $d'$ is the unique integer determined by the conditions

$$d' \in D := \{36, 42, 57, 77, 115, 135, 136, 142, 144, 165, 180, 187, 195, 216\}$$

and $d' \equiv d \pmod{14}$. To avoid too many details we do not include these diagrams here. However in the appendix we present an explicit description for Conder’s generators, which allows to restore all these diagrams. As the numbers in $D$ give all residues modulo 14, each $n$ big enough can be written in the form

$$n = 42a + 14b + d', \quad a \geq 2, \ b \in \{0, 1, 2\}, \ d' \in D.$$

So, if $\Omega$ is a set of cardinality $n$, one can take

$$\Omega := G_1 \cup \cdots \cup G_a \cup H_d \cup \Omega_0,$$

where $d \equiv n \pmod{14}$, each $G_i$ is a copy of $G$ and $\Omega_0$ is empty if $b = 0$, whereas $\Omega_0$ coincides with $A$ if $b = 1$ or with $E$ if $b = 2$. As the diagram $G$ has 3 handles and each of the diagrams $A, E$ and $H_d$ has at least one handle, repeated application of Lemma 3.3 with respect to $X_1$, gives a transitive permutational representation of $T(2, 3, 7)$ over $\Omega$. In fact it is possible to join the $a$ copies of $G$ into a chain, join $G_a$ with $H_d$ and, if necessary, with $A$ or $E$. There is a certain degree of flexibility in making the joins, depending on the choice of the handles. But, no matter how the joins are performed, any representation $\psi : T(2, 3, 7) \to \text{Sym}(\Omega)$ obtained in this way has the following properties. Set

$$\psi(X) = x, \quad \psi(Y) = y.$$

Then we have (see appendix):

(i) $\{x, y\}^{9 \cdot 11 \cdot 13}$ fixes each vector in $\Omega \setminus (H_d \cup G_a)$;

(ii) there exists a multiple $k = k(d)$ of $9 \cdot 11 \cdot 13$ such that

$$c := [x, y]^k$$

is a cycle of prime length $r \notin \{2, 3, 11, 13\}$, with support $\Gamma \subseteq H_d$;

(iii) $\Gamma$ contains an orbit of $x$ and two points from an orbit of $y$;

(iv) $|\Gamma \cup \Gamma y| \geq r + 3$.

In the appendix the cycle $c$ is written explicitly in each case.

In particular, $\langle c, c^3 \rangle = \text{Alt}(\Gamma \cup \Gamma y)$ and the normal closure of this group under the transitive subgroup $\langle x, y \rangle$ of $\text{Alt}(\Omega)$ is easily seen to be $\text{Alt}(\Omega)$.

Similar considerations, with more specific arguments for some values of $n$, lead to the following:

**Theorem 3.4 (Conder, [8]).** The alternating group $A_n$ is Hurwitz for all $n > 167$ and for the values of $n$ displayed in the following table.
Actually, in the paper [8], the previous theorem is obtained as a corollary of the following remarkable result. Whenever \( n > 167 \), the symmetric group \( S_n \) is an epimorphic image of \( T^*(2, 3, 7) \). To prove this it is essential to have a third generator \( T \), which corresponds to a symmetry in the vertical axis of each of the 17 diagrams mentioned above.

There are variations of these results in several directions. For example, in [14], Conder shows that all but finitely many of the alternating groups \( A_n \) can be generated by a \((2, 3, 7)\)-generating triple \((x, y, xy)\) satisfying the further relation \([x, y]^{84} = 1\).

Many authors considered other triangle groups. Conder, [9], obtained the following result.

**Theorem 3.5 (Conder, [9]).** For each \( k \geq 7 \), there exists an \( n_k \) such that, for all \( n \geq n_k \), \( A_n \) is an epimorphic image of \( T(2, 3, k) \).

Actually he proves more than he claims, because a careful analysis of his diagrams leads to the stronger conclusion that \( A_n \) is \((2, 3, k)\)-generated.

Using a similar technique Q. Mushtaq and G.-C. Rota, [64], proved

**Theorem 3.6 (Mushtaq and Rota, [64]).** Let \( k \) be even, \( k \geq 6 \) and \( l \geq 5k - 3 \). For sufficiently large \( n \), the group \( A_n \) is a homomorphic image of \( T(2, k, l) \).

An even more striking generalization of the above results was given by B. Everitt.

**Theorem 3.7 (Everitt, [29]).** Any Fuchsian group surjects almost all of the alternating groups.

Recently, M.W. Liebeck and A. Shalev, [51], gave another proof of this result. Their proof uses character-theoretic and probabilistic methods and it is totally independent from Higman’s and Conder’s diagrams.

### 3.2. Classical groups

Already in [77] constructive, permutational methods had been used by the first author of this survey to show that, for all \( n \geq 25 \) and all prime powers \( q \), the special linear group
SL\(_n(q)\) can be generated by an element of order 2 and an element of order 3, i.e., is an epimorphic image of the modular group PSL\(_2(\mathbb{Z})\). This result actually originated from the following:

**Theorem 3.8** (Tamburini and Wilson, [79]). Let \(A\) and \(B\) be finite groups which are non-trivial. If \(|A||B| \geq 12\), then for all \(n \geq |A||B| + 12\) the group \(\text{PSL}_n(q)\) has subgroups \(\tilde{A} \simeq A\) and \(\tilde{B} \simeq B\) such that \(\langle \tilde{A}, \tilde{B} \rangle = \text{PSL}_n(q)\).

A key tool in the proof of this theorem is a simple and beautiful idea of H. Wielandt, [89], which requires the hypothesis that at least one of the groups \(A\) and \(B\) has order \(\geq 4\). But an appropriate variation of the proof of this theorem was used in [77] to establish the similar result for the smallest possible values of \(|A|\) and \(|B|\), namely 2 and 3. And, indeed, the \((2, 3)\)-generation of the projective special linear groups PSL\(_n(q)\), provided \(n > 4\) when \(q\) is odd and \(n > 12\) when \(q\) is even, has been established by L. Di Martino and N. Vavilov in [26] and [27], in a constructive way which involves quite a lot of case by case analysis and heavy computation.

A combination of the linear methods in [77] with the \((2, 3, 7)\) generators for Alt(\(\Omega\)) of Theorem 3.4, gives the following:

**Theorem 3.9** (Lucchini, Tamburini and Wilson, [56]). For all \(n \geq 287\), the special linear group \(\text{SL}_n(q)\) is Hurwitz.

The authors take \(n = |\Omega|\) big enough in order to guarantee that the representation of \(T(2, 3, 7)\) affording the generators \(x, y\) of Alt(\(\Omega\)) in (4) has a couple of handles \((e_1, e_2)\) and \((e'_1, e'_2)\). Thus they can apply Lemma 3.3 to extend the generating triple \((x, y, x\cdot y)\) of Alt(\(\Omega\)) to a \((2, 3, 7)\)-generating triple \((xX_2, y, xX_2\cdot y)\) of SL\(_n(q)\). In the definition of \(X_2\) they take \(t \neq 2\) to be a generator of GF\((q)\) (as a ring). Their proof consists in showing that

\[ \text{Alt(\(\Omega\))} \leq \langle xX_2, y \rangle \]

and their claim follows from the fact that for \(n \geq 6\), SL\(_n(q)\) is generated by \(X_2\) and Alt(\(\Omega\)).

An inspection of the proof shows that SL\(_n(q)\) (and SL\(_n(\mathbb{Z})\)) are \((2, 3, 7)\)-generated for all \(n\) in the set

\[ \{14m + d \mid m \geq 6, \; d \in D\} \cup \{42 + d \mid d \in D\}, \]

where \(D\) is as in (3). There are 93 integers less than 286 in this set. Further improvement was made in [86], where 60 new values of \(n\) were found. In particular it follows that SL\(_n(q)\) is Hurwitz for all \(n \geq 252\) and for 118 more values of \(n\), the smallest of which is 49.

Duplication of the \((2, 3)\)-generators of SL\(_n(q)\), according to a well-known embedding of this group into the classical groups of degree \(2n\) or \(2n + 1\) over GF\((q)\), had already been used in [81] and [80] to show that classical groups of sufficiently large rank are \((2, 3)\)-generated. In a similar way, duplication of the \((2, 3, 7)\)-generators of SL\(_n(q)\), and an application of Lemma 3.3 with appropriate choices of \(X_1\), leads to the following results.
THEOREM 3.10 (Lucchini and Tamburini, [55]). For each \( n \geq 371 \), the following classical groups are Hurwitz:

- \( \text{Sp}_{2n}(q) \), \( \text{SU}_{2n}(q) \), \( \Omega_{2n}^+(q) \), all \( q \);
- \( \text{SU}_{2n+7}(q) \), \( \Omega_{2n+7}(q) \), \( q \) odd.

As above, analysis of the proof shows that this result holds for all \( n = 42a + 14b + d \) with \( d \in D \) as in (3), and either \( a \geq 4 \) or \( a = 3 \) and \( b = 0 \). There are many such integers less than 371, the smallest of which is 162.

As in the permutational case, further generalizations to other triangle groups are possible. For example, A. Lucchini in [54] and J.S. Wilson in [91] independently proved that for any \( k \geq 7 \) there is an integer \( n_k \) such that the group \( \text{SL}_n(q) \) is \( (2, 3, k) \)-generated provided \( n \geq n_k \). In fact, their result is a consequence of a more general statement; see Theorem 5.9.

It may be worth noting that the problem of determining which finite classical groups are Hurwitz deserves further investigation, having received only partial answers. In fact there are two classes which are not even considered in the above theorem, namely the orthogonal groups \( \Omega_{2n}^-(q) \) and the unitary groups \( \text{SU}_{2n+1}(2^t) \). Moreover, although the above results are satisfactory if considered asymptotically with respect to the ranks of the groups under consideration, the lower bounds for their ranks are certainly much higher than necessary for the existence of Hurwitz generators. Some evidence for this claim will be given in Section 4.3 dedicated to groups of small rank. In fact it will be shown that the reason why the Hurwitz subgroups of \( \text{PSL}_n(F) \), with \( n \leq 4 \), are essentially those discovered by Macbeath is rigidity. On the other hand, already for \( n = 5 \) and \( n = 7 \) there are new \( (2, 3, 7) \)-generated projective subgroups. In the theorems mentioned above, the assumptions on the lower bounds are forced by the permutational approach, based on the diagrams of Conder, which has the advantage of being constructive and allowing rather uniform proofs.

But the treatment of linear groups of relatively small ranks requires different techniques, which may in any case involve quite a lot of computation and consideration of special cases. In fact the property of being Hurwitz for groups of small rank depends also on the size of the field and its characteristic. Evidence of this is given by the above mentioned result of Macbeath and also by the results in [60, 61] and [82, 78] which will be mentioned in the following sections.

### 3.3. Exceptional groups of Lie type and sporadic simple groups

A key tool in the approach to these groups has been the use of multiplication constants defined as follows. Let \( X_1, \ldots, X_r \) denote the conjugacy classes of a finite group \( G \). For a fixed \( z \in X_k \), the number of pairs \((x, y)\) such that \( x \in X_i \), \( y \in X_j \) and \( xy = z \) coincides with the number

\[
a_{ijk} := \frac{|X_i| |X_j|}{|G|} \sum_{\chi \in \text{Irr}(G)} \frac{\chi(g_i) \chi(g_j) \overline{\chi(g_k)}}{\chi(1)},
\]

where \( g_\ell \in X_\ell \), \( 1 \leq \ell \leq r \). (See [33, Theorem 2.12], for example.) These numbers, also called the multiplication constants, are useful in determining whether \( G \) is Hurwitz. One
first computes the class constants for each choice of classes of elements of order 2, 3 and 7 in $G$. Clearly, if they are all 0, then $G$ has no Hurwitz subgroup. But, apart from this trivial case, one can use rather sophisticated techniques, based on additional information about $G$. Like calculating the class constants in appropriate subgroups of $G$, in order to evaluate how many solutions generate proper subgroups. A refined version of this technique is due to Philip Hall, [35] (see also [41]).

Using the Green–Deligne–Lusztig parameterizations of characters of groups of Lie type, G. Malle studied the Hurwitz generation of many exceptional simple groups of Lie type. For a more detailed description we refer also to the survey [22] of L. Di Martino, and to the survey [41] of G. Jones.

**Theorem 3.11** (Malle, [60,61]). As to the exceptional simple groups of Lie type:

1. $G_2(p^m)$ are Hurwitz if and only if $p^m \geq 5$;
2. $2G_2(3^{2m+1})$ are Hurwitz if and only if $m \neq 1$;
3. $3D_4(p^m)$ are Hurwitz if and only if $p \neq 3$, $p^m \neq 4$;
4. $2F_4(2^{2m+1})'$ are Hurwitz if and only if $m \equiv 1 \pmod{3}$.

For the Ree groups $2G_2(3^{2m+1})$ see also, [40] and [67]. The results of Theorem 3.11 do not produce explicit $(2,3,7)$-generators. With different methods two explicit matrices $x, y \in \text{SL}_7(p)$, $p \geq 5$, are constructed in [87] such that $x^2 = y^3 = (xy)^7 = [x, y]^{2p} = I$ and $(x, y)$ is isomorphic to $G_2(p)$.

By the contribution of several authors, the problem of determining which of the 26 sporadic simple groups are Hurwitz, has now a complete answer.

**Theorem 3.12.** The sporadic simple groups which are Hurwitz are the following:

1. $J_1$ (Sah, [67]);
2. $J_2$ (Finkelstein and Rudvalis, [31]);
3. $Co_3$ (Worboys, [97] and Woldar, [95]);
4. $He$, $Ru$, $HN$, $Ly$, $Fi_{22}$, $J_4$ (Woldar, [95,96]);
5. $Th$ (Linton, [52]);
6. $Fi_{24}$ (Linton and Wilson, [53]);
7. $M$ (Wilson, [94]).

The orders of $M_{11}$, $M_{12}$ and $J_3$ are not divisible by 7. The proof that the remaining simple groups are not Hurwitz comes from the determination of their symmetric genus, [16,44]. The technique is a combination of the above method of multiplication constants and Scott’s formula. The latter will be discussed in Section 4.

### 4. Low-dimensional representations of Hurwitz groups

Clearly, a $(2,3,7)$-generated group $G$ must have order divisible by 2, 3 and 7 and it must be perfect, i.e. with trivial abelianization $G/G'$. Moreover, for any subgroup $S$ of $G$, its index $n$, when finite, must satisfy the classical genus formula (2). But there are other methods which can be used to exclude that a group $G$ is $(2,3,7)$-generated. They will be illustrated in this section.
4.1. Scott’s formula and negative results

Theorem 4.1 below, which is a special case of a result of L.L. Scott, provides a very efficient tool to show that certain groups are not \((2, 3, 7)\)-generated. We state this result only in the form needed for our purposes. However the original theorem of Scott applies to a more general context and deals with representations of any finitely generated group \(H\). To state the theorem, we need some notation. Given a group \(H\) and a representation \(f : H \to \text{GL}_n(F)\) let \(V\) be the vector space \(F^n\). For any subset \(A\) of \(H\), define \(VA\) as the subspace of fixed points of \(f(A)\) and denote by \(d_A^V\) its dimension over \(F\). In symbols:

\[
VA := \{v \in V \mid f(a)v = v, \text{ for all } a \in A\}, \quad d_A^V := \dim(V_A).
\]  

(5)

Define \(\hat{d}_A^V\) in the same way, with respect to the dual representation, namely set

\[
\hat{V}_A := \{v \in V \mid (f(a))^t v = v, \text{ for all } a \in A\}, \quad \hat{d}_A^V := \dim(\hat{V}_A).
\]  

(6)

In the above notations:

**THEOREM 4.1 (Scott, [68]).** Assume that \(H\) is generated by \(x\) and \(y\). Then

\[
d_x^V + d_y^V + d_{xy}^V \leq n + d_H^V + \hat{d}_H^V.
\]  

(7)

**PROOF.** Consider \(V\) as an \(H\)-module via \(f\). Set \(z = (xy)^{-1}\) and let \(C\) be the direct sum

\[
C := (1 - x)V \oplus (1 - y)V \oplus (1 - z)V.
\]

Define the linear transformations \(\beta : V \to C\) and \(\delta : C \to V\) respectively by

\[
v \mapsto ((1 - x)v, (1 - y)v, (1 - z)v),
\]

\[
(v_1, v_2, v_3) \mapsto v_1 + xv_2 + z^{-1}v_3.
\]

We have \(\dim(\text{Im} \beta) = n - d_H^V\), since \(\text{Ker} \beta\) is the space of fixed points of \(H\). Using the identity \(a(1 - b) = (1 - a)(b - 1) + (1 - b)\), for all \(a, b\) in the group algebra \(F H\), it is easy to deduce that \(\text{Im} \delta\) coincides with the subspace \((1 - x)V + (1 - y)V + (1 - z)V\) and, moreover, that \(\text{Im} \delta\) is \(H\)-invariant. Thus \(\text{Im} \delta\) is the smallest \(H\)-submodule of \(V\) with trivial action on the quotient. Let \(B = B_0 \cup B_1\) be a basis of \(V\) such that \(B_0\) is a basis of \(\text{Im} \delta\). With respect to \(B\), \(f(H)\) consists of matrices of the form

\[
\begin{pmatrix}
* & * \\
0 & I
\end{pmatrix}.
\]
This observation easily implies that $|B_1| = \hat{d}^H_V$, hence $\dim(\text{Im} \delta) = n - \hat{d}^H_V$.

From $\text{Im} \beta \leq \text{Ker} \delta \leq C$ we deduce

$$\dim C = \dim(\text{Im} \beta) + \dim \frac{\text{Ker} \delta}{\text{Im} \beta} + \dim \frac{C}{\text{Ker} \delta} \geq \dim(\text{Im} \beta) + \dim(\text{Im} \delta).$$

We conclude

$$\dim C = (n - d^x_V) + (n - d^y_V) + (n - d_V^H) \geq (n - \hat{d}^H_V)$$

whence $d^x_V + d^y_V + d^{xy}_V \leq n + d^H_V + \hat{d}^H_V$. \hfill \square

As noticed by L.L. Scott in [68], the genus formula (2) itself is a consequence of (7). The argument is the following. Assume that a $(2, 3, 7)$-generated group $G$ has a subgroup $S$ of index $n$. Let $f : G \to \text{GL}_n(\mathbb{C})$ be the linear representation of $G$ induced by the transitive permutational action on the (left) cosets of $S$ and let $V = \mathbb{C}^n$ be the corresponding $G$-module. For every $g \in G$ of prime order $r$ whose cyclic structure consists of $\ell$ non-trivial cycles, we have

$$n - d^g_V = (r - 1)\ell \leq (r - 1) \left\lfloor \frac{n}{r} \right\rfloor.$$

So, if $(x, y, xy)$ is a $(2, 3, 7)$-generating triple for $G$, then

$$\left\lfloor \frac{n}{2} \right\rfloor \geq n - d^x_V, \quad 2\left\lfloor \frac{n}{3} \right\rfloor \geq n - d^y_V, \quad 6\left\lfloor \frac{n}{7} \right\rfloor \geq n - d^{xy}_V,$$

hence

$$\left\lfloor \frac{n}{2} \right\rfloor + 2\left\lfloor \frac{n}{3} \right\rfloor + 6\left\lfloor \frac{n}{7} \right\rfloor \geq 3n - (d^x_V + d^y_V + d^{xy}_V).$$

By the transitivity, the multiplicity of the trivial representation is 1. Hence $d^G_V = \hat{d}^G_V = 1$, and Scott’s formula gives

$$d^x_V + d^y_V + d^{xy}_V \leq n + 2.$$

We conclude

$$\left\lfloor \frac{n}{2} \right\rfloor + 2\left\lfloor \frac{n}{3} \right\rfloor + 6\left\lfloor \frac{n}{7} \right\rfloor \geq 2n - 2.$$
and SL₂(3). But a more systematic application was first made in [25], where this formula was applied essentially to the following representations of an absolutely irreducible subgroup $H$ of SLₙ(F), with (2, 3, 7)-generating triple $(x, y, xy)$.

1. The conjugation action of $H$ on $M = \text{Mat}_n(F)$. In this case, the fixed-points subspace of $M$ is the centralizer of $H$ and therefore, by Schur's lemma, it consists of scalar matrices. Hence Scott's formula reads

$$d^x_M + d^y_M + d^{xy}_M \leq n^2 + 2.$$  

The values of the left-hand side of this equation are easily calculated using a well-known formula, due to F.G. Frobenius (e.g., see [39, p. 207, Theorem 3.16]). Namely, let $n_1 \leq \cdots \leq n_s$ be the degrees of the similarity invariants of $a \in M$. Then

$$d^a_M = \sum_{j=1}^s (2s - 2j + 1)n_j = (2s + 1)n - 2 \sum_{j=1}^s jn_j.$$  

(9)

In particular $d^a_M \geq n + s^2 - s$.

2. The diagonal action of $H$ on the symmetric square $S$ of $F^n$. Scott's formula takes the shape

$$d^x_M + d^y_M + d^{xy}_M \leq \frac{n(n+1)}{2} + 2.$$  

Moreover if $\frac{n(n+1)}{2} < d^x_M + d^y_M + d^{xy}_M \leq \frac{n(n+1)}{2} + 2$ then $H$ is orthogonal for $p \neq 2$ and $H$ is symplectic for $p = 2$. This claim was first stated in Lemma 4.1 of [25], but the proof in characteristic 2 was inaccurate. For a revised proof see [78, Lemma 2.1] or [84].

The values of the left-hand side of (10), for the relevant elements $g \in \text{GL}_n(F)$, are afforded by the following formulas (see [25]). Assume first that $g$ is semisimple. If $\nu$ is an eigenvalue of $g$, let $m_\nu$ denote the multiplicity of $\nu$. Then

$$d^g_S = \frac{m_1(m_1 + 1) + m_{-1}(m_{-1} + 1)}{2} + \sum m_\nu m_{\nu^{-1}},$$

where the summation runs over all pairs $\nu, \nu^{-1}$ of eigenvalues of $g$ in $F$ with $\nu \neq \nu^{-1}$. Next assume that $g$ is unipotent, of prime order $p$. Let $k_i$ be the number of similarity invariants of $g$ of degree $i$, $1 \leq i \leq p$. Then, if $p = 2$

$$d^g_S = \frac{k_1^2 + 2k_2^2}{2} + k_1k_2 + \frac{k_1 + 2k_2}{2};$$

otherwise

$$d^g_S = \sum_{i=1}^p \frac{ik_i^2}{2} + \sum_{i=1}^{p-1} \sum_{j=i+1}^p ik_ik_j + \sum_{i=0}^{p-1} \frac{k_{i+1}^2}{2}.$$
A detailed analysis of conjugacy classes and comparison of (8) with (10) lead to conclude that many classical linear groups of rank $\leq 19$ are not Hurwitz. As an example, we quote some of the results.

**Theorem 4.2** (Di Martino, Tamburini and Zalesski, [25]). Let $H$ denote an irreducible subgroup of $\text{SL}_n(F)$, with $n \in \{4, 5, 6, 7, 10\}$. Assume that $H$ is not contained in an orthogonal group if $p \neq 2$, and that $H$ is not contained in a symplectic group if $p = 2$. If $n = 6$ and $p = 2$, assume further that $H = \text{SL}_6(q)$ or $\text{SU}_6(q)$. Then $H$ is not $(2, 3, 7)$-generated. In particular, if $n \in \{4, 5, 6, 7, 10\}$, then

1. the groups $\text{SL}_n(q)$, $\text{Sp}_n(q)$, $\text{SU}_n(q^2)$ are not Hurwitz, with the only possible exception of $\text{Sp}_6(2^t)$, $n \geq 6$;
2. every complex irreducible character of degree $n$ of a $(2, 3, 7)$-generated group is real.

**Remark 4.3.** The case of $\text{Sp}_4(2^t)$ was not excluded in [25]. But the fact that the symplectic groups $\text{Sp}_4(2^t)$ are not Hurwitz can be deduced either from [50], where it is shown that they are not even $(2, 3)$-generated, or from our Theorem 4.16.

**Theorem 4.4** (Di Martino, Tamburini and Zalesski, [25]). Let $H$ denote an absolutely irreducible subgroup of $\text{SL}_n(Q)$, with $n \leq 19$ or $n = 22$. If $H$ is not contained in an orthogonal group, then $H$ is not $(2, 3, 7)$-generated. In particular, for these values of $n$, the group $\text{SL}_n(Z)$ is not $(2, 3, 7)$-generated.

We will not prove Theorems 4.2 and 4.4 here. We just observe that, when $n = 4, 5$, a stronger result holds. In fact there is now a complete classification of the irreducible Hurwitz subgroups of $\text{PSL}_4(\mathbb{F})$ and $\text{PSL}_5(\mathbb{F})$. For details see Section 4.3 and [82]. Recently R. Vincent and A. Zalesskii have extended Theorems 4.2 and 4.4 to other values of $n$, [84]. Their results depend on the residues of $q$ modulo 42.

### 4.2. Rigidity

The following definition is a special case of a more general one. Among the first who used it we quote G.V. Belyi, [1], and J.G. Thompson, [83]. But for more complete historical information, we refer to [62].

**Definition 4.5.** Let $a_1, a_2, a_3 \in \text{GL}_n(\mathbb{F})$ be such that $a_1a_2 = a_3$. The triple $(a_1, a_2, a_3)$ is called linearly rigid if, whenever $b_1, b_2, b_3$ are matrices such that $b_1b_2 = b_3$ and each $b_i$ is conjugate to $a_i$, there exists $g \in \text{GL}_n(\mathbb{F})$ such that $gb_ig^{-1} = a_i$, for $i = 1, 2, 3$.

Rigid generators of finite groups have been studied in the inverse Galois problem (see [62] and [85]). The same concept, under the name of physical rigidity (for $\mathbb{F} = \mathbb{C}$) appeared in the totally different context of linear differential equations and local systems on the sphere (see [43]).
In Section 4.3 we will illustrate some applications of linear rigidity to the context of Hurwitz generation, based on a useful criterion for recognizing rigid triples. In order to describe this criterion we recall some notation.

As above we set \( M = \text{Mat}_n(\mathbb{F}) \) and, for each \( a \in M \), \( \dim C_M(a) = d_M^a \).

**Theorem 4.6** (Strambach and Völklein, [74]). Assume that \( a_1, a_2 \in \text{GL}_n(\mathbb{F}) \) generate an irreducible subgroup. Set \( a_3 = a_1 a_2 \) and suppose that

\[
\sum_{i=1}^{3} d_M^{a_i} = n^2 + 2. \tag{11}
\]

Then the triple \( a_1, a_2, a_3 \) is linearly rigid.

**Proof.** Let \( b_i = a_i^{b_i} \) for \( i \leq 3 \), with \( b_3 = b_1 b_2 \). Consider the linear transformations \( \sigma_i \) of \( M = \text{Mat}_n(\mathbb{F}) \) defined by

\[
m \mapsto b_i^{-1} m a_i.
\]

For \( c \in M \), let us denote by \( \lambda_c \) and \( \rho_c \) the endomorphisms of \( M \) given by left and right multiplication by \( c \). Then

\[
\sigma_i = \lambda_{b_i^{-1}} \rho_{a_i} = \lambda_{g_i}^i \lambda_{g_i}^{-1} \lambda_{g_i} \rho_{a_i} = \lambda_{g_i}^{-1} (\lambda_{g_i} \rho_{a_i}) \lambda_{g_i}.
\]

Thus \( \sigma_i \) is conjugate in \( \text{GL}(M) \) to conjugation by \( a_i \). Therefore \( d_M^{a_i} = \dim C_M(a_i) \), hence \( d_M^{a_i} = d_M^{a_i} \). Set \( H = \langle \sigma_1, \sigma_2 \rangle \). Since \( H \leq \text{GL}(M) \), we may consider \( M \) as an \( H \)-module. Thus, applying Theorem 4.1, we obtain

\[
\sum_{i=1}^{3} d_M^{\sigma_i} \leq n^2 + d_M^H + \hat{d}_M^H.
\]

Together with assumption (11) this yields

\[
n^2 + 2 \leq n^2 + d_M^H + \hat{d}_M^H.
\]

It follows that \( d_M^H > 0 \) or \( \hat{d}_M^H > 0 \). Thus there exists a non-zero matrix \( g \) such that either \( b_i^{-1} g a_i = g \) or \( b_i^{-1} g a_i^t = g \), for \( i = 1, 2, 3 \). We claim that \( g \) is non-singular. Otherwise let \( W \) be the eigenspace of \( g \) relative to the eigenvalue 0. In the first case \( W \) would be invariant under the irreducible subgroup \( \langle a_1, a_2 \rangle \). In the second case it would be invariant under its transpose, again a contradiction. We conclude either \( b_i = a_i^{g^{-1}} \) or \( b_i = a_i^{g^t} \), for \( i \leq 3 \). \( \square \)

The above proof shows that \( d_M^H > 0 \) if and only if \( \hat{d}_M^H > 0 \). Moreover, in this case, both of them must be 1 because, by Schur’s lemma, the centralizer in \( \text{Mat}_n(\mathbb{F}) \) of the irreducible group \( \langle a_1, a_2 \rangle \) consists of scalar matrices.
REMARK 4.7. If there are $a_1, a_2$ and $a_3 = a_1 a_2 \in \text{GL}_n(\mathbb{F})$ that satisfy (11) but generate a reducible subgroup of $\text{GL}_n(\mathbb{F})$, then theorem 4.6 also implies that no other triple with the same set of similarity invariants can generate an irreducible subgroup of $\text{GL}_n(\mathbb{F})$. However there may be more that one conjugacy class of triples generating reducible subgroups.

To conclude this section we quote the following result which, ultimately, depends on a well known theorem of S. Lang and R. Steinberg (see, e.g., [72]).

**THEOREM 4.8.** Let $(a_1, a_2, a_3)$ be a linearly rigid triple, with $a_i \in \text{GL}_n(\mathbb{F})$. Let $C_i$ be the conjugacy class of $a_i$ and suppose that $C_i \cap \text{GL}_n(q)$ (respectively, $C_i \cap U(n, q^2)$) is non-empty, for $i = 1, 2, 3$. Then there exists $g \in \text{GL}_n(\mathbb{F})$ such that $g a_i g^{-1} \in \text{GL}_n(q)$ (respectively $\in U(n, q^2)$), for $i = 1, 2, 3$.

### 4.3. Classical groups of small rank which are Hurwitz

**THEOREM 4.9 (Macbeath, [57]).** Let $k$ be a prime number $\geq 7$. If $p > 0$ and $p \neq k$ denote by $n$ the order of $p$ modulo $k$. The group $\text{PSL}_2(\mathbb{F})$ contains exactly one isomorphism type of $(2, 3, k)$-generated subgroups, namely

1. $\Delta(2, 3, k)$, if $p = 0$;
2. $\text{PSL}_2(p)$, if $p = k$;
3. $\text{PSL}_2(p^n)$, if $p \neq k$ and $n$ is odd;
4. $\text{PSL}_2(p^2)$, if $p \neq k$ and $n$ is even.

Moreover, if $p > 0$, there is just one conjugacy class of such groups.

REMARK 4.10. We recall that $\text{PSL}_2(q)$ has order $q(q+1)(q-1)/(2, q-1)$. So the groups listed in items (2), (3) and (4) in the statement correspond precisely to the smallest power of $p$ such that $\text{PSL}_2(q)$ has order divisible by $k$.

**PROOF.** The assumption $k \geq 7$ implies that any $(2, 3, k)$-generated group is perfect, hence non-soluble. The latter fact will be used in the following without further mention.

If $p \neq k$, define $\varepsilon \in \mathbb{F}$ to be a primitive $k$-th root of unity. If $p = k$, put $\varepsilon = 1$.

First, we consider the case when $p = 0$. Note that, for each $\ell$ such that $(\ell, k) = 1$, the projective image of

\[
\begin{pmatrix}
0 & -1 \\
1 & 0
\end{pmatrix}, \quad
\begin{pmatrix}
0 & \varepsilon^{-\ell} \\
-e^\ell & -1
\end{pmatrix}, \quad
\begin{pmatrix}
\varepsilon^\ell & 1 \\
0 & \varepsilon^{-\ell}
\end{pmatrix}
\]

(12)
is a $(2, 3, k)$-triple. The group $\langle x, y_1 \rangle$ is isomorphic to $\langle x, y_\ell \rangle$, for each $\ell$, under the automorphism of $\text{Mat}_2(\mathbb{Z}[\varepsilon])$ induced by the map $\varepsilon \mapsto \varepsilon^\ell$. On the other hand, a slight modification of the arguments given in the proof of theorem 1 in [25] shows that the preimage of a $(2, 3, k)$-generated subgroup of $\text{PSL}_2(\mathbb{F})$ must be conjugate to $\langle x, y_1 \rangle$, for some $\ell$. Our claim follows from the classical embedding of $T(2, 3, k)$ into $\text{PSL}_2(\mathbb{C})$ (see [59, Theorem 2.8]).
Now assume $p > 0$. Let $q = p$ if $p = k$; $q = p^n$ if $p \neq k$ and $n$ is odd; $q = p^{n/2}$ if $p \neq k$ and $n$ is even. So $\theta_{\ell} = \epsilon^\ell + \epsilon^{-\ell}$ is an element of $\text{GF}(q)$. For each $\ell$ such that $1 \leq \ell \leq \frac{k-1}{2}$, define
\[
x = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad y_\ell = \begin{pmatrix} b & a \\ a - \theta_{\ell} & -1 - b \end{pmatrix},
\]
x_{xy_\ell} = \begin{pmatrix} \theta_{\ell} - a & 1 + b \\ b & a \end{pmatrix},
(13)
where
\[
a(a - \theta_{\ell}) + b(1 + b) = -1.
(14)
\]
If $p = 2$, we can take $a = b = (1 + \theta_{\ell})^{-1}$. For $p > 2$, equation (14) is equivalent to
\[
(2a - \theta_{\ell})^2 + (2b + 1)^2 = -3 + \theta_{\ell}^2,
\]
which is always solvable over $\text{GF}(q)$ since every element of a finite field is a sum of two squares. Thus, $(x, y_\ell) \leq \text{SL}_2(q)$ and the projective image of $(x, y_\ell, x_{xy_\ell})$ is a $(2, 3, k)$-triple. As observed in Remark 4.10, $n$ and $q$ are defined so that $\text{SL}(2, q_0)$ does not have elements of projective order $k$, for any proper divisor $q_0$ of $q$. It follows from Dickson’s classification of the subgroups of $\text{PSL}_2(q)$ (see, for example, [20, Chapter XII] or [36, 8.27]) that the perfect group $\langle x, y_\ell \rangle$ coincides with $\text{SL}_2(q)$.

On the other hand, every triple $(x', y', x'y')$ in $\text{SL}_2(\mathbb{F})$, whose projective image is a $(2, 3, k)$-triple, is such that $x' \sim x$, $y' \sim y_\ell$ and $x'y' \sim x_{xy_\ell}$ for some $\ell$, where $x$ and $y_\ell$ are as in (13). Moreover we can assume that $1 \leq \ell \leq \frac{k-1}{2}$. Thus our final claim follows from Theorem 4.6.

Actually the factorizations in (12) and (13), may be viewed as a special case of the following (constructive) factorization theorem for matrices.

**Theorem 4.11** (Sourour, [70]). Let $a \in \text{GL}_n(\mathbb{F})$ be non-scalar, and let $\beta_i$ and $\gamma_i$ $(1 \leq i \leq n)$ be elements of $\mathbb{F}$ such that
\[
\prod_{j=1}^{n} \beta_{ij} \gamma_{ij} = \det a.
\]
There exist $b$ and $c$ in $\text{GL}_n(\mathbb{F})$, with respective eigenvalues $\beta_i$ and $\gamma_i$, such that $a = bc$.

In order to classify the Hurwitz subgroups of $\text{PSL}_n(\mathbb{F})$, it is necessary to keep in mind the irreducible representations of $\text{PSL}_2(q)$, when Hurwitz. In the natural characteristic they are described in [3]. But we prefer to give an independent proof of what is relevant for us.
In what follows we consider a field \( K \) of characteristic \( p \), and describe a bunch of absolutely irreducible representations of \( \text{SL}_2(\mathbb{K}) \) over \( K \). For any automorphism \( \sigma \) of the field \( K \), \( \text{SL}_2(\mathbb{K}) \) acts on the polynomial ring \( \mathbb{K}[t_1, t_2] \) via

\[
t_2^l \mapsto (\sigma(a_{11})t_1 + \sigma(a_{21})t_2)^l (\sigma(a_{12})t_1 + \sigma(a_{22})t_2)^l,
\]

where \( (a_{11}, a_{12}) \) is in \( \text{SL}_2(\mathbb{K}) \). For any \( m \) and any \( \sigma \), the space of homogeneous polynomials of degree \( m \) is invariant under this action. We will denote this module by \( V_\sigma^m \).

**Theorem 4.12.** Let \( m \leq p - 1 \) if \( p > 0 \) and \( m \geq 1 \) if \( p = 0 \). The \( \text{SL}_2(\mathbb{K}) \)-module \( V_\sigma^m \) is absolutely irreducible for each \( \sigma \).

**Proof.** Let \( \overline{K} \) be the algebraic closure of \( K \) and let \( U \) be a non-zero \( \text{SL}_2(\mathbb{K}) \)-invariant subspace of \( V_\sigma^m \otimes \overline{K} \). Let us fix \( 0 \neq f \in U \). We first show that \( t_2^m \in U \). This is clear if \( f = \lambda_2 t_2^m \), otherwise write

\[
f(t_1, t_2) = \lambda_d t_1^d t_2^{m-d} + \lambda_{d-1} t_1^{d-1} t_2^{m-d+1} + \cdots,
\]

where \( \lambda_d \neq 0 \). Let \( \Delta \) be the difference operator

\[
(\Delta f)(t_1, t_2) = f(t_1, t_2) - f(t_1 - t_2, t_2).
\]

Then \( \Delta f \in U \) and its \( d \)-th iterate \( \Delta^{(d)} \) satisfies

\[
(\Delta^{(d)} f)(t_1, t_2) = d! \lambda_d t_2^m.
\]

Note that \( d! \neq 0 \) in \( \mathbb{K} \) for any \( p \) (for \( p > 0 \) we use \( d \leq p - 1 \)). Therefore \( t_2^m \in U \). Now, for \( j = 0, 1, \ldots, m \), the polynomials \( t_2^j \) are in \( U \) and are linearly independent. We conclude that \( U = V_m \otimes \overline{K} \). \( \square \)

For each \( \sigma \) and \( m \geq 1 \), the above action has a non-trivial kernel, namely \((-I)\), exactly when \( q \) is odd and \( m \) is even. Otherwise it is faithful. Moreover, this action consists of linear transformations of determinant 1. Thus it gives an embedding of \( \text{PSL}_2(\mathbb{K}) \) into \( \text{SL}_{m+1}(\mathbb{K}) \) when \( m \) is even, and into \( \text{PSL}_{m+1}(\mathbb{K}) \) when \( m \) is odd. Thus the above result, together with Theorem 4.9, allows to construct projective representations of \( T(2, 3, 7) \) of degree \( m \), for each \( m \leq p \).

**Theorem 4.13** (Cohen, [6]). Any Hurwitz subgroup \( \overline{H} \) of \( \text{PSL}_3(\mathbb{F}) \) has a preimage \( H \) in \( \text{SL}_3(\mathbb{F}) \) which is also Hurwitz. If \( H \) is reducible, then \( p = 2 \), \( H \cong \text{PSL}_2(8) \cong \text{SL}_2(8) \), and there are two conjugacy classes of such groups. If \( H \) is irreducible, one of the following holds:

1. \( p \neq 7 \) and \( H \cong \text{PSL}_2(7) \);
2. \( p \equiv 0, \pm 1 \pmod{7} \) and \( H = \Omega_3(p) \cong \text{PSL}_2(p) \), or \( 2 < p \equiv \pm 2, \pm 3 \pmod{7} \) and \( H = \Omega_3(p^3) \cong \text{PSL}_2(p^3) \).

Moreover, in both cases (1) and (2) there is just one conjugacy class of Hurwitz groups.
PROOF. Let \((\tilde{x}, \tilde{y}, \tilde{x} \tilde{y} = \tilde{z})\) be a \((2, 3, 7)\)-generating triple for \(\overline{H}\). First, we show that we can choose \(x, y \in \text{SL}_3(\overline{F})\) such that \(x \mapsto \tilde{x}, y \mapsto \tilde{y}\) and \((x, y, xy = z)\) is also a \((2, 3, 7)\)-triple. This is true if the characteristic of \(\overline{F}\) is 3 since \(\text{PSL}_3(\overline{F}) = \text{SL}_3(\overline{F})\) in that case. So suppose that the characteristic is different from 3. Multiplying \(x\) and \(y\) by scalar matrices, if necessary, we can assume that \(x^2 = I\) and \(z^2 = (xy)^7 = I\). In particular, \(x\) has eigenvalues \(-1\) (with multiplicity 2) and 1. We claim that \(y\) cannot be a matrix of order 9 such that \(y^3\) is scalar. In fact, in this case, \(y\) is diagonalizable with eigenvalues \(\eta\), with multiplicity 2, and \(\eta^7\), for some primitive 9-th root of unity \(\eta\). Therefore there exists a non-zero vector \(v\) such that \(xv = -v, yv = \eta v\). It follows \(zv = -\eta^2 v\), a contradiction since \(z\) has order 7. We conclude that \(y\) must have order 3.

Set \(H = \langle x, y \rangle\). Assume that \(H\) is reducible. Replacing \(H\) with its transpose, if necessary, we may assume that \(H\) fixes a one-dimensional space \(U\), hence a non-zero vector. In odd characteristic, the action of \(x\) modulo \(U\) should be scalar. From here it is easy to deduce that \(H\) would be soluble, a contradiction. Thus the characteristic is 2 and, by Theorem 4.9, \(H\) should induce on the quotient \(\overline{F}^3 / U\) the group \(\text{PSL}_2(8)\). In particular, up to conjugation, we can take

\[
\begin{bmatrix}
\varepsilon & 1 & 0 \\
0 & \varepsilon^{-1} & 0 \\
0 & 0 & 1 
\end{bmatrix},
\begin{bmatrix}
0 & 1 & 0 \\
1 & 0 & 0 \\
\alpha & \alpha & 1 
\end{bmatrix},
\]

where \(\varepsilon^7 = 1\) and \(\alpha\) is either 0 or 1. Direct computation shows that both choices of \(\alpha\) give a group isomorphic to \(\text{PSL}_2(8)\). Moreover these two groups are not conjugate.

Now suppose that \(H\) is irreducible. Then the similarity invariants must be \(t + 1, t^2 - 1\) for \(x, t^3 - 1\) for \(y\) and \((t - \varepsilon^\ell)(t - \varepsilon^{2\ell})(t - \varepsilon^{4\ell})\) with \(\ell = 1, 3\) or \((t - 1)(t - \varepsilon^\ell)(t - \varepsilon^{-\ell})\) with \(\ell = 1, 2, 3\) for \(z\). Thus, by Theorem 4.6, \(H\) belongs to at most five conjugacy classes (just one, if \(p = 7\), corresponding to the five possibilities for \(xy\).

1. If \(p \neq 7\), then \(\text{PSL}_2(7)\) has two dual irreducible representations of degree 3 over \(\overline{F}\).

   This fact can be deduced from the knowledge of its ordinary and Brauer characters.

   These representations exhaust the first two possibilities for \(xy\).

2. As to the remaining possibilities for \(xy\), let us consider the embedding of \(\text{PSL}_2(q)\) into \(\text{SL}_3(q)\) described just before Theorem 4.12, with \(\sigma = 1\) and \(m = 2\). Under this embedding, the three non-conjugate \((2, 3, 7)\)-triples generating \(\text{PSL}_2(q)\) are mapped to three non-conjugate triples in \(\text{SL}_3(\overline{F})\). Moreover the image of \(\text{PSL}_2(q)\) preserves a symmetric (non-zero) bilinear form on the space of homogeneous polynomials of degree 2. This form is non-degenerate precisely when \(p > 2\) (see also the proof of Theorem 1 in [25]). Thus, when \(p = 2\), the remaining possibilities for \(xy\) do not give rise to irreducible subgroups of \(\text{SL}_3(\overline{F})\), by Remark 4.7. On the other hand, when \(p > 2\), this embedding exhausts the remaining possibilities for \(xy\) giving rise to an irreducible subgroup of \(\Omega_3(q)\). Our claim (2) follows from the isomorphism \(\text{PSL}_2(q) \cong \Omega_3(q)\) and Theorem 4.9. \(\square\)

The following result deals with classical groups of rank 4. The first statement is related to a fact already proved by Macbeath in [57]. Namely that, when \(p \equiv \pm 1 \pmod{7}\), there are three normal subgroups of \(T(2, 3, 7)\) with quotient isomorphic to \(\text{PSL}_2(p)\).
LEMMA 4.14. Let \( p > 0 \) and let \( \overline{H} \) be a Hurwitz subgroup of \( \text{PSL}_2(\mathbb{F}) \times \text{PSL}_2(\mathbb{F}) \). Assume that \( \overline{H} \) is not isomorphic to the Hurwitz subgroup of \( \text{PSL}_2(\mathbb{F}) \). Then \( p \equiv \pm 1 \pmod{7} \) and \( \overline{H} \) is conjugate to \( \text{PSL}_2(p) \times \text{PSL}_2(p) \).

PROOF. For \( i = 1, 2 \) let \( \pi_i \) be the projections of \( \text{PSL}_2(\mathbb{F}) \times \text{PSL}_2(\mathbb{F}) \) onto \( \text{PSL}_2(\mathbb{F}) \). Since \( \text{Ker} \pi_i \cong \text{PSL}_2(\mathbb{F}) \), our assumption implies that \( 1 \neq \pi_1(\overline{H}) \) and \( 1 \neq \pi_2(\overline{H}) \). Hence \( \pi_1(\overline{H}) \cong \pi_2(\overline{H}) \cong \text{PSL}_2(q) \), with \( q \in \{ p, p^3 \} \) as in Theorem 4.9, case \( k = 7 \). By the same theorem, up to conjugation we may assume that \( \overline{H} \) is the image of \( \langle (x, x), (y_\ell, y_m) \rangle \leq \text{SL}_2(q) \times \text{SL}_2(q) \), \( \ell \neq m \) by our assumptions. In particular, \( p \neq 7 \).

If \( q = p^3 \) with \( p \equiv \pm 2, \pm 3 \pmod{7} \), then there exists a field automorphism \( \sigma \) of \( \text{PSL}_2(q) \) such that \( x = \sigma(x) \) and \( y_m = \sigma(y_\ell) \). We conclude that \( \overline{H} \) is isomorphic to \( \text{PSL}_2(q) \).

Now assume \( q = p \equiv \pm 1 \pmod{7} \). We note that

\[
\pi_i(\text{Ker} \pi_j \cap \overline{H}) \subseteq \pi_i(\overline{H}) = \text{PSL}_2(q).
\]

If \( \pi_1(\text{Ker} \pi_2 \cap \overline{H}) \) is trivial, it follows that \( \text{Ker} \pi_1 \cap \overline{H} = 1 \). In this case the restriction \( \pi_1 : \overline{H} \to \text{PSL}_2(q) \) would be an isomorphism, against our assumption. Thus \( \pi_1(\text{Ker} \pi_2 \cap \overline{H}) = \text{PSL}_2(q) \) by the simplicity of \( \text{PSL}_2(q) \), i.e. \( \text{Ker} \pi_2 \cap \overline{H} = \text{PSL}_2(q) \). From \( \pi_2(\overline{H}) = \text{PSL}_2(q) \) we conclude that \( \overline{H} = \text{PSL}_2(q) \times \text{PSL}_2(q) \).

For any field \( K \), consider the homomorphism \( \varphi : \text{SL}_2(K) \times \text{SL}_2(K) \to \text{SL}_4(K) \)

\[
(a, b) \mapsto a \otimes b.
\]

The kernel of \( \varphi \) is \( (\{-I\}, \{-I\}) \). Moreover the image of \( \varphi \) preserves the bilinear symmetric form defined by the matrix

\[
\begin{pmatrix}
0 & 1 \\
-1 & 0
\end{pmatrix} \otimes \begin{pmatrix}
0 & 1 \\
-1 & 0
\end{pmatrix} = \text{antidiag}(1, -1, -1, 1).
\]

Hence the image of \( \varphi \) is an orthogonal group in odd characteristic, a symplectic group in characteristic 2.

REMARK 4.15. Assume \( K = \text{GF}(q) \), with \( q \in \{ p, p^3 \} \) as in the theorem of Macbeath, \( p \neq 7 \), and let \( H \) be defined as in (15) with \( \ell \neq m \). Then \( y_\ell \otimes y_m \) does not have the eigenvalue 1. It follows that \( \varphi(H) \) does not fix any one-dimensional subspace and this fact easily implies that it is absolutely irreducible. Moreover \( \varphi(H) \) is a subgroup of \( \Omega^+_4(q) \) for \( p \) odd, of \( \text{Sp}_4(8) \) for \( p = 2 \). In particular, when \( p \equiv \pm 1 \pmod{7} \), by order reasons \( \varphi \) induces an isomorphism from the Hurwitz central product \( \text{SL}_2(p) \circ \text{SL}_2(p) \) onto
\( \Omega_4^+(p) \). Factorizing this central product by its center we obtain the Hurwitz direct product 
\( \text{PSL}_2(p) \times \text{PSL}_2(p) \cong \Omega_4^+(p) \). The details are left to the reader.

**Theorem 4.16.** If \( p > 0 \), the irreducible subgroups of \( \text{PSL}_4(\mathbb{F}) \) which are Hurwitz are isomorphic to:

1. \( \text{PSL}_2(p) \times \text{PSL}_2(p) \cong \Omega_4^+(p) \), when \( p \equiv \pm 1 \) (mod 7);
2. \( \text{PSL}_2(q) \) with \( q \in \{ p, p^3 \} \) as in Macbeath’s theorem;
3. \( \text{PSL}_2(7) \), when \( p \neq 2 \).

In particular \( \Omega_4^+(q) \cong \text{PSL}_2(q) \times \text{PSL}_2(q) \) is Hurwitz if and only if \( q = p \equiv \pm 1 \) (mod 7), whereas the following groups are never Hurwitz: \( \text{PSL}_4(q) \cong \Omega_6^-(q) \), \( \text{PSU}_4(q^2) \cong \Omega_6^+(q) \), and \( \text{PSp}_4(q) \cong \Omega_5(q) \).

**Proof.** Let \( x, y \in \text{SL}_4(\mathbb{F}) \) be such that the projective image of \( (x, y, xy) \) is a \((2, 3, 7)\)-generating triple of an irreducible subgroup of \( \text{PSL}_4(\mathbb{F}) \). Multiplying \( x \) and \( y \) by scalar matrices of determinant 1, if necessary, we may assume that \( y^3 = (xy)^7 = 1 \) and

\[
x^2 = I, \quad \text{if} \ p = 2; \quad x^2 \in (iI), \quad \text{where} \ i \in \mathbb{F} \ \text{has order} \ 4, \ \text{if} \ p > 2.
\]

By Scott’s formula, in the notation of Section 4.1,

\[
d_M^x + d_M^y + d_M^{xy} \leq 18. \quad (17)
\]

Direct calculation based on the formula (9) of Frobenius shows that

\[
d_M^x \geq 8, \quad d_M^y \geq 6, \quad d_M^{xy} \geq 4. \quad (18)
\]

It follows that in (17) and (18) we have all equalities. In particular \( x \) must have two equal similarity invariants, namely \( t^2 - 1, t^2 - 1 \), or \( t^2 + 1, t^2 + 1 \). In the first case \( x^2 = I \), in the second \( x^2 = -I \). On the other hand \( y \) must have similarity invariants \( t - 1 \) and \( t^3 - 1 \); \( xy \) must have a unique similarity invariant. Thus, when \( p = 7 \), \( xy \) is conjugate to the Jordan block of order 4. When \( p \neq 7 \), \( xy \) has 4 different eigenvalues and its similarity invariants can only be either

\[
(t - \epsilon^\ell)(t - \epsilon^{-\ell})(t - \epsilon^{2\ell})(t - \epsilon^{-2\ell}), \quad \text{with} \ \ell = 1, 2, 3,
\]

or

\[
(t - 1)(t - \epsilon^\ell)(t - \epsilon^{2\ell})(t - \epsilon^{4\ell}), \quad \text{with} \ \ell = 1, 3.
\]

Assume first \( x^2 = I \) and \( p > 2 \). By what shown in [25], \( (x, y) \) is a subgroup of the orthogonal group \( \Omega_4(\mathbb{F}, f) \), where \( f \) is a non-degenerate quadratic form of Witt index 2. Since \( \text{PSL}_4(\mathbb{F}, f) \) is isomorphic to \( \text{PSL}_2(\mathbb{F}) \times \text{PSL}_2(\mathbb{F}) \) (see [21]), by the previous lemma the projective image of \( (x, y) \) can only be of type (1) or of type (2).

Remark 4.15 tells us that (1) actually occurs, and that we obtain an irreducible subgroup of type (2) whenever \( p \neq 7 \).
Now assume either \( p = 2 \) and \( x^2 = I = -I \), or \( p \) odd and \( x^2 = -I \). Equality in relation (17) implies that \( (x, y, xy) \) is a linearly rigid triple, by Theorem 4.6. Thus \( (x, y) \) belongs to at most five conjugacy classes of irreducible subgroups (just one, if \( p = 7 \)) corresponding to the five possibilities for \( xy \). Moreover, if \( SL_4(F) \) contains a reducible subgroup generated by a triple \((x', y', x'y')\) such that \( x \sim x'\), \( y \sim y'\), \( xy \sim x'y'\), then the triple \((x, y, xy)\) cannot generate an irreducible subgroup; see Remark 4.7.

Let \( p \in \{p, p^2\} \) be defined as in the theorem of Macbeath. If \( p \geq 2 \), we consider the embeddings of \( SL_2(q) \) into \( SL_4(F) \) arising from the action of \( SL_2(q) \) on homogeneous polynomials of degree 3. They are irreducible when \( p \neq 3 \). If \( p = 2 \), we consider the irreducible embeddings of \( SL_2(8) \) into \( SL_4(F) \) corresponding to the projective image of \( \varphi(H) \), as in Remark 4.15. These embeddings of \( SL_2(q) \) exhaust the first three possibilities for \( xy \). Hence they give rise to a unique conjugacy class of irreducible subgroups, whenever \( p \neq 3 \). On the other hand, when \( p = 3 \), they are reducible. By what was observed above there is no irreducible subgroup generated by a triple of this kind.

Finally, when \( p \neq 2 \), \( SL_2(7) \) has faithful irreducible representations of degree 4 over \( F \), the existence of which can be deduced from the knowledge of ordinary, modular and Brauer characters of \( SL_2(7) \). They give rise to projective representations of \( SL_2(7) \) which exhaust the remaining cases for \( xy \). On the other hand, when \( p = 2 \), there are copies of \( SL_2(7) \sim SL_3(2) \) arising from the embedding of \( SL_3(F) \) into \( SL_4(F) = PSL_4(F) \) which exhaust the possibilities for \( xy \). Since this embedding is reducible, there is no irreducible subgroup of this kind when \( p = 2 \).

For the reader’s convenience, we give explicitly these representations of \( SL_2(7) \) for \( p \neq 2 \). When \( p = 7 \), we have the above representation on the homogeneous polynomials of degree 3. So we may assume \( p \neq 7 \). Let \( \varepsilon \) be a primitive 7-th root of unity in \( F \) and let \( \tau = \varepsilon + \varepsilon^2 + \varepsilon^4 \) or \( \tau = \varepsilon^3 + \varepsilon^6 + \varepsilon^5 \). Define

\[
x := \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad y := \begin{pmatrix} 1 & 0 & 0 & \tau \\ 0 & 1 & 0 & 1 + \tau \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.
\]

Direct calculation shows that for both values of \( \tau \) we have

\[
x^2 = -I, \quad y^3 = (xy)^7 = I, \quad [x, y]^4 = -I,
\]

and, moreover, that \( (x, y) \) fixes only the zero vector. It follows easily that this group is irreducible. Since \( SL_2(7) \) has the presentation \( x^2 = y^3 = (xy)^7 = [x, y]^4 = 1 \), the projective image of \( (x, y) \) is isomorphic to \( SL_2(7) \).

The remaining claims which do not follow directly from what shown above are consequences of Theorem 4.9 and of the isomorphisms in the statement. These isomorphisms are well known and can be found in [17, page xii].

**Theorem 4.17** (Tamburini and Zalesskii, [82]). Assume that \( k \geq 7 \) is a prime number. If \( k = p \), set \( q = p \). Otherwise, let \( n \) be the order of \( p \) modulo \( k \), and suppose \( n \not\equiv 0 \pmod{4} \). Set \( q = p^n \) if \( n \) is odd, \( q = p^2 \) if \( n \) is even.
(1) The following groups are $(2, 3, k)$-generated:
- $\text{PSL}_5(q)$, if $p \equiv 1 \pmod{5}$;
- $\text{PSU}_5(q^2)$, if $p \equiv -1 \pmod{5}$;
- $\text{PSU}_5(q^4)$, if $p \equiv \pm 2 \pmod{5}$.

(2) If $k = 7$, the only other irreducible Hurwitz subgroups of $\text{PSL}_5(F)$ are isomorphic to $\text{PSL}_2(q)$ with $q$ as above, except $\text{PSL}_2(8)$ and $\text{PSL}_2(27)$.

Note that the assumptions on $n$ are certainly satisfied if $k \equiv 3 \pmod{4}$. In particular, for $k = 7$, this theorem gives a new example of a Hurwitz group in each characteristic $p \neq 0, 5$. The proof, which avoids calculations, relies on the following results.

We recall that a pair $(b, c)$ of elements in $\text{Mat}_n(F)$ is said to be spectrally complete with respect to the product if, for every $\lambda_1, \ldots, \lambda_n \in F$ such that $\lambda_1 \cdots \lambda_n = \det(bc)$, there exist $b'$ conjugate to $b$ and $c'$ conjugate to $c$ under $\text{GL}_n(F)$, such that $b'c'$ has eigenvalues $\lambda_1, \ldots, \lambda_n$.

**Theorem 4.18** (Silva, [69]). Assume $n \geq 3$ and let $b, c \in \text{GL}_n(F)$. Denote by

$$f_1(t), \ldots, f_r(t), g_1(t), \ldots, g_s(t)$$

the non-trivial similarity invariants of $tI_n - b$ and of $tI_n - c$, respectively. Then $(b, c)$ is spectrally complete with respect to the product provided that $r + s \leq n$ and at least one of the polynomials $f_i(t)$ or $g_j(t)$ has degree $\neq 2$.

This theorem, at least in characteristic $p \neq k$, guarantees the existence of a $(2, 3, 5k)$ triple $(x, y, \eta z)$, where $(\eta z)^k$ is scalar, of order 5. It follows immediately that $(x, y)$ is an irreducible subgroup of $\text{SL}_5(F)$. Then the proof that the projective image of $(x, y)$ coincides with one of the above groups is based on the knowledge of the subgroups of $\text{SL}_5(K)$, when $K$ is a finite field, and Theorem 4.8.

5. Related results

5.1. Number-theoretic aspects

In this section we give an alternative description of some triangle groups $T(2, 3, k)$, and in particular $T(2, 3, 7)$, as two-dimensional projective unitary groups over rings of algebraic integers. The corresponding unitary groups can be identified, in turn, with groups of principal units in certain orders of generalized quaternion algebras. This description sheds more light on number-theoretic aspects of Hurwitz generation and provides a new view on some classical results like Macbeath's Theorem 3.1. The treatment in this section mainly follows, [88].

Let $F = \mathbb{C}$ and define $\varepsilon \in \mathbb{C}$ to be a primitive $2k$-th root of unity if $k$ is even, and a primitive $k$-th root of unity if $k$ is odd. We also set

$$\eta = \varepsilon - \varepsilon^{-1}, \quad \theta = \varepsilon + \varepsilon^{-1}.$$
It follows from the proof of Theorem 4.9 that \( T(2, 3, k) \) is isomorphic to the projective image of the group generated by matrices \( x \) and \( z = z_1 \) (or \( x \) and \( y_1 \)), which were defined in (12). It is immediate to verify that both \( x \) and \( z \) preserves the Hermitian form defined by the matrix
\[
B = \begin{pmatrix}
\eta^2 & \eta \\
-\eta & \eta^2
\end{pmatrix},
\]
which is non-degenerate provided \( k \neq 6 \). Thus, \( T(2, 3, k) \) is isomorphic to the projective image of a subgroup of
\[
SU(2, B, \mathbb{Z}[\varepsilon]) = \{ A \in SL(2, \mathbb{Z}[\varepsilon]) \mid \bar{A}^T BA = B \}.
\]
The problem when \( \langle x, z \rangle \) coincides with \( SU(2, B, \mathbb{Z}[\varepsilon]) \) was studied in [88].

**THEOREM 5.1 (Vsemirnov, Mysovskikh and Tamburini, [88]).** The equality \( \langle x, z \rangle = SU(2, B, \mathbb{Z}[\varepsilon]) \) holds if and only if \( k \in \{ 2, 3, 4, 5, 7, 9, 11 \} \). In particular, \( T(2, 3, k) \cong PSU(2, B, \mathbb{Z}[\varepsilon]) \) precisely for the values of \( k \) listed above.

In fact, the proof in [88] shows more. Namely, for \( k \) odd \( \geq 13 \) or \( k \) even \( \geq 6 \), the group generated by \( x \) and \( z \) has infinite index in \( SU(2, B, \mathbb{Z}[\varepsilon]) \). As C. Maclachlan noticed in a private correspondence, [58], this ‘negative’ result can be also deduced from the description of all arithmetic Fuchsian groups given by K. Takeuchi, [75, 76]. However, the ‘positive’ part of Theorem 5.1 is more delicate: it ensures not only that \( T(2, 3, k) \) is arithmetic for \( k = 7, 9, \) and 11, or equivalently, that it has a finite index in \( PSU(2, B, \mathbb{Z}[\varepsilon]) \), but also proves the coincidence of these two groups.

It is convenient to treat the elements of \( SU(2, B, \mathbb{Z}[\varepsilon]) \) as quaternions of norm 1. For this purpose, we note that the matrices
\[
1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad j = \begin{pmatrix} \eta & 1 \\ 1 & -\eta \end{pmatrix}, \quad k = \begin{pmatrix} -1 & \eta \\ \eta & 1 \end{pmatrix}
\]
form a standard basis over \( \mathbb{Q}(\theta) \) for the generalized quaternion algebra \( (-1, \theta^2 - 3) \). Moreover, for any element (given as a matrix) in this algebra, its quaternion norm coincides with the determinant. It is then easy to see that \( SU(2, B, \mathbb{Z}[\varepsilon]) \) is exactly the set \( \mathcal{H}_1^{*} \) of all quaternions of norm 1 in the subring
\[
\mathcal{H} = \{ a_01 + a_1i + a_2j + a_3k \mid 2a_i \in \mathbb{Z}[\theta], \ a_0 - a_3 - a_2\theta \in \mathbb{Z}[\theta], \ a_1 + a_2 - a_3\theta \in \mathbb{Z}[\theta] \}.
\]
Thus, Theorem 5.1 can be restated in the following way.
THEOREM 5.2. We have $\langle x, z \rangle = H_1^+$ if and only if $k \in \{2, 3, 4, 5, 7, 9, 11\}$. In particular, $T(2, 3, k) \cong H_1^*/\{\pm1\}$ precisely for the values of $k$ listed above.

Note that the norm in the corresponding quaternion algebra is given by

$$a_0^2 + a_1^2 - (\theta^2 - 3)a_2^2 - (\theta^2 - 3)a_3^2.$$  

For a given $k$ these norm forms depend on $\theta$, i.e. on the choice of a primitive root of unity, and there is a natural action of $\text{Gal}(\mathbb{Q}(\theta)/\mathbb{Q})$ on their coefficients. If $k = 2, 3, 4$ or 5 all corresponding forms are positively definite, while for each $k = 7, 9$ and 11 there is exactly one indefinite form. The proof of Theorems 5.1 and 5.2 is effective in the following sense. It provides a procedure for representing any element of $H_1^+$ (or $\text{SU}(2, B, \mathbb{Z}[e])$) as a word in the generators $x, z$. For example, when $k = 7, 9$, or 11 let us choose the quaternion algebra $H$ which corresponds to the indefinite form above. Then a more detailed analysis of the proof in [88] shows that any $u = a_01 + a_1i + a_2j + a_3k \in H_1^+$ can be written as a word in $x$ and $z$ of length $O(\log(a_3^2 + a_2^2))$.

To conclude this section we indicate some relations with Macbeath’s theorem. For the sake of simplicity we deal only with Hurwitz groups, i.e., we assume $k = 7$. However, a similar observation can be applied when $k = 9$ or 11. It is well known that, for $k = 7$, the ring $\mathbb{Z}[\theta]$ is a principal ideal domain (see [66]; the tables from [66] are reproduced in [19, pp. 141–145]; also see [2, Table 7]). In addition, by a special case of a theorem due to E. Kummer (see, e.g., [28, §2.11, Corollary 1]) we have that for any rational prime $p$ the following holds:

- $p$ remains a prime in $\mathbb{Z}[\theta]$ if and only if $p \equiv \pm 2, \pm 3$ (mod 7);
- $p$ splits into a product of three different primes in $\mathbb{Z}[\theta]$ if and only if $p \equiv \pm 1$ (mod 7);
- $p$ ramifies in $\mathbb{Z}[\theta]$ if and only if $p = 7$.

In particular, if $p$ is a prime in $\mathbb{Z}[\theta]$ lying over a rational prime $p$, then $\mathbb{Z}[\theta]/p = \mathbb{F}(q)$, where $q = p$ for $p \equiv 0, \pm 1$ (mod 7) and $q = p^3$ otherwise.

If $p$ lies over an odd rational prime, then there is a natural residue homomorphism

$$\psi : H \rightarrow \left(\frac{-1, \theta^2 - 3}{\mathbb{Z}[\theta]/p}\right) \cong \text{Mat}_2(\mathbb{F}(q)).$$

Now, Theorem 4.9 combined with Theorem 5.2 asserts that the restriction of $\psi$ to $H_1^+$ is onto $\text{SL}_2(q)$. It is interesting to notice that we can also go in the opposite direction and deduce Macbeath’s theorem from the fact that the above restriction is onto. From this point of view it would be very interesting to find a purely number-theoretic proof of Macbeath’s theorem not using Dickson’s classification of subgroups of $\text{SL}_2(q)$.

We remark that results similar to the above were obtained, independently, in [71].

5.2. Other groups which are $(2, 3, 7)$-generated

The $(2, 3, 7)$-generation of $\text{SL}_n(q)$, for $n \geq 287$, was actually obtained as a special case of a more general result, which has many other applications.
Hurwitz groups and Hurwitz generation

Namely, given a ring $R$ with identity element, let $E_n(R)$ denote the group generated by the set of elementary matrices:

$$\{1 + re_{ij} \mid r \in R, 1 \leq i \neq j \leq n\}.$$

Clearly $E_n(R)$ can only be a finitely generated group if $R$ is a finitely generated ring. On the other hand, applications of the permutational methods of Higman and Conder in a linear context, as explained in Section 3.1, led to the following result.

**Theorem 5.3** (Lucchini, Tamburini and Wilson, [56]). Let $R$ be a ring which is generated by elements $\alpha_1, \ldots, \alpha_m$, where $2\alpha_1 - \alpha_1^2$ is a unit of $R$ of finite multiplicative order. Then $E_n(R)$ is $(2, 3, 7)$-generated for all $n \geq 287 + 84(m - 1)$.

For $n \geq 3$, the groups $E_n(R)$ and $SL_n(R)$ coincide, in particular, if $R$ is commutative and either semi-local or a Euclidean domain, see, e.g., [34, 1.2.11 and 4.3.9]. As the hypothesis on $R$ of the above theorem holds with $m = 1$ if $R$ is a finite field or the ring $\mathbb{Z}$ of integers, the same theorem implies that $SL_n(q)$ and $SL_n(\mathbb{Z})$ are $(2, 3, 7)$-generated for $n \geq 287$. This lower bound for $n$ was improved to 252 in [86].

An easy corollary of Theorem 5.3 also shows that the derived group of the automorphism group $Aut(F_n)$ of a free group of rank $n$ is $(2, 3, 7)$-generated, provided that $n \geq 329$.

But there are other applications of this theorem, which shed further light on the class of $(2, 3, 7)$-generated groups.

Let $p$ be a prime number. For each positive integer $l$, let $N_l$ be the kernel of the epimorphism $SL_n(\mathbb{Z}) \to SL_n(\mathbb{Z}/p^l\mathbb{Z})$. Thus, $N_1/N_l$ is a finite $p$-group for each $l$, and $\bigcap_{l \geq 1} N_l = 1$. Since $N_1$ (the group of matrices in $SL_n(\mathbb{Z})$ congruent to 1 modulo $p$) is non-soluble, we conclude that there is no bound on the derived lengths of the groups $N_1/N_l$.

Applying Theorem 5.3 we have the following result:

**Corollary 5.4** [56]. Let $n \geq 287$ and let $p$ be a prime. There exist Hurwitz groups which are extensions of $p$-groups of arbitrarily large derived length by the group $SL_n(p)$.

Since the direct product of two Hurwitz groups without common composition factors is again a Hurwitz group, one is led to study direct powers of simple groups. To this purpose, consider the polynomial ring $R = GF(q)[t_1, \ldots, t_l]$. Then $R$ can be generated by $l + 1$ elements, the first of which can be chosen to be a non-zero element $\alpha$ of $GF(q)$ satisfying $2\alpha - \alpha^2 \neq 0$. Let $I$ be the intersection of the kernels of the homomorphisms from $R$ to $GF(q)$ which extend the identity map on $GF(q)$. By the Chinese remainder theorem, $R/I$ is isomorphic to the direct product of $q^l$ copies of $GF(q)$, and the quotient map induces an epimorphism from $E_n(R)$ to the direct product of $q^l$ copies of $SL_n(q)$. We conclude from the theorem that this direct product is a Hurwitz group provided that $n \geq 287 + 84l$.

Therefore we have the following result, which shows that, for large $n$, the direct power of many copies of $SL_n(q)$ is a Hurwitz group.

**Corollary 5.5** [56]. Let $q$ be a prime power and let $n \geq 287$. Then the direct product of $r$ copies of $SL_n(q)$ is a Hurwitz group, where $r = q^{\lfloor (n-287)/84 \rfloor}$.
In [90], J. Wilson constructed a family of simple non-commutative, finitely generated rings $S$, with the property that, for each $n \geq 8$, the central quotient group $PE_n(S)$ of $E_n(S)$ is simple. Moreover, it was shown that the groups $PE_n(S)$ arising from rings $S$ in this family fall into $2^{\aleph_0}$ isomorphism classes. It follows from Theorem 5.3 that each of these groups is $(2, 3, 7)$-generated for $n$ sufficiently large. Therefore we have the following result, which makes unrealistic any attempt of classifying all $(2, 3, 7)$-generated groups.

**Corollary 5.6** [56]. There are $2^{\aleph_0}$ isomorphism classes of infinite simple $(2, 3, 7)$-generated groups.

Of course, there are no more than $2^{\aleph_0}$ isomorphism classes of finitely generated groups.

Another application of Theorem 5.3, recently made by M. Conder, gives a complete answer to the question of what centres are possible in finite quotients of the triangle group $T(2, 3, 7)$. This question was raised in 1965 by John Leech, [48], who later produced two infinite families of Hurwitz groups with centres of order 2 and 4. In [10], M. Conder used similar methods to prove the existence of infinitely many Hurwitz groups with a centre of order 3 and in [13] he constructed a family of central products of 2-dimensional special linear groups to show that the centre of a Hurwitz group could be an elementary Abelian 2-group of arbitrarily large order. Actually the centre of a Hurwitz group can be anything, in virtue of the following:

**Theorem 5.7** (Conder, [15]). Given any finite Abelian group $A$, there exist infinitely many Hurwitz groups $G$ such that the centre $Z(G)$ of $G$ is isomorphic to $A$.

We give a sketch of the elegant proof, which consists in taking a product of appropriately chosen special linear Hurwitz groups. Indeed, let $A$ be any finite Abelian group and write

$$A = C_{m_1} \times \cdots \times C_{m_s}$$

as a direct product of cyclic groups. Now choose any prime $p$ such that $(p, |A|) = 1$ and let $q = p^e$ with $e = \phi(|A|)$, where $\phi$ denotes Euler’s totient function. Then $q - 1 = \ell_i m_i$ for some integer $\ell_i$ (1 $\leq i \leq s$). Further, if $k_i$ is any positive integer coprime to $\ell_i$, then

$$(k_i m_i, q - 1) = (k_i m_i, \ell_i m_i) = m_i.$$

In particular there are infinitely many possibilities for each $k_i$, and all can be chosen such that $k_i m_i \neq k_j m_j$ for $i \neq j$ and $k_i m_i \geq 287$, for $1 \leq i \leq s$. Next let $H_i = \text{SL}_{k_i m_i}(q)$ and set

$$G = \prod_{1 \leq i \leq s} H_i.$$

Using Theorem 5.3, it is easily seen that $G$ has the required properties.
We mention another application of the same theorem, to the maximal parabolic subgroups of $E_{n+\bar{n}}(R)$. Namely, let:

$$P_{n,\bar{n}}(R) = \left\{ \begin{pmatrix} A & B \\ 0 & \bar{A} \end{pmatrix} \mid A \in E_n(R), \ \bar{A} \in E_{\bar{n}}(R), \ B \in \text{Mat}_{n,\bar{n}}(R) \right\}.$$ 

**THEOREM 5.8** (Di Martino and Tamburini, [24]). Let $R$ be a ring generated by elements $\alpha_1, \ldots, \alpha_m$, where $2\alpha_1 - \alpha_2^2$ is a unit of $R$ of finite multiplicative order. Then $P_{n,\bar{n}}(R)$ (and therefore also the Levi subgroup $L$) is $(2,3,7)$-generated for all $n, \bar{n} \geq 84(m+1) + 396$.

Finally, we mention that A. Lucchini in [54] and, independently, J.S. Wilson in [91] have generalized Theorem 5.3 to any $k \geq 7$. Lucchini’s version reads:

**THEOREM 5.9** (Lucchini, [54]). Let $R$ be a ring which is generated by elements $\alpha_1, \ldots, \alpha_m$, where $2\alpha_1 - \alpha_2^2$ is a unit of $R$ of finite multiplicative order. For any fixed $k \geq 7$ there exist two integers $n_k$ and $a_k$ such that $E_n(R)$ is $(2,3,k)$-generated for all $n \geq n_k + a_km$.

**Acknowledgements**

We are very grateful to A. Lucchini and A. Zalesskii for useful remarks and valuable advice on many parts of the manuscript.

**Appendix**

As mentioned in Section 3, to define a $(2,3,7)$-generating triple $(x, y, z)$ of $\text{Alt}(\Omega)$ when $n > 167$ and $n \neq 173, 174, 181, 188, 202$, Conder uses $3 + 14$ transitive permutational representations of $T(2,3,7)$, each of which is depicted by a diagram (see [8]). Here we describe explicitly each of these representations and give some information about the commutator $[x, y]$, useful to understand the considerations above theorem 3.4. We assume that the joins of diagrams, in order to obtain $[x, y]$, are made as described there: in particular each diagram of type $H_d$ is joined to a diagram of type $G$.

The following rule, which can easily be checked, has repeated application. Let $\psi$ and $\psi'$ be permutation representations of $T(2,3,7) = \langle X, Y \rangle$ on disjoint sets $\Delta$ and $\Delta'$, with respective handles $(e_1, e_2)$ and $(e'_1, e'_2)$. Let $e_3$ be the image of $e_2$ under $\psi(Y)$ and $e'_3$ be the image of $e'_2$ under $\psi'(Y)$. Consider the representation $\varphi$ on the set $\Delta \cup \Delta'$ defined by

$$\varphi(X) = \psi(X)\psi'(X)(e_1, e'_1)(e_2, e'_2), \quad \varphi(Y) = \psi(Y)\psi'(Y)$$

as described in Lemma 3.3. Assume, for simplicity, that $e_1$ and $e_3$ are in the same cycle of the commutator $[\psi(X), \psi'(Y)]$ and let $\Gamma$ be the support of this cycle. Let $\Gamma''$ be the union of the supports of the cycles of $[\psi'(X), \psi'(Y)]$ (not necessarily distinct) which contain $e'_1$ and $e'_3$. Then the set $\Gamma \cup \Gamma''$ is invariant under the commutator $[\varphi(X), \varphi(Y)]$. Moreover, setting
The cycle structure of the restriction \( [\varphi(X), \varphi(Y)]_{\Gamma \cup \Gamma'} \) consists of two cycles of length \( m \) if \( s = 0 \) and of one cycle of length \( 2m + 1 \) if \( s = 1 \). Furthermore, the cycle structure of \( [\varphi(X), \varphi(Y)] \) on \( \Delta \cup \Delta' \setminus \{ \Gamma \cup \Gamma' \} \) is the juxtaposition of the cycles of \( [\psi(X), \psi(Y)] \) and \( [\psi'(X), \psi'(Y)] \).

In view of the application to the linear context, our treatment differs slightly from that of Conder. In particular we avoid the use of the odd involution \( t \in \text{Sym}(\Omega) \), which centralizes \( x \) and inverts \( y \) by conjugation. But the two approaches are essentially equivalent in virtue of the obvious relation \((xyt)^2 = [x, y]^t\).

**NOTATION.** The cycle structure of a permutation will be denoted by

\[
(i, \ldots) \ell_1^{k_1} \ell_2^{k_2} \ldots.
\]

This means a cycle of length \( j \) whose support contains \( i \), followed by \( k_1 \) cycles of length \( \ell_1 \), followed by \( k_2 \) cycles of length \( \ell_2 \), etc.

**DIAGRAM G:** degree 42. Handles: (2, 3), (14, 15), (32, 33).

\[
x_G = (1, 4)(5, 7)(6, 10)(8, 12)(9, 24)(11, 29)(13, 16)(17, 19)(18, 25)(20, 27)
(14)(15)(32)(33).
\]

\[
y_G = \prod_{i=0}^{13} (3i + 1, 3i + 2, 3i + 3).
\]

\[
[x_G, y_G] = (2, \ldots, 1, \ldots) (14, \ldots, 13, \ldots) (32, \ldots, 31, \ldots) 1^3.
\]

**REMARK.** This representation is given by the action of \( \text{PSL}_2(13) \) on the cosets of a subgroup \( N \) of index 42.

**DIAGRAM A:** degree 14. Handle: (1, 2).

\[
x_A = (3, 4)(5, 9)(6, 11)(7, 10)(8, 13)(12, 14)(1)(2).
\]

\[
y_A = \prod_{i=0}^{3} (3i + 1, 3i + 2, 3i + 3)(13)(14).
\]

\[
[x_A, y_A] = (1, \ldots, 3, \ldots) 1^1.
\]
REMARK. This representation is given by the action of PSL$_2(13)$ on the points of the projective line.

**Diagram E:** degree 28. Handle: (1, 2).

\[ x_E = (3, 4)(5, 9)(6, 11)(7, 10)(8, 13)(12, 24)(14, 26)(15, 16)(18, 19)(21, 22) \\

\[ y_E = \prod_{i=0}^{8} (3i + 1, 3i + 2, 3i + 3)(28). \]

\[ [x_E, y_E] = (1, \ldots, 3, \ldots)^{1}9^{2}. \]

**Diagram H$_0$:** degree 42 $\equiv 0$ (mod 14). Handle: (1, 2).

\[ x_{H_0} = (3, 5)(4, 12)(6, 7)(8, 11)(9, 17)(10, 32)(13, 21)(14, 29)(15, 16)(18, 36) \\
(20)(26)(35)(41). \]

\[ y_{H_0} = \prod_{i=0}^{13} (3i + 1, 3i + 2, 3i + 3). \]

\[ [x_{H_0}, y_{H_0}] = (1, \ldots) (3, \ldots)^{1}3^{1}11^{1}17^{1}. \]

\[ c = [x, y]^{3}5^{1}11^{1}13^{2}3 \\
= (14, 31, 38, 20, 24, 40, 29, 19, 25, 15, 34, 22, 27, 39, 16, 30, 23). \]

**Diagram H$_1$:** degree 57 $\equiv 1$ (mod 14). Handle: (16, 17).

\[ x_{H_1} = (2, 9)(3, 5)(4, 35)(6, 12)(7, 41)(8, 11)(10, 13)(14, 15)(18, 20) \\
(1)(16)(17)(50)(56). \]

\[ y_{H_1} = \prod_{i=0}^{18} (3i + 1, 3i + 2, 3i + 3). \]
\[ [x_{H_1}, y_{H_1}] = (16, \ldots) (18, \ldots) 1^1 3^1 5^1 7^2 12^2. \]
\[
c = [x, y]^{1^3 3^7 11^3 13^2} = (1, 6, 9, 2, 12).
\]

**Diagram H₂:** degree 142 \(\equiv 2\) (mod 14). Handle: (77, 78).

\[
x_{H_2} = (1, 8)(2, 20)(3, 4)(5, 6)(7, 14)(9, 10)(11, 19)(12, 22)(13, 24)(15, 16)
(73, 74)(75, 81)(76, 83)(79, 80)(82, 84)(85, 87)(88, 100)(89, 98)(91, 96)
(118, 125)(121, 132)(123, 131)(124, 126)(127, 128)(129, 142)
\]

\[
y_{H_2} = \prod_{i=0}^{18} (3i + 1, 3i + 2, 3i + 3)(58) \prod_{i=20}^{47} (3i - 1, 3i, 3i + 1).
\]

\[
[x_{H_2}, y_{H_2}] = (77, 79, \ldots) 1^5 3^1 11^3 12^4 17^1 23^1.
\]
\[
c = [x, y]^{1^3 3^7 11^3 13^2} = (115, 125, 126, 121, 132, 124, 118, 116, 129, 136, 137, 131, 119, 123, 133,
138, 142).
\]

**Diagram H₃:** degree 115 \(\equiv 3\) (mod 14). Handle: (1, 2).

\[
x_{H_3} = (3, 4)(5, 9)(6, 11)(7, 10)(8, 14)(12, 24)(13, 17)(15, 26)(16, 20)(18, 38)
(70, 81)(71, 77)(73, 74)(75, 76)(78, 80)(82, 83)(86, 93)(87, 105)
\]

\[
y_{H_3} = \prod_{i=0}^{8} (3i + 1, 3i + 2, 3i + 3)(28) \prod_{i=10}^{38} (3i - 1, 3i, 3i + 1).
\]

\[
[x_{H_3}, y_{H_3}] = (1, 3, \ldots) 1^1 2^2 5^2 11^4 15^2 17^1.
\]
\[(x, y)^{2 \cdot 3 \cdot 5 \cdot 11 \cdot 13} = (13, 17, 30, 21, 27, 37, 40, 25, 16, 29, 20, 23, 42, 39, 28, 35, 32).\]

**Diagram H₄:** Degree 144 \(\equiv 4 \pmod{14}\). Handle: (64, 65).

\[x_{H₄} = (1, 10)(2, 7)(3, 4)(5, 76)(6, 77)(8, 13)(9, 11)(12, 28)(14, 37)(15, 17)
(51, 52)(54, 56)(57, 58)(59, 72)(61, 70)(63, 68)(66, 67)(69, 71)(73, 82)
(74, 85)(75, 80)(78, 79)(81, 83)(84, 102)(86, 103)(87, 89)(88, 95)
(90, 91)(92, 111)(93, 118)(94, 100)(96, 97)(98, 126)(99, 133)(101, 141)
(117, 119)(120, 121)(122, 142)(123, 124)(125, 131)(127, 128)

\[y_{H₄} = \prod_{i=0}^{47} (3i + 1, 3i + 2, 3i + 3).\]

\[[x_{H₄}, y_{H₄}] = (64, 66, \ldots) 1^{3} 5^{3} 8^{2} 11^{2} 17^{1} 30^{2}.\]

\[c = [x, y]^{8 \cdot 3 \cdot 5 \cdot 11 \cdot 13} = (108, 129, 115, 120, 128, 110, 135, 113, 125, 131, 112, 137, 116, 127, 121, 114, 130).\]

**Diagram H₅:** Degree 187 \(\equiv 5 \pmod{14}\). Handle: (1, 2).

\[x_{H₅} = (3, 4)(5, 11)(6, 21)(8, 13)(9, 17)(10, 14)(12, 20)(15, 75)(16, 19)(18, 69)
(66, 78)(68, 70)(71, 111)(72, 73)(74, 76)(77, 79)(80, 82)(83, 114)
(84, 85)(86, 95)(87, 88)(89, 99)(91, 92)(93, 94)(96, 97)(98, 100)
(116, 128)(119, 125)(120, 129)(121, 133)(122, 134)(123, 124)(126, 127)
(1)(2)(7).\]
\[ y_{H_6} = \prod_{i=0}^{51} (3i + 1, 3i + 2, 3i + 3)(157) \prod_{i=53}^{62} (3i - 1, 3i, 3i + 1). \]

\[ [x_{H_6}, y_{H_6}] = (1, 3, \ldots) \cdot 1^34^29^210^212^14^15^24^3. \]

\[ c = [x, y]^{4^9\cdot 5^{11\cdot 13}} = (7, 100, 62, 58, 41, 18, 94, 76, 52, 107, 71, 92, 70, 50, 72, 113, 95, 66, 53, 13, 102, 82, 80, 40, 8, 44, 78, 86, 115, 73, 49, 68, 91, 111, 109, 51, 74, 93, 69, 103, 57, 64, 98). \]

**Diagram H_6:** degree 216 \equiv 6 \pmod{14}. Handle: (70, 71).

\[ x_{H_6} = (1, 9)(2, 5)(4, 32)(6, 11)(7, 26)(8, 10)(12, 13)(14, 15)(16, 40)(17, 41) \]
\[ \hspace{0.5cm} (18, 19)(20, 23)(21, 36)(22, 25)(24, 35)(27, 28)(29, 31)(33, 34)(37, 46) \]
\[ \hspace{1.5cm} (38, 49)(39, 44)(42, 43)(45, 47)(48, 66)(50, 68)(51, 52)(53, 59)(54, 61) \]
\[ \hspace{2.5cm} (55, 170)(56, 58)(57, 169)(60, 62)(63, 64)(65, 78)(67, 76)(69, 74)(72, 73) \]
\[ \hspace{3.5cm} (72, 73)(75, 77)(79, 83)(80, 121)(81, 92)(82, 88)(84, 85)(86, 87)(89, 91) \]
\[ \hspace{4.5cm} (90, 162)(93, 94)(95, 97)(96, 114)(98, 127)(99, 100)(101, 116)(102, 103) \]
\[ \hspace{5.5cm} (104, 211)(105, 106)(107, 208)(108, 109)(110, 115)(111, 112)(113, 150) \]
\[ \hspace{6.5cm} (117, 118)(119, 141)(120, 142)(122, 182)(123, 124)(125, 134)(126, 128) \]
\[ \hspace{7.5cm} (129, 136)(130, 131)(132, 133)(135, 137)(138, 139)(140, 163)(143, 180) \]
\[ \hspace{8.5cm} (144, 145)(146, 155)(147, 148)(149, 159)(151, 152)(153, 154)(156, 157) \]
\[ \hspace{9.5cm} (158, 160)(161, 185)(164, 181)(165, 166)(167, 173)(168, 175)(171, 172) \]
\[ \hspace{10.5cm} (174, 176)(177, 178)(179, 186)(183, 184)(187, 194)(188, 206)(189, 190) \]
\[ \hspace{11.5cm} (191, 192)(193, 200)(195, 196)(197, 205)(198, 214)(199, 216)(201, 202) \]
\[ \hspace{12.5cm} (203, 204)(207, 209)(210, 212)(213, 215)(3, 30)(70)(71). \]

\[ y_{H_6} = \prod_{i=0}^{71} (3i + 1, 3i + 2, 3i + 3). \]

\[ [x_{H_6}, y_{H_6}] = (70, \ldots, 72, \ldots) \cdot 1^54^29^16^27^211^112^13^217^232^2. \]

\[ c = [x, y]^{32^3\cdot 7^1\cdot 11^1\cdot 13^1} = (1, 3, 9, 11, 6). \]

**Diagram H_7:** degree 77 \equiv 7 \pmod{14}. Handle: (50, 51).

\[ x_{H_7} = (1, 22)(2, 23)(3, 4)(5, 8)(6, 21)(7, 11)(9, 20)(10, 13)(14, 16)(18, 19) \]
\[ \hspace{0.5cm} (24, 25)(26, 30)(27, 32)(28, 31)(29, 35)(33, 45)(34, 37)(36, 47)(38, 67) \]
\[ \hspace{1.5cm} (39, 40)(41, 69)(42, 43)(44, 46)(48, 49)(52, 53)(54, 58)(55, 60)(56, 59) \]
\[ \hspace{2.5cm} (57, 63)(61, 73)(62, 66)(64, 75)(65, 68)(70, 71)(72, 74)(76, 77)(12, 15) \]
\[ \hspace{3.5cm} (17)(50)(51). \]
Hurwitz groups and Hurwitz generation

\[ y_{H_7} = \prod_{i=0}^{15} (3i + 1, 3i + 2, 3i + 3) (49) \prod_{i=17}^{25} (3i - 1, 3i, 3i + 1) (77). \]

\[ [x_{H_7}, y_{H_7}] = (50, \ldots, 52, \ldots) 1^3 2^2 4^2 9^4 17^1. \]

\[ c = [x, y]^{4-9-7-11-13} = (1, 22, 24, 3, 26, 5, 33, 19, 36, 12, 47, 18, 45, 8, 30, 4, 25). \]

**Diagram H_8:** degree 36 \(\equiv 8 \pmod{14} \). Handle: (16, 17).

\[ x_{H_8} = (1, 9)(2, 5)(4, 32)(6, 11)(7, 27)(8, 10)(12, 13)(14, 15)(18, 19)(20, 23)

\[ y_{H_8} = \prod_{i=0}^{11} (3i + 1, 3i + 2, 3i + 3). \]

\[ [x_{H_8}, y_{H_8}] = (16, \ldots, 18, \ldots) 1^1 4^2 5^1 11^1. \]

\[ c = [x, y]^{4-3-11-13} = (1, 11, 3, 6, 9). \]

**Diagram H_9:** degree 135 \(\equiv 9 \pmod{14} \). Handle: (124, 125).

\[ x_{H_9} = (3, 4)(5, 8)(6, 21)(7, 10)(9, 20)(11, 48)(12, 13)(14, 16)(17, 42)(18, 19)
(56, 58)(59, 61)(64, 73)(65, 70)(66, 67)(71, 76)(72, 74)(75, 93)(77, 100)
(78, 79)(80, 95)(81, 83)(82, 88)(84, 85)(86, 87)(89, 94)(90, 91)(92, 103)

\[ y_{H_9} = \prod_{i=0}^{44} (3i + 1, 3i + 2, 3i + 3). \]

\[ [x_{H_9}, y_{H_9}] = (124, 126, \ldots) 1^4 3^1 4^2 5^2 8^2 11^2 19^1 21^2. \]

\[ c = [x, y]^{8-3-5-7-11-13} = (1, 11, 31, 29, 18, 3, 40, 32, 52, 8, 5, 50, 27, 38, 4, 19, 34, 28, 48). \]
**Diagram $H_{10}$**: degree $136 \equiv 10 \pmod{14}$. Handle: $(55, 56)$.

\[
x_{H_{10}} = (1, 9)(2, 5)(4, 32)(6, 11)(7, 26)(8, 10)(12, 13)(14, 15)(16, 37)(17, 38)
(57, 58)(60, 62)(63, 64)(65, 78)(67, 79)(68, 80)(69, 70)(71, 75)(72, 77)
(73, 76)(81, 82)(83, 87)(84, 89)(85, 88)(86, 92)(90, 102)(91, 94)
(93, 104)(95, 131)(96, 97)(98, 128)(99, 100)(101, 103)(105, 106)
\]

\[
y_{H_{10}} = \prod_{i=0}^{34}(3i + 1, 3i + 2, 3i + 3)(106) \prod_{i=36}^{45}(3i - 1, 3i, 3i + 1).
\]

\[
[x_{H_{10}}, y_{H_{10}}] = (55, \ldots, 57, \ldots) 15^{42} 5^{1} 11^{3} 12^{6}.
\]

\[
c = [x, y]^{4 \cdot 3 \cdot 11 \cdot 13} = (1, 11, 3, 6, 9).
\]

**Diagram $H_{11}$**: degree $165 \equiv 11 \pmod{14}$. Handle: $(160, 161)$.

\[
x_{H_{11}} = (1, 26)(2, 27)(3, 4)(5, 8)(6, 21)(7, 11)(9, 20)(10, 14)(15, 16)(18, 19)
(57, 76)(59, 63)(60, 64)(62, 81)(65, 83)(66, 67)(68, 74)(70, 72)(71, 73)
(75, 77)(78, 79)(80, 93)(82, 91)(84, 89)(85, 98)(86, 99)(87, 88)(90, 92)
(94, 103)(95, 100)(96, 97)(101, 112)(102, 104)(105, 123)(106, 118)
(119, 124)(120, 121)(122, 133)(126, 127)(128, 141)(129, 148)
(142, 143)(144, 145)(147, 149)(150, 151)(152, 159)(154, 157)
\]

\[
y_{H_{11}} = \prod_{i=0}^{54}(3i + 1, 3i + 2, 3i + 3).
\]

\[
[x_{H_{11}}, y_{H_{11}}] = (160, 162, \ldots) 15^{2} 2^{4} 5^{4} 8^{4} 11^{6} 19.
\]

\[
c = [x, y]^{8 \cdot 3 \cdot 5 \cdot 11 \cdot 13}
\]

\[
= (1, 53, 3, 4, 39, 26, 8, 49, 24, 18, 29, 31, 12, 22, 34, 19, 25, 48, 5).
\]
**Diagram** $H_{12}$: degree $180 \equiv 12 \pmod{14}$. Handle: $(5, 6)$.


$y_{H_{12}} = \prod_{i=0}^{59}(3i + 1, 3i + 2, 3i + 3)$.

$[x_{H_{12}}, y_{H_{12}}] = (4, 5, 6, \ldots) \cdot 1^{3} \cdot 2^{5} \cdot 3^{2} \cdot 7^{2} \cdot 11^{2} \cdot 13^{2} \cdot 19^{1} \cdot 47^{1}$.

$c = [x, y]^{8 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 19}$

$= (36, 155, 113, 66, 157, 101, 137, 67, 138, 179, 55, 124, 111, 163, 171, 100, 142, 63, 174, 178, 45, 120, 104, 166, 167, 96, 117, 59, 175, 158, 72, 133, 97, 161, 165, 109, 153, 46, 177, 140, 68, 141, 99, 172, 64, 112, 123)$.  

**Diagram** $H_{13}$: degree $195 \equiv 13 \pmod{14}$. Handle: $(180, 178)$.


$y_{H_{13}} = \prod_{i=0}^{64}(3i + 1, 3i + 2, 3i + 3)$.  

**Hurwitz groups and Hurwitz generation**

423
\[ [x_{H_{13}}, y_{H_{13}}] = (179, \ldots, 180, \ldots) 1^1 2^2 5^2 6^2 7^2 13^4 14^2 23^1. \]

\[ c = [x, y]^{1^2 3^2 5^2 7^2 11^2 13^2} = (2, 19, 37, 32, 42, 24, 9, 18, 7, 27, 39, 31, 41, 29, 11, 28, 16, 26, 44, 34, 22, 17, 20). \]

References


Hurwitz groups and Hurwitz generation

[57] C. Maclachlan, private correspondence.


# Braids, their Properties and Generalizations

V.V. Vershchinin*

Département des Sciences Mathématiques, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier cedex 5, France
E-mail: vershini@math.univ-montp2.fr

Sobolev Institute of Mathematics, Novosibirsk, 630090, Russia
E-mail: versh@math.nsc.ru

## Contents

1. Introduction ........................................ 429
2. Historical remarks .................................... 429
3. Definitions and general properties ........................ 431
   3.1. Systems of \( n \) curves in three-dimensional space and braid groups .................. 431
   3.2. Braid groups and configuration spaces ...................................................... 433
   3.3. Braid groups as automorphism groups of free groups and the word problem ............ 434
   3.4. Commutator subgroup and other presentations ............................................. 435
   3.5. Presentation of the pure braid group and Markov normal form .......................... 437
4. Garside normal form, center and conjugacy problem .......... 438
5. Ordering of braids .................................... 441
6. Representations ........................................ 442
   6.1. Burau representation .................................................. 442
   6.2. Lawrence–Krammer representation ......................................................... 442
7. Generalizations of braids .................................. 443
   7.1. Configuration spaces of manifolds ......................................................... 443
   7.2. Artin–Brieskorn braid groups ............................................................... 444
   7.3. Braid groups of surfaces ................................................................. 446
   7.4. Braid groups in handlebodies ............................................................... 446
   7.5. Braids with singularities ................................................................. 448
8. Homological properties ..................................... 452
   8.1. Configuration spaces and \( K(\pi, 1) \)-spaces ............................................. 452
   8.2. Cohomology of pure braid groups ......................................................... 453
   8.3. Homology of braid groups ............................................................... 455

*The author was supported in part by the by CNRS-NSF grant No 17149, INTAS grant No 03-5-3251 and the ACI project ACI-NIM-2004-243 “Braids and Knots”.

HANDBOOK OF ALGEBRA, VOL. 4
Edited by M. Hazewinkel
© 2006 Elsevier B.V. All rights reserved
Abstract

In the chapter we give a survey on braid groups and subjects connected with them. We start with the initial definition, then we give several interpretations as well as several presentations of these groups. Burau presentation for the pure braid group and the Markov normal form are given next. Garside normal form and his solution of the conjugacy problem are presented as well as more recent results on the ordering and on the linearity of braid groups. Next topics are the generalizations of braids, their homological properties and connections with the other mathematical fields, like knot theory (via Alexander and Markov theorems) and homotopy groups of spheres.
1. Introduction

Braid groups describe intuitive concept of classes of continuous deformations of braids, which are collections of intertwining strands whose endpoints are fixed. Mathematically, they can be considered from various points of view. The first intuitive approach is formalized naturally as isotopy classes of a collection of \( n \) connected curves (strings) in 3-dimensional space. This point of view is connected with the definition of a braid group as the fundamental group of a configuration space of \( n \) points on a plane. Also braids can be interpreted as a mapping class group of a punctured disc and as a subgroup of the automorphism group of a free group (Section 3.3).

The present survey is organized as follows. In Section 2 we make some historical remarks. Definition and general properties are considered in Section 3. Configuration spaces appear in Section 3.2. Connections with groups of automorphisms of free groups are given in Section 3.3. Presentations of the braid group which appeared quite recently are observed in Section 3.4. Section 4 is devoted to F.A. Garside’s classical work, [91], and Section 5 to that of P. Dehornoy on ordering for braids. Representations and in particular linearity are discussed in Section 6. In Section 7 various generalization of braids are presented. Homological properties are observed in Section 8. In the last Section 9 we discuss connection with the knot theory given by the Alexander and Markov theorems and with the homotopy groups of spheres.

2. Historical remarks

Braids were rigorously defined by E. Artin, [7], in 1925, although the roots of this natural concept are seen in the works of A. Hurwitz ([117], 1891), R. Fricke and F. Klein ([89], 1897) and even in the notebooks of C.-F. Gauss. E. Artin, [7], gave the presentation of the braid group (see formulas (3.2) in Section 3) which is common now. Already in the book of Felix Klein, [126], published in 1926 there appeared a chapter about braids. Essential topics about braids were also presented in the Reidemeister’s Knotentheorie, [175], published in 1932.

In the 30ies there appeared a series of papers of Werner Burau, [48–50], where he in particular gave the presentation of the pure braid group (see Section 3.5) and introduced the Burau representation (Section 6.1). Wilhelm Magnus in his work [138] published in 1934 established relations between braid groups and the mapping class groups. At the same time there appeared the work of A.A. Markov, [149], which together with the Alexander theorem, [3], builds a bijection between links and equivalence classes of braids. It became an essential ingredient in study of links and knots (in the work of V.F.R. Jones, [121], for example). In 1936–37 were published papers of O. Zariski, [200,201], where he discovered connections between braid groups and the fundamental group of the complement of the discriminant of the general polynomial

\[
f_n(t) = a_0 t^n + a_1 t^{n-1} + \cdots + a_{n-1} t + a_n,
\]

a point of view later rediscovered by V.I. Arnold, [5]. Zariski also understood connections between braids and configuration spaces, gave the presentation of the braid group of the
spheres, and studied the braid groups of Riemann surfaces. Amazingly and unfortunately these works of Zariski were not noticed by the specialists on braids and are not mentioned even in books and papers where the presentations of braids of surfaces are discussed.

In the beginning of 60ies R. Fox and L. Neuwirth, [86], and E. Fadell and L. Neuwirth, [80], studied configuration spaces which turned out to be $K(\pi, 1)$-spaces and so give a natural geometrical model of the classifying spaces for the braid groups. Later, V.I. Arnold, [5], in this direction proved the first results on the cohomology of braids. The motivation for his study was a connection (which he discovered) with the problem of representing algebraic functions in several variables by superposing algebraic functions in fewer variables. Also, in 1969 V.I. Arnold completely described the cohomology of pure braid groups, [4].

In 1969 there appeared the publication of F.A. Garside’s work [91] where he suggests a new normal form of elements in the braid group and with its help gives a new solution of the word problem and also solves the conjugacy problem. In 1968 was published a two-page note of G.S. Makanin, [142], where he sketches his algorithm for the solution of the conjugacy problem. The complete publication of Makanin’s work didn’t appear (as far as the author is aware).

In the 70ies the study of cohomology of braids was continued independently and by different methods by D.B. Fuks, [90], who determined the mod 2 cohomology, and F.R. Cohen, [54–56], who described the homology with coefficients in $\mathbb{Z}$ and in $\mathbb{Z}/p$ as modules over the Steenrod algebra.

In 1984–85, independently, N.V. Ivanov, [118], and J. McCarthy, [151], proved the “Tits alternative” for the mapping class groups of surfaces and as a consequence it is true for the braid groups. Namely, they proved that every subgroup of the mapping class group either contains an Abelian subgroup of finite index, or contains a non-Abelian free group.

The question of whether braid groups are linear attracted significant attention. It was realized that the Burau representation is faithful for $Br_3$, [92,141]. Then, after a long break, in 1991 J.A. Moody, [153], proved that Burau representation is unfaithful for $n \geq 9$. This bound was improved to $n \geq 6$ by D.D. Long and M. Paton, [137], and to $n = 5$ by S. Bigelow, [28]. In 1999–2000 there appeared preprints of papers of D. Krammer, [128,129], and S. Bigelow, [29], who proved that $Br_n$ is linear for all $n$ (using the other representation).

At the beginning of nineties P. Dehornoy, [68–70], proved that there exists a left order in braid groups.

Interesting generalizations of braids were introduced in the work of E. Brieskorn, [42]. The configuration space can also be considered as the orbit space of the complement of the complexification of the arrangement of hyperplanes corresponding to the Coxeter group $A_{n-1} = \Sigma_n$. Generalizing this approach to any finite Coxeter group, E. Brieskorn defined the so-called generalized braid groups which are also called Artin groups.

Another way of generalization is to consider braid groups in 3-manifolds, possibly with a boundary. The simplest examples are braid groups in handlebodies. A.B. Sossinsky, [182], was the first who studied them. Such a group can be interpreted as the fundamental group of the configuration space of a plane without $g$ points where $g$ is the genus of the handlebody. The generalized braid group of type $C$ is isomorphic to the braid group in the solid torus.
In the context of the influence of the theory of Vassiliev–Goussarov (finite-type) invariants singular braids were introduced. The corresponding algebraic structures are the Baez–Birman monoid, [10,33], and the braid-permutation group by R. Fenn, R. Rimányi and C. Rourke, [83,84]. Various properties of these objects were studied in [85,202,93,94,66,119,62,104].

3. Definitions and general properties

3.1. Systems of $n$ curves in three-dimensional space and braid groups

First of all, as was already mentioned, braids naturally arise as objects in 3-space. Let us consider two parallel planes $P_0$ and $P_1$ in $\mathbb{R}^3$, which contain two ordered sets of points $A_1, \ldots, A_n \in P_0$ and $B_1, \ldots, B_n \in P_1$. These points are lying on parallel lines $L_A$ and $L_B$ respectively. The space between the planes $P_0$ and $P_1$ we denote by $\Pi$. Suppose that the point $B_i$ is lying under the point $A_i$, as a result of the orthogonal projection of the plane $P_0$ onto the plane $P_1$. Let us connect the set of points $A_1, \ldots, A_n$ with the set of points $B_1, \ldots, B_n$ by simple nonintersecting curves $C_1, \ldots, C_n$ lying in the space $\Pi$ and such that each curve meets only once each parallel plane $P_t$ lying in the space $\Pi$ (see Figure 1). This object is called a braid and the curves are called the strings of a braid. Usually braids are depicted by projections on the plane passing through the lines $L_A$ and $L_B$. This projection is supposed to be in general position so that there is only finite number of double points of intersection which are lying on pairwise different levels and intersections are transversal. The simplest braid $\sigma_i$ (Figure 2) corresponds to the transposition $(i, i + 1)$.

Let us introduce the following equivalence relation on the set of all braids with $n$ strings and with fixed $P_0$, $P_1$, $A_i$ and $B_i$. It is defined by homeomorphisms $h : \Pi \to \Pi$, which are the identity on $P_0 \cup P_1$ and such that $h(P_t) = P_t$. Braids $\beta$ and $\beta'$ are equivalent if there exists a homeomorphism $h$ such that $h(\beta) = \beta'$. On the set $Br_n$ of equivalence classes under the considered relation the structure of a group is introduced as follows. We put a copy $\Pi'$ of the domain $\Pi$ under the $\Pi$ in such a way that $P_0'$ coincides with $P_1$ and each $A_i$ coincides with $B_i$ and we glue the braids $\beta$ and $\beta'$. This gluing gives a composition of braids $\beta \beta'$ (Figure 3). The unit element is the equivalence class containing a braid of $n$ parallel intervals, the braid $\beta^{-1}$ inverse to $\beta$ is defined by reflection of $\beta$ with respect to

![Fig. 1.](image-url)
the plane \( P_{1/2} \). A string \( C_i \) of a braid \( \beta \) connects the point \( A_i \) with the point \( B_{ki} \), defining a permutation \( S^\beta \). If this permutation is the identity then the braid \( \beta \) is called \textit{pure}. The map \( \beta \rightarrow S^\beta \) defines an epimorphism \( \tau_n \) of the braid group \( Br_n \) on the permutation group \( \Sigma_n \) with the kernel consisting of all pure braids:

\[
1 \rightarrow P_n \rightarrow Br_n \xrightarrow{\tau_n} \Sigma_n \rightarrow 1.
\] (3.1)

The following presentation of the braid group \( Br_n \) with generators \( \sigma_i \), \( i = 1, \ldots, n - 1 \), and two types of relations:

\[
\begin{align*}
\sigma_i \sigma_j &= \sigma_j \sigma_i, & \text{if } |i - j| > 1, \\
\sigma_i \sigma_{i+1} \sigma_i &= \sigma_{i+1} \sigma_i \sigma_{i+1}
\end{align*}
\] (3.2)

is the algebraic expression of the fact that any isotopy of braids can be broken down into “elementary moves” of two types that correspond to the two types of relations.
If we add a vertical interval to the system of curves on Figure 1 we can get a canonical inclusion $j_n$ of the group $Br_n$ into the group $Br_{n+1}$

$$j_n : Br_n \to Br_{n+1}.$$  

If the symmetric group $\Sigma_n$ is given by its canonical presentation with generators $s_i$, $i = 1, \ldots, n - 1$, and relations:

$$\begin{align*}
  s_i s_j &= s_j s_i, \quad \text{if } |i - j| > 1, \\
  s_i s_{i+1} s_i &= s_{i+1} s_i s_{i+1}, \\
  s_i^2 &= 1, \\
\end{align*}$$

(3.3)

then the homomorphism $\tau_n$ is given by the formula

$$\tau_n(\sigma_i) = s_i, \quad i = 1, \ldots, n - 1.$$

It is possible to consider braids as classes of equivalence of braid diagrams which are generic projections of three-dimensional braids on a plane. The classes of equivalence are defined by the Reidemeister moves depicted in Figure 4.

### 3.2. Braid groups and configuration spaces

If we look at Figure 1, then this picture can be interpreted as a graph of a loop in the configuration space of $n$ points on a plane, that is the space of unordered sets of $n$ points on a plane, see Figure 5. So, it is possible to interpret the braid group as the fundamental
group of the configuration space. Formally it is done as follows. The symmetric group $\Sigma_m$ acts on the Cartesian power $(\mathbb{R}^2)^m$ of the space $\mathbb{R}^2$:

$$w(y_1, \ldots, y_m) = (y_{w^{-1}(1)}, \ldots, y_{w^{-1}(m)}), \quad w \in \Sigma_m.$$ (3.4)

Denote by $F(\mathbb{R}^2, m)$ the space of $m$-tuples of pairwise different points in $\mathbb{R}^2$:

$$F(\mathbb{R}^2, m) = \{(p_1, \ldots, p_m) \in (\mathbb{R}^2)^m : p_i \neq p_j \text{ for } i \neq j\}.$$ This is the space of regular points of our action. We call the orbit space of this action $B(\mathbb{R}^2, m) = F(\mathbb{R}^2, m)/\Sigma_m$ the configuration space of $n$ points on a plane. The braid group $Br_m$ is the fundamental group of this configuration space

$$Br_m = \pi_1(B(\mathbb{R}^2, m)).$$

The pure braid group $P_m$ is the fundamental group of the space $F(\mathbb{R}^2, m)$. The covering $p : F(\mathbb{R}^2, m) \to B(\mathbb{R}^2, m)$ defines an exact sequence:

$$1 \to \pi_1(F(\mathbb{R}^2, m)) \xrightarrow{p_*} \pi_1(B(\mathbb{R}^2, m)) \to \Sigma_n \to 1,$$ (3.5)

which is equivalent to the sequence (3.1).

It can be used for proving the canonical presentation of the braid group (3.2) as is done, for example, in the book of J. Birman, [32].

Such considerations were done by R. Fox and L. Neuwirth, [86].

3.3. Braid groups as automorphism groups of free groups and the word problem

Another important approach to the braid group is based on the fact that this group may be considered as a subgroup of the automorphism group of a free group.

Let $F_n$ be the free group of rank $n$ with the set of generators $\{x_1, \ldots, x_n\}$. Denote by $\text{Aut } F_n$ the automorphism group of $F_n$.

We have the standard inclusions of the symmetric group $\Sigma_n$ and the braid group $Br_n$ into $\text{Aut } F_n$. For the braid group it may be described as follows. Let $\tilde{\sigma}_i \in \text{Aut } F_n, i = 1, 2, \ldots, n - 1$, be given by the following formula, which describes its action on the generators:

$$\begin{align*}
\tilde{x}_i &\mapsto x_{i+1}, \\
\tilde{x}_{i+1} &\mapsto x_i^{-1}x_{i+1}x_i, \\
\tilde{x}_j &\mapsto x_j, \quad j \neq i, i + 1.
\end{align*}$$ (3.6)

Let us define a map $\nu$ of the generators $\sigma_i, i = 1, \ldots, n - 1$, of the braid group $Br_n$ to these automorphisms:

$$\nu(\sigma_i) = \tilde{\sigma}_i.$$ (3.7)
THEOREM 3.1. Formulas (3.7) define correctly a homomorphism

\[ \nu : \text{Br}_n \to \text{Aut } F_n \]

which is a monomorphism.

Theorem 3.1 gives a solution of the word problem for the braid groups. This was done first by E. Artin, [7].

The free group \( F_n \) is a fundamental group of a disc \( D_n \) without \( n \) points and the generator \( x_i \) corresponds to a loop going around the \( i \)-th point. The braid group \( \text{Br}_n \) is the mapping class group of a disc \( D_n \) with its boundary fixed, [32], and so it acts on the fundamental group of \( D_n \). This action is described by the formulas (3.6) where \( x_i \) corresponds to the canonical loops on \( D_n \) which form the generators of the fundamental group. Geometrically this action is depicted in Figure 6.

3.4. Commutator subgroup and other presentations

Let us define a homomorphism from the braid group to the integers by taking the sum of exponents of the entries of the generators \( \sigma_i \) in the expression of any element of the group through these canonical generators:

\[ \deg : \text{Br}_n \to \mathbb{Z}, \quad \deg(b) = \sum_j m_j, \text{ where } b = (\sigma_i_1)^{m_1} \cdots (\sigma_i_k)^{m_k}. \]

PROPOSITION 3.1. The homomorphism

\[ \deg : \text{Br}_n \to \mathbb{Z} \]

gives the Abelianization of the braid group and the commutator subgroup \( \text{Br}'_n \) is characterized by the condition

\[ b \in \text{Br}'_n \text{ if and only if } \deg(b) = 0. \]
Proof. Let \( a : Br_n \rightarrow A \) be a homomorphism to any other Abelian group \( A \), then from the relations (3.2) we have:

\[
a(\sigma_i)a(\sigma_{i+1})a(\sigma_i) = a(\sigma_{i+1})a(\sigma_i)a(\sigma_{i+1}).
\]

The commutativity of \( A \) gives that \( a(\sigma_{i+1}) = a(\sigma_i) \). This means that the homomorphism \( \text{deg} \) is universal. \( \square \)

Of course, there exist another presentations of the braid group. Let

\[
\sigma = \sigma_1 \sigma_2 \cdots \sigma_{n-1},
\]

then the group \( Br_n \) is generated by \( \sigma_1 \) and \( \sigma \) because

\[
\sigma_{i+1} = \sigma_i^{1} \sigma_1 \sigma^{-i}, \quad i = 1, \ldots, n-2.
\]

The relations for the generators \( \sigma_1 \) and \( \sigma \) are the following

\[
\begin{align*}
\sigma_1 \sigma_1' \sigma_1^{-i} &= \sigma_1' \sigma_1 \sigma^{-i} \sigma_1, \quad \text{for } 2 \leq i \leq n/2, \\
\sigma^n &= (\sigma \sigma_1)^{n-1}.
\end{align*}
\]

(3.8)

This was observed by Artin in the initial paper [7].

An interesting series of presentations was given by V. Sergiescu, [181]. For every planar graph he constructed a presentation of the group \( Br_n \), where \( n \) is the number of vertices of the graph, with generators corresponding to edges and relations reflecting the geometry of the graph. Artin’s presentation in this context corresponds to the graph consisting of the interval from 1 to \( n \) with the natural numbers (from 1 to \( n \)) as vertices and with segments between them as edges. For generalizations of braids graph presentations of these type were considered by P. Bellingeri and V. Vershinin, [17,21].

J.S. Birman, K.H. Ko and S.J. Lee, [35], introduced the presentation with the generators \( a_{ts} \) with \( 1 \leq s < t \leq n \) and relations

\[
\begin{align*}
a_{ts}a_{rq} &= a_{rq}a_{ts}, \quad \text{for } (t-r)(t-q)(s-r)(s-q) > 0, \\
a_{ts}a_{sr} &= a_{sr}a_{ts} = a_{sr}a_{tr}, \quad \text{for } 1 \leq r < s < t \leq n.
\end{align*}
\]

(3.9)

The generators \( a_{ts} \) are expressed by the canonical generators \( \sigma_i \) in the following form:

\[
a_{ts} = (\sigma_{t-1} \sigma_{t-2} \cdots \sigma_{s+1}) \sigma_i (\sigma_{s+1}^{-1} \cdots \sigma_{t-2}^{-1} \sigma_{t-1}^{-1}) \quad \text{for } 1 \leq s < t \leq n.
\]

(3.10)

Geometrically the generators \( a_{s,t} \) are depicted in Figure 7.

The set of generators for braid groups were even enlarged in the work of Jean Michel, [152], as follows. Let \( | : \Sigma_n \rightarrow \mathbb{Z} \) be the length function on the symmetric group with respect to the generators \( s_i \): for \( x \in \Sigma_n \), \( |x| \) is the smallest natural number \( k \) such that \( x \) is a product of \( k \) elements of the set \( \{s_1, \ldots, s_{n-1}\} \). It is known ([41], Section 1, Example 13(b)) that two minimal expressions for an element of \( \Sigma_n \) are equivalent by using only the relations (3.2). This implies that the canonical projection \( \tau_n : Br_n \rightarrow \Sigma_n \) has
a unique set-theoretic section $r : \Sigma_n \to B_{r_n}$ such that $r(s_i) = \sigma_i$ for $i = 1, \ldots, n - 1$ and $r(xy) = r(x)r(y)$ whenever $|xy| = |x| + |y|$. Then the group $B_{r_n}$ admits a presentation by generators $\{r(x) \mid x \in \Sigma_n\}$ and relations $r(xy) = r(x)r(y)$ for all $x, y \in \Sigma_n$ such that $|xy| = |x| + |y|$.

### 3.5. Presentation of the pure braid group and Markov normal form

Let $f(y_1, \ldots, y_m)$ be a word with (possibly empty) entries of $y^\epsilon_i$, where the $y_i$ are some letters and $\epsilon$ may be $\pm 1$. If $y_i$ are elements of a group $G$ then $f(y_1, \ldots, y_m)$ will be considered as the corresponding element of $G$.

Let us define the elements $s_{i,j}$, $1 \leq i < j \leq m$, of the braid group $B_{r_m}$ by the formula:

$$s_{i,j} = \sigma_{j-1} \cdots \sigma_{i+1} \sigma_i^2 \sigma_{i+1}^2 \cdots \sigma_j^{-1}.$$ 

These elements satisfy the following Burau relations ([48,150], see also [134]):

$$\begin{align*}
s_{i,j}s_{k,l} &= s_{k,l}s_{i,j} & \text{for } i < j < k < l, \\
s_{i,j}s_{i,k}s_{j,k} &= s_{i,k}s_{j,k}s_{i,j} & \text{for } i < j < k, \\
s_{i,k}s_{j,k} &= s_{j,k}s_{i,k} & \text{for } i < j < k, \\
s_{i,k}s_{j,k}s_{j,k}^{-1}s_{i,k} &= s_{j,k}s_{j,k}^{-1}s_{i,k} & \text{for } i < j < k.
\end{align*}$$

W. Burau and later A.A. Markov proved that the elements $s_{i,j}$ with the relations (3.11) give a presentation of the pure braid group $P_m$, [150]. The following formula is a consequence of the Burau relations and is also due to A.A. Markov:

$$[s_{i,j}, s_{j,k}^\epsilon] = f(s_{i,j}, \ldots, s_{l-1,j}), \quad \epsilon = \pm 1, \quad k < l.$$ 

Let us define the elements $\sigma_{k,l}$, $1 \leq k \leq l \leq m$, by the formulas

$$\sigma_{k,k} = e,$$

$$\sigma_{k,l} = \sigma_{k}^{-1} \cdots \sigma_{j}^{-1}.$$ 

Let $P_m^k$ be the subgroup of $P_m$ generated by the elements $s_{i,j}$ with $k < j$. 

---

**Fig. 7.**
THEOREM 3.2 (A.A. Markov).

(i) Every element of the group $\text{Br}_m$ can be uniquely written in the form

$$f_m(s_{1,m}, \ldots, s_{m-1,m}) \cdots f_j(s_{1,j}, \ldots, s_{j-1,j}) \cdots f_2(s_{1,2}) \sigma_{i_m,m} \cdots \sigma_{i_j,j} \cdots \sigma_{i_2,2}.$$  \hspace{1cm} (3.13)

(ii) The factor group $P^k_m/P^k_{m+1}$ is the free group on free generators $s_{i,k+1}, 1 \leq i \leq k$.

The form (3.13) is called the Markov normal form, it also gives the solution of the word problem for the braid groups.

4. Garside normal form, center and conjugacy problem

An essential role in Garside’s work [91] is played by the monoid of positive braids $\text{Br}_n^+$, that is the monoid which has a presentation with generators $\sigma_i, i = 1, \ldots, n$, and relations (3.2). In other words each element of this monoid can be represented as a word on the elements $\sigma_i, i = 1, \ldots, n$, with no entrances of the $\sigma_i^{-1}$. Two positive words $A$ and $B$ in the alphabet $\{\sigma_i, i = 1, \ldots, n - 1\}$ will be said to be positively equal if they are equal as elements of $\text{Br}_n^+$. In this case we shall write $A \equiv B$.

First of all Garside proves the following statement.

PROPOSITION 4.1. In $\text{Br}_n^+$ for $i, k = 1, \ldots, n - 1$, given $\sigma_i A \equiv \sigma_k B$, it follows that

- if $k = i$, then $A \equiv B$,
- if $|k - i| = 1$, then $A \equiv \sigma_k \sigma_i Z, B \equiv \sigma_i \sigma_k Z$ for some $Z$,
- if $|k - i| \geq 2$, then $A \equiv \sigma_k Z, B \equiv \sigma_i Z$ for some $Z$.

The same is true for the right multiples of $\sigma_i$.

COROLLARY 4.1. If $A \equiv P, B \equiv Q, AXB \equiv PYQ, (L(A) \geq 0, L(B) \geq 0)$, then $X \equiv Y$. That is, monoid $\text{Br}_n^+$ is left and right cancellative.

Garside’s fundamental word $\Delta$ in the braid group $\text{Br}_{n+1}$ is defined by the formula:

$$\Delta = \sigma_1 \cdots \sigma_n \sigma_1 \cdots \sigma_{n-1} \cdots \sigma_1 \sigma_2 \sigma_1.$$ 

If we use Garside’s notation $\Pi_i \equiv \sigma_1 \cdots \sigma_i$, then $\Delta \equiv \Pi_{n-1} \cdots \Pi_1$.

For a positive word $W$ in $\sigma_i, i = 1, \ldots, n$, we say that $\Delta$ is a factor of $W$ or simply $W$ contains $\Delta$, if $W \equiv A \Delta B$ with $A$ and $B$ being arbitrary positive words, probably empty. If $W$ does not contain $\Delta$ we shall say $W$ is prime to $\Delta$.

Garside’s transformation of words $R$ is defined by the formula

$$R(\sigma_i) \equiv \sigma_{n-i}.$$ 

This gives the automorphism of $\text{Br}_n$ and the positive braid monoid $\text{Br}_n^+$. 
PROPOSITION 4.2. In $B_r$ 

$$\sigma_i \Delta \doteq \Delta R(\sigma_i).$$

Geometrically this commutation is shown on Figure 8 ($\Delta \sigma_3 = \sigma_1 \Delta$).

PROPOSITION 4.3. If $W$ is an arbitrary positive word in $B_r^+$ such that either

$$W \doteq \sigma_1 A_1 \doteq \sigma_2 A_2 \doteq \cdots \doteq \sigma_{n-1} A_{n-1},$$

or

$$W \doteq B_1 \sigma_1 \doteq B_2 \sigma_2 \doteq \cdots \doteq B_{n-1} \sigma_{n-1},$$

then $W \doteq \Delta Z$ for some $Z$.

PROPOSITION 4.4. The canonical homomorphism

$$B_r^+ \to B_r$$

is a monomorphism.

Among positive words on the alphabet $\{\sigma_1 \cdots \sigma_n\}$ let us introduce a lexicographical ordering with the condition that $\sigma_1 < \sigma_2 < \cdots < \sigma_n$. For a positive word $W$ the base of $W$ is
the smallest positive word which is positively equal to $W$. The base is uniquely determined. If a positive word $A$ is prime to $\Delta$, then for the base of $A$ the notation $\bar{A}$ will be used.

**Theorem 4.1 (F.A. Garside).** Every word $W$ in $Br_{n+1}$ can be uniquely written in the form $\Delta^m \bar{A}$, where $m$ is an integer.

The form of a word $W$ established in this theorem we call the Garside left normal form and the index $m$ we call the power of $W$. The same way the Garside right normal form is defined and the corresponding variant of Theorem 4.1 is true. The Garside normal form also gives a solution to the word problem in the braid group.

**Theorem 4.2 (F.A. Garside).** The necessary and sufficient condition that two words in $Br_{n+1}$ are equal is that their Garside normal forms are identical.

Garside normal form for the braid groups was precised in the subsequent works of S.I. Adyan, [1], W. Thurston, [78], E. El-Rifai and H.R. Morton, [76]. Namely, there was introduced the left-greedy form (in the terminology of W. Thurston, [78])

$$\Delta^t A_1 \cdots A_k,$$

where $A_i$ are the successive possible longest fragments of the word $\Delta$ (in the terminology of S.I. Adyan, [1]) or positive permutation braids (in the terminology of E. El-Rifai and H.R. Morton, [76]). Certainly, the same way the right-greedy form is defined. With the help of this form it was proved that the braid group is biautomatic.

The center of the braid group was first found by W.-L. Chow, [53]. Namely, as follows from the presentation of braid groups with two generators $\sigma_1$ and $\sigma_2$ and relations (3.8) given in Section 3.1 the element $\sigma^n$ commutes with $\sigma_2$ and so with $\sigma_1$. Chow proved that it generates the center. Garside normal form gives an elegant proof of the following theorem.

**Theorem 4.3.**

(i) When $n = 1$, the center of the group $Br_{n+1}$ is generated by $\Delta$.

(ii) When $n > 1$ the center of the group $Br_{n+1}$ is generated by $\Delta^2$.

Let $\alpha$ be a positive word such that $\Delta = \alpha X$, where $X$ is a positive word, possibly empty. For any word $W$ in $Br_{n+1}$, the word $\alpha^{-1} W \alpha$, reduced to Garside normal form is called an $\alpha$-transformation of $W$.

For any word $W$ in $Br_{n+1}$ with the Garside normal form $\Delta^m \bar{A} \equiv W_1$ consider the following chains of $\alpha$-transformations: take all the $\alpha$-transformations of $W_1$ and let those which are of power $\geq m$ and which are distinct from each other be $W_2, W_3, \ldots, W_t$. Now repeat the process for each of the words $W_2, W_3, \ldots, W_t$ in turn, denoting successively by $W_{t+1}, W_{t+2}, \ldots$ any new words occurring, the condition being always that each new word must be of power $\geq m$. Continue to repeat the process for every new distinct word arising, as the sequence $W_1, W_2, W_{t+2}, \ldots$, expands.
PROPOSITION 4.5. The set $W_1, W_2, W_{t+2}, \ldots$, is finite.

Suppose that in the set $W_1, W_2, W_{t+2}, \ldots$, the highest power reached is $s$ and that the words of power $s$ form the subset $V_1, V_2, \ldots$. Then this set $V_1, V_2, \ldots$ is called the summit set of $W$.

THEOREM 4.4 (F.A. Garside). Two elements $A$ and $B$ of the group $Br_{n+1}$ are conjugate if and only if their summit sets are identical.

J.S. Birman, K.H. Ko, and S.J. Lee considered the word $\delta = a_n(a_{n-1}) \cdots a_2a_1 = \sigma_{n-1} \cdots \sigma_2\sigma_1$, as a fundamental in their system of generators and proved that every element in $Br_n$ has a representative $W = \delta^j A_1 A_2 \cdots A_k$ with positive $A_i$ in a unique way in some sense. Based on this form they gave an algorithm for the word problem in $B_n$ which runs in time $(nm^2)$ for a given word of length $m$.

5. Ordering of braids

A group $G$ is said totally (or linearly) left (correspondingly right) ordered if it has a total order $<$ invariant by left (right) multiplication, i.e. if $a < b$, then $ca < cb$ for any $c \in G$. If this order is also invariant by right (left) multiplication, then the group $G$ is called ordered.

For any left ordered group $G$ denote by $P$ the set of positive elements $\{x \in G : x > 1\}$, then the set of negative elements is defined by the formula: $P^{-1} = \{x \in G : x \in P\}$. The total character of an order on $G$ is expressed by the partition

$$G = P \sqcup \{1\} \sqcup P^{-1}.$$  

The invariance of multiplication is expressed by the inclusion $P^2 \subset P$, where $P^2$ is formed by products of couples of elements of $P$. Conversely, if there exists a subset $P$ of a group $G$ with the properties:

$$G = P \sqcup \{1\} \sqcup P^{-1}, \quad P^2 \subset P,$$

then $G$ is left ordered by the order defined by: $x < y$ if and only if $x^{-1}y \in P$. A group $G$ then is ordered if and only if $xPx^{-1} \subset P$ for all $x \in G$.

Let $i \in \{1, \ldots, n\}$ and considered a word $w$ on the alphabet $\{\sigma_1, \ldots, \sigma_n\}$ expressed in the form

$$w_0\sigma_i w_1 \sigma_i \cdots \sigma_i w_r,$$

where the subwords $w_0, \ldots, w_r$ are the words on the letters $\sigma_j^{\pm 1}$ with $j > i$. Then such a word is called $\sigma_i$-positive. This means that all entries of $\sigma_j^{\pm 1}$ in the word $w$ with $i$ minimal must be positive. If all such entries are negative then a word $w$ is called $\sigma_i$-negative. A braid of $Br_{n+1}$ is called $\sigma_i$-positive ($\sigma_i$-negative) if there exists its expression as a word on the standard generators which is $\sigma_i$-positive ($\sigma_i$-negative). A braid is called $\sigma$-positive ($\sigma$-negative) if there exists a number $i$, such that it is $\sigma_i$-positive ($\sigma_i$-negative).
THEOREM 5.1 (P. Dehornoy). Every braid in $Br_{n+1}$ different from 1 is either $\sigma$-positive or $\sigma$-negative.

COROLLARY 5.1. For all $n$ the braid group $Br_{n+1}$ is left ordered.

6. Representations

6.1. Burau representation

Let us map the generators of the braid group $Br_n$ to the following elements of the group $GL_n \mathbb{Z}[t, t^{-1}]

$\sigma_i \mapsto \begin{pmatrix} E_{i-1} & 0 & 0 & 0 \\ 0 & 1-t & t & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & E_{n-i-1} \end{pmatrix}$, \hspace{1cm} (6.1)

where $E_i$ is the unit $i \times i$ matrix. The formula (6.1) gives a well-defined representation of the braid group in $GL_n \mathbb{Z}[t, t^{-1}]:$

$r : Br_n \rightarrow GL_n \mathbb{Z}[t, t^{-1}]$, 

which is called Burau representation, [50].

THEOREM 6.1. Burau representation is faithful for $n = 3$.

THEOREM 6.2 (J.A. Moody, D.D. Long and M. Paton, S. Bigelow). The Burau representation is not faithful for $n \geq 5$.

The case $n = 4$ remains open.

6.2. Lawrence–Krammer representation

Consider the ring $K = \mathbb{Z}[q^{\pm 1}, t^{\pm 1}]$ of Laurent polynomials in two variables $q, t$, and the free $K$-module

$V = \bigoplus_{1 \leq i < j \leq n} K x_{i,j}$.

For $k \in \{1, 2, \ldots, n - 1\}$, define the action of the braid generators $\sigma_k$ on the specified basis of $V$ by the formula:
Braids, their properties and generalizations

\[ \sigma_k(x_{i,j}) = \begin{cases} 
  x_{i,j}, & k < i - 1 \text{ or } j < k; \\
  x_{i-1,j} + (1-q)x_{i,j}, & k = i - 1; \\
  tq(q - 1)x_{i,i+1} + qx_{i+1,j}, & k = i < j - 1; \\
  tq^2 x_{i,j}, & k = i = j - 1; \\
  x_{i,j} + tq^{k-i} (q - 1)^2 x_{k,k+1}, & i < k < j - 1; \\
  x_{i,j-1} + tq^{j-i} (q - 1)x_{j-1,j}, & k = j - 1; \\
  (1-q) x_{i,j} + qx_{i,j+1}, & k = j. 
\]  

(6.2)

Direct computation shows that this defines a representation

\[ \rho_n : Br_n \to GL(V), \]

which was firstly defined by R. Lawrence, [133], in topological terms and in the explicit form (6.2) by D. Krammer, [129].

**Theorem 6.3 (S. Bigelow, [29], D. Krammer, [129]).** The representation

\[ \rho_n : Br_n \to GL(V) \]

is faithful for all \( n \geq 1 \).

**Remark 6.1.** Actually, S. Bigelow, [29], proved this theorem for the representation \( \rho_n \) characterized in homological terms and D. Krammer, [129], proved the following. Let \( K = \mathbb{R}[t^\pm 1], q \in \mathbb{R}, \) and \( 0 < q < 1 \). Then the representation \( \rho_n \) defined by (6.2) is faithful for all \( n \geq 1 \). This result implies Theorem 6.3: if a representation over \( \mathbb{Z}[q^{\pm 1}, t^{\pm 1}] \) becomes faithful after assigning a real value to \( q \), then it is faithful itself.

M.G. Zinno, [203], established a connection between the Birman–Murakami–Wenzl algebra, [40,155], and the Lawrence–Krammer representation. Namely, he proved that the Lawrence–Krammer representation is identical to the irreducible representations of the Birman–Murakami–Wenzl algebra parametrized by Young diagrams of shapes \((n - 2)\) and \((1^{n-2})\). This means that the Young diagram in the case considered consists of one row (respectively of one column) only, with \( n - 2 \) boxes. It follows that Lawrence–Krammer representation is irreducible.

**7. Generalizations of braids**

**7.1. Configuration spaces of manifolds**

The notion of a configuration space as in Section 3.2 can be naturally generalized for a configuration space of a manifold as follows. Let \( Y \) be a connected topological manifold and let \( W \) be a finite group acting on \( Y \). A point \( y \in Y \) is called regular if its stabilizer \( \{ w \in W : wy = y \} \) is trivial, i.e. consists only of the unit of the group \( W \). The set \( \tilde{Y} \)
of all regular points is open. Suppose that it is connected and nonempty. The subspace \( \text{ORB}(Y, W) \) of the space of all orbits \( \text{Orb}(Y, W) \) consisting of the orbits of all regular points is called the \textit{space of regular orbits}. There is a free action of \( W \) on \( \tilde{Y} \) and the projection \( p: \tilde{Y} \to \tilde{Y}/W = \text{ORB}(Y, W) \) defines a covering. Let us consider the initial segment of the long exact sequence of this covering:

\[
1 \to \pi_1(\tilde{Y}, y_0) \xrightarrow{p_*} \pi_1(\text{ORB}(Y, W), p(y_0)) \to W \to 1.
\]

The fundamental group \( \pi_1(\text{ORB}(Y, W), p(y_0)) \) of the space of regular orbits is called the \textit{braid group of the action of} \( W \text{ on } Y \) and is denoted by \( \text{Br}(Y, W) \). The fundamental group \( \pi_1(\tilde{Y}, y_0) \) is called the \textit{pure braid group of the action of} \( W \text{ on } Y \) and is denoted by \( \text{P}(Y, W) \).

The spaces \( \tilde{Y} \) and \( \text{ORB}(Y, W) \) are path connected, so the pair of these groups is defined uniquely up to isomorphism and we may omit mentioning the base point \( y_0 \) in the notations.

For any space \( Y \) the symmetric group \( \Sigma_m \) acts on the Cartesian power \( Y^m \) of the space \( Y \) by the formulas (3.4). We denote by \( F(Y, m) \) the space of \( m \)-tuples of pairwise different points in \( Y \):

\[
F(Y, m) = \{ (p_1, \ldots, p_m) \in Y^m : p_i \neq p_j \text{ for } i \neq j \}.
\]

This is the space of regular points of this action. In the case when \( Y \) is a connected topological manifold \( M \) without boundary and \( \dim M \geq 2 \), the space of regular orbits \( \text{ORB}(M^m, \Sigma_m) \) is open, connected and nonempty. We call \( \text{ORB}(M^m, \Sigma_m) \) the \textit{configuration space of the manifold} \( M \) and denote by \( \text{B}(M, m) \). The braid group \( \text{Br}(M^m, \Sigma_m) \) is called the \textit{braid group on} \( m \text{ strings of the manifold} \) \( M \) and is denoted by \( \text{Br}(m, M) \).

Analogously, we call the group \( \text{P}(M^m, \Sigma_m) \) the \textit{pure braid group on} \( m \text{ strings of the manifold} \) \( M \) and denote it by \( \text{P}(m, M) \). These definitions of braid groups were given by R. Fox and L. Neuwirth, [86].

\section{Artin–Brieskorn braid groups}

The braid groups are included in the series of so-called generalized braid groups (this was their name in the work of E. Brieskorn of 1971, [42]), or Artin groups (as they were called by E. Brieskorn and K. Saito in the paper of 1972, [45]). They were defined by E. Brieskorn, [42], so we call them Artin–Brieskorn groups.

Let \( V \) be a finite-dimensional real vector space \( (\dim V = n) \) with Euclidean structure. Let \( W \) be a finite subgroup of \( \text{GL}(V) \) generated by reflections. Let \( \mathcal{M} \) be the set of hyperplanes such that \( W \) is generated by orthogonal reflections with respect to \( M \in \mathcal{M} \). We suppose that for every \( w \in W \) and every hyperplane \( M \in \mathcal{M} \) the hyperplane \( w(M) \) belongs to \( \mathcal{M} \).

The group \( W \) is generated by the reflections \( w_i = w_i(M_i), \ i \in I, \) satisfying only the following relations

\[
(w_i w_j)^{m_{i,j}} = e, \quad i, j \in I,
\]
where the natural numbers \( m_{i,j} = m_{j,i} \) form the Coxeter matrix of \( W \) from which the Coxeter graph \( \Gamma(W) \) of \( W \) is constructed, [41]. We use the following notation of P. Deligne, [74]: \( \prod(m; x, y) \) denotes the product \( x y x y \cdots (m \text{ factors}) \). The generalized braid group (or Artin–Brieskorn group) \( Br(W) \) of \( W \), [42,74], is defined as the group with generators \( \{ s_i, i \in I \} \) and relations:

\[
\prod(m_{i,j}; s_i, s_j) = \prod(m_{j,i}; s_j, s_i).
\]

From this we obtain the presentation of the group \( W \) by adding the relations:

\[
s_i^2 = e; \quad i \in I.
\]

We will see in Theorem 7.1 that this definition of the generalized braid group agrees with our general definition of a braid group of an action of a group \( W \) (Section 7.1). We denote by \( \tau_W \) the canonical homomorphism from \( Br(W) \) to \( W \). The classical braids on \( k \) strings \( Br_k \) are obtained by this construction if \( W \) is the symmetric group on \( k + 1 \) symbols.

The classification of irreducible (with connected Coxeter graph) Coxeter groups is well known (see, for example, Theorem 1, Chapter VI, §4 of [41]). It consists of the three infinite series: \( A, C \) (which is also denoted by \( B \) because in the corresponding classification of simple Lie algebras two different series \( B \) and \( C \) have this group as their Weyl group) and \( D \) as well as the exceptional groups \( E_6, E_7, E_8, F_4, G_2, H_3, H_4 \) and \( I_2(p) \).

Now let us consider the complexification \( V_C \) of the space \( V \) and the complexification \( M_C \) of \( M \in M \). Let \( Y_W = V_C - \bigcup_{M \in M} M_C \). The group \( W \) acts freely on \( Y_W \). Let \( X_W = Y_W / W \) then \( Y_W \) is a covering over \( X_W \) corresponding to the group \( W \). Let \( y_0 \in A_0 \) be a point in some chamber \( A_0 \) and let \( x_0 \) stand for its image in \( X_W \). We are in the situation described in Section 7.1 in the definition of the braid group of the action of the group \( W \). This braid group is defined as the fundamental group of the space of regular orbits of the action of \( W \). In our case \( ORB(V_C, W) = X_W \). So, the generalized braid group is \( \pi_1(X_W, x_0) \). For each \( j \in I \), let \( \ell_j \) be the homotopy class of paths in \( Y_W \) starting from \( y_0 \) and ending in \( w_j(y_0) \) which contains a polygon line with successive vertices: \( y_0, y_0 + iy_0, w_j(y_0) + iy_0, w_j(y_0) \). The image \( \ell_j \) of the class \( \ell_j \) in \( X_W \) is a loop with base point \( x_0 \).

**Theorem 7.1.** The fundamental group \( \pi_1(X_W, x_0) \) is generated by the elements \( \ell_j \) satisfying the following relations:

\[
\prod(m_{j,k}; \ell_j, \ell_k) = \prod(m_{k,j}; \ell_k, \ell_j).
\]

This theorem was proved by E. Brieskorn, [43].

The word problem and the conjugacy problem for Artin–Brieskorn groups were solved by E. Brieskorn and K. Saito, [45], and P. Deligne, [74]. The biautomatic structure of these groups was established by R. Charney, [51].

In the case when \( V \) is complex finite-dimensional space and \( W \) is a finite subgroup of \( GL(V) \) generated by pseudo-reflections the corresponding braid groups were studied by M. Broué, G. Malle and R. Rouquier, [47], and also by D. Bessis and J. Michel, [27].
7.3. Braid groups of surfaces

The braid groups of a sphere $Br_n(S^2)$ also have simple geometric interpretation as a group of isotopy classes of braids lying in a layer between two concentric spheres. It has the presentation with generators $\delta_i, i = 1, \ldots, n - 1$, and relations:

$$
\begin{align*}
\delta_i \delta_j &= \delta_j \delta_i, \quad \text{if } |i - j| > 1, \\
\delta_i \delta_{i+1} \delta_i &= \delta_{i+1} \delta_i \delta_{i+1}, \\
\delta_1 \delta_2 \cdots \delta_{n-2} \delta_{n-1} \delta_{n-2} \cdots \delta_2 \delta_1 &= 1.
\end{align*}
$$

(7.2)

This presentation was found by O. Zariski, [200], in 1936 and then rediscovered by E. Fadell and J. Van Buskirk, [81], in 1961.

Presentations of braid groups of all closed surfaces were obtained by G.P. Scott, [180], and others.

7.4. Braid groups in handlebodies

The subgroup $Br_{1,n+1}$ of the braid group $Br_{n+1}$ consisting of braids with the first string fixed can be interpreted also as the braid group in a solid torus. Here we study braids in a handlebody of the arbitrary genus $g$.

Let $H_g$ be a handlebody of genus $g$. The braid group $Br_g^n$ on $n$ strings in $H_g$ was first considered by A.B. Sossinsky, [182]. Let $Q_g$ denote a subset of the complex plane $\mathbb{C}$, consisting of $g$ different points, $Q_g = \{z_0^1, \ldots, z_0^g\}$, say, $z_0^j = j$. The interior of the handlebody $H_g$ may be interpreted as the direct product of the complex plane $\mathbb{C}$ without $g$ points: $\mathbb{C} \setminus Q_g$, and an open interval, for example, $(-1, 1)$:

$$
\hat{H}_g = (\mathbb{C} \setminus Q_g) \times (-1, 1).
$$

The space $F(\mathbb{C} \setminus Q_g, n)$ can be interpreted as the complement of the arrangement of hyperplanes in $\mathbb{C}^{g+n}$ given by the formulas:

$$
H_{jk}: z_j - z_k = 0 \quad \text{for all } j, k; \\
H_{ji}^j: z_j = z_i^0 \quad \text{for } i = 1, \ldots, g; \quad j = 1, \ldots, n.
$$

The braids in $Br_g^n$ are considered as lying between the planes with coordinates $z = 0$ and $z = 1$ and connecting the points $(g + 1, 0), \ldots, (g + n, 0))$. So $Br_g^n$ can be considered as a subgroup of the classical braid group $Br_{g+n}$ on $g + n$ strings such that the braids from $Br_g^n$ leave the first $g$ strings unbraided. In this subsection we denote by $\sigma_j$ the standard generators of the group $Br_{g+n}$. Let $\tau_k, k = 1, 2, \ldots, g$, be the following braids:

$$
\tau_k = \sigma_{g+1} \sigma_{g+2} \cdots \sigma_{k+1} \sigma_k^{-1} \sigma_{k+1}^{-1} \cdots \sigma_{g+1}^{-1} \sigma_g^{-1}.
$$

Such a braid is depicted in Figure 9. The elements $\tau_k, k = 1, 2, \ldots, g$, generate a free subgroup $F_g$ in the braid group $Br_{g+n}$. It follows for example from the Markov normal form that the elements $\tau_k, k = 1, 2, \ldots, g$, together with the standard generators
\[
\sigma_i \sigma_j = \sigma_j \sigma_i \quad \text{if } |i - j| > 1,
\]
\[
\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1},
\]
\[
\tau_k \sigma_i = \sigma_i \tau_k \quad \text{if } k \geq 1, \ i \geq 2,
\]
\[
\tau_k \sigma_1 \tau_k \sigma_1 = \sigma_1 \tau_k \sigma_1 \tau_k, \quad k = 1, 2, \ldots, g,
\]
\[
\tau_k \sigma_1^{-1} \tau_{k+1} \sigma_1 = \sigma_1^{-1} \tau_{k+1} \sigma_1 \tau_k, \quad k = 1, 2, \ldots, g - 1; \ l = 1, 2, \ldots, g - k.
\]

The relation of the fourth type in (7.3) is the relation of the braid group of type \(B(C)\). The relations of the fifth type in (7.3) describe the interaction between the generators of the free group and their closest neighbor \(\sigma_1\). Geometrically this is seen in Figure 10. If we introduce new generators \(\theta_k, \ k = 1, 2, \ldots, g - 1\); by the formulas:

\[
\theta_k = \sigma_1^{-1} \tau_k \sigma_1
\]

we obtain the "positive" presentation of the group \(B_n^g\) with generators of the types \(\sigma_i, \tau_k, \theta_k\) and relations:

\[
\sigma_i \sigma_j = \sigma_j \sigma_i \quad \text{if } |i - j| > 1,
\]
\[
\sigma_i \sigma_{i+1} \sigma_i = \sigma_{i+1} \sigma_i \sigma_{i+1},
\]
\[
\tau_k \sigma_i = \sigma_i \tau_k \quad \text{if } k \geq 1, \ i \geq 2,
\]
\[
\tau_k \sigma_1 \tau_k \sigma_1 = \sigma_1 \tau_k \sigma_1 \tau_k, \quad k = 1, 2, \ldots, g,
\]
\[
\tau_k \theta_{k+l} = \theta_{k+l} \tau_k, \quad k = 1, 2, \ldots, g - 1; \ l = 1, 2, \ldots, g - k,
\]
\[
\sigma_1 \theta_k = \tau_k \sigma_1, \quad k = 1, 2, \ldots, g - 1.
\]

There is an analog of Markov Theorem 3.2 for the group \(B_n^g\), [189,192].
7.5. Braids with singularities

Let $BP_n$ be the subgroup of $\text{Aut} F_n$, generated by both sets of the automorphisms $\sigma_i$ of (3.6) and $\xi_i$ of the following form:

$$
\begin{align*}
X_i & \mapsto x_{i+1}, \\
X_{i+1} & \mapsto x_i, \\
x_j & \mapsto x_j, \quad j \neq i, i+1.
\end{align*}
$$

This is the braid-permutation group. R. Fenn, R. Rimányi and C. Rourke proved, [83,84], that this group is given by the set of generators: $\{\xi_i, \sigma_i, \ i = 1, 2, \ldots, n - 1\}$ and relations:

$$
\begin{align*}
\xi_i^2 &= 1, \\
\xi_i \xi_j &= \xi_j \xi_i, \quad \text{if } |i - j| > 1, \\
\xi_i \xi_{i+1} \xi_i &= \xi_{i+1} \xi_i \xi_{i+1}.
\end{align*}
$$

The symmetric group relations

$$
\begin{align*}
\sigma_i \sigma_j &= \sigma_j \sigma_i, \quad \text{if } |i - j| > 1, \\
\sigma_i \sigma_{i+1} \sigma_i &= \sigma_{i+1} \sigma_i \sigma_{i+1}.
\end{align*}
$$

The braid group relations
\[
\begin{align*}
\sigma_i \xi_j &= \xi_j \sigma_i, \quad \text{if } |i - j| > 1, \\
\xi_i \xi_{i+1} \sigma_i &= \sigma_{i+1} \xi_{i} \xi_{i+1}, \\
\sigma_i \sigma_{i+1} \xi_i &= \xi_{i+1} \sigma_i \sigma_{i+1}.
\end{align*}
\]

The mixed relations

R. Fenn, R. Rimányi and C. Rourke also gave a geometric interpretation of \(BP_n\) as a group of welded braids. First they define a welded braid diagram on \(n\) strings as a collection of \(n\) monotone arcs starting from \(n\) points at a horizontal line of a plane (the top of the diagram) and going down to \(n\) points at another horizontal line (the bottom of the diagram). The diagrams may have crossings of two types: (1) the same as usual braids as for example on Figure 2 or (2) welds as depicted in Figure 11.

Composition of welded braid diagrams on \(n\) strings is defined by stacking one diagram under the other. The diagram with no crossings or welds is the identity with respect to composition. So the set of welded braid diagrams on \(n\) strings forms a semigroup which is denoted by \(WD_n\).

R. Fenn, R. Rimányi and C. Rourke defined the allowable moves on welded braid diagrams. They consist of the usual Reidemeister moves (Figure 4) and the specific moves depicted in Figures 12, 13, 14. The automorphisms of \(F_n\) which lie in \(BP_n\) can be characterized as follows. Let \(\pi \in \Sigma_n\) be a permutation and \(w_i, i = 1, 2, \ldots, n\), be words in \(F_n\). Then the mapping

\[x_i \mapsto w_i^{-1} x_{\pi(i)} w_i\]

determines an injective endomorphism of \(F_n\). If it is also surjective, we call it an automorphism of permutation-conjugacy type. The automorphisms of this type comprise a subgroup of \(\text{Aut } F_n\) which is precisely \(BP_n\).
The Baez–Birman monoid $SB_n$ or singular braid monoid, [10,33], is defined as the monoid with generators $g_i, g_i^{-1}, a_i, i = 1, \ldots, n - 1$, and relations

\[
\begin{align*}
g_i g_j &= g_j g_i \quad \text{if } |i - j| > 1, \\
a_i a_j &= a_j a_i \quad \text{if } |i - j| > 1, \\
a_i g_j &= g_j a_i \quad \text{if } |i - j| \neq 1, \\
g_i^{g_i+1} g_i &= g_i^{g_i+1} g_i, \\
g_i g_i^{a_i+1} g_i &= g_i g_i^{a_i+1} g_i, \\
g_i g_i^{a_i+1} a_i &= a_i g_i^{a_i+1} g_i, \\
g_i g_i^{a_i+1} g_i &= g_i^{-1} g_i = 1.
\end{align*}
\]

In these pictures $g_i$ corresponds to canonical generator of the braid group and $a_i$ represents an intersection of the $i$-th and $(i + 1)$-st strand as in Figure 15. A more detailed geometric interpretation of the Baez–Birman monoid can be found in the article of J. Birman, [33]. R. Fenn, E. Keyman and C. Rourke proved, [82], that the Baez–Birman monoid embeds in a group $SG_n$ which they called the singular braid group:

$$SB_n \to SG_n.$$ 

So, in $SG_n$ the elements $a_i$ become invertible and all relations of $SB_n$ remain true.
The analogue of the Birman–Ko–Lee presentation for the singular braid monoid was obtained in [198]. Namely, it was proved that the monoid $SB_n$ has a presentation with generators $a_{is}, a_{is}^{-1}$ for $1 \leq s < t \leq n$ and $b_{qp}$ for $1 \leq p < q \leq n$ and relations

\[
\begin{align*}
\{a_{is}a_{rq} &= a_{rq}a_{is} \quad \text{for } (t-r)(t-q)(s-r)(s-q)>0, \\
a_{is}a_{sr} &= a_{ts}a_{is} = a_{sr}a_{tr} \quad \text{for } 1 \leq r < s < t \leq n, \\
a_{ts}a_{is}^{-1} &= a_{is}^{-1}a_{ts} = 1 \quad \text{for } 1 \leq s < t \leq n, \\
a_{is}b_{rq} &= b_{rq}a_{is} \quad \text{for } (t-r)(t-q)(s-r)(s-q)>0, \\
a_{is}b_{ts} &= b_{ts}a_{is} \quad \text{for } 1 \leq s < t \leq n, \\
a_{ts}b_{sr} &= b_{sr}a_{ts} \quad \text{for } 1 \leq r < s < t \leq n, \\
a_{ts}b_{ts} &= b_{ts}a_{ts} \quad \text{for } 1 \leq r < s < t \leq n, \\
b_{ts}b_{rq} &= b_{rq}b_{ts} \quad \text{for } (t-r)(t-q)(s-r)(s-q)>0.
\end{align*}
\]

(7.5)

The elements $a_{is}$ are defined the same way as in (3.10) and the elements $b_{qp}$ for $1 \leq p < q \leq n$ are defined by

\[
b_{qp} = (\sigma_{q-1}\sigma_{q-2}\cdots\sigma_{p+1})x_p(\sigma_{p+1}^{-1}\cdots\sigma_{q-2}^{-1}\sigma_{q-1}^{-1}) \quad \text{for } 1 \leq p < q \leq n.
\]

(7.6)

Geometrically the generators $b_{s,t}$ are depicted in Figure 16.
8. Homological properties

8.1. Configuration spaces and $K(\pi, 1)$-spaces

Let $(q_i)_{i \in \mathbb{N}}$ be a fixed sequence of distinct points in the manifold $M$ and put $Q_m = \{q_1, \ldots, q_m\}$. We use

$$Q_{m,l} = (q_{l+1}, \ldots, q_{l+m}) \in F(M \setminus Q_l, m)$$

as the standard base point of the space $F(M \setminus Q_l, m)$. If $k < m$ we define the projection

$$\text{proj} : F(M \setminus Q_l, m) \to F(M \setminus Q_l, k)$$

by the formula: $\text{proj}(p_1, \ldots, p_m) = (p_1, \ldots, p_k)$. The following theorems were proved by E. Fadell and L. Neuwirth, [80].

**Theorem 8.1.** The triple $\text{proj} : F(M \setminus Q_l, m) \to F(M \setminus Q_l, k)$ is a locally trivial fiber bundle with fiber $\text{proj}^{-1} Q_{k,l}$ homeomorphic to $F(M \setminus Q_{k,l}, m-k)$.

Consideration of the sequence of fibrations

$$F(M \setminus Q_{m-1}, 1) \to F(M \setminus Q_{m-2}, 2) \to M \setminus Q_{m-2},$$

$$F(M \setminus Q_{m-2}, 2) \to F(M \setminus Q_{m-3}, 3) \to M \setminus Q_{m-3},$$

$$\cdots$$

$$F(M \setminus Q_1, m-1) \to F(M, m) \to M$$

leads to the following theorem.

**Theorem 8.2.** For any manifold $M$

$$\pi_i(F(M \setminus Q_1, m-1)) = \bigoplus_{k=1}^{m-1} \pi_i(M \setminus Q_k)$$

for $i \geq 2$. If $\text{proj} : F(M, m) \to M$ admits a section then

$$\text{proj}_i \pi_i(F(M, m)) = \bigoplus_{k=0}^{m-1} \pi_i(F(M \setminus Q_k)), \quad i \geq 2.$$

**Corollary 8.1.** If $M$ is Euclidean $r$-space, then

$$\pi_i(F(M, m)) = \bigoplus_{k=0}^{m-1} \pi_i(S^{r-1} \vee \cdots \vee S^{r-1}), \quad i \geq 2.$$
Corollary 8.2. If $M$ is Euclidean 2-space, then $F(\mathbb{R}^2, m)$ is the $K(P_m, 1)$-space and $B(\mathbb{R}^2, m)$ is the $K(\text{Br}_m, 1)$-space.

Let $X_W$ be the space defined in Section 7.2.

Theorem 8.3. The universal covering of $X_W$ is contractible, and so $X_W$ is a $K(\pi; 1)$-space.

This theorem for the groups of types $C_n, G_2$, and $I_2(p)$ was proved by E. Brieskorn, [42], in much the same way as E. Fadell and L. Neuwirth, [80], proved Theorems 8.1, 8.2 and Corollary 8.2. For the groups of types $D_n$ and $F_4$ E. Brieskorn used this method with minor modifications. In the general case Theorem 8.3 was proved by P. Deligne, [74].

It follows from Theorem 8.2 that $F(\mathbb{C} \setminus Q_k, n)$ and $B(\mathbb{C} \setminus Q_k)$ are $K(\pi, 1)$-spaces, that $\pi_1 B(\mathbb{C} \setminus Q_k) = \text{Br}_n$, and so, $B(\mathbb{C} \setminus Q_k)$ can be considered as the classifying space of $\text{Br}_n$.

8.2. Cohomology of pure braid groups

The cohomology of pure braid groups was first calculated by V. I. Arnold, [4]. The map

$$\phi : S^{n-1} \rightarrow F(\mathbb{R}^n, 2),$$

described by the formula $\phi(x) = (x, -x)$, is a $\Sigma_2$-equivariant homotopy equivalence. Denote by $A$ the generator of $H^{n-1}(F(\mathbb{R}^n, 2), \mathbb{Z})$ that is mapped by $\phi^*$ to the standard generator of $H^{n-1}(S^{n-1}, \mathbb{Z})$. For $i$ and $j$, such that $1 \leq i, j \leq m$, $i \neq j$, specify $\pi_{i,j}: F(\mathbb{R}^n, m) \rightarrow F(\mathbb{R}^n, 2)$ by the formula $\pi_{i,j}(p_1, \ldots, p_m) = (p_i, p_j)$. Put

$$A_{i,j} = \pi_{i,j}^*(A) \in H^{n-1}(F(\mathbb{R}^n, m), \mathbb{Z}).$$

It follows that $A_{i,j} = (-1)^n A_{j,i}$ and $A_{i,i}^2 = 0$. For $w \in \Sigma_m$ there is an action $w(A_{i,j}) = A_{w^{-1}(i), w^{-1}(j)}$, since $\pi_{i,j} w = \pi_{w^{-1}(i), w^{-1}(j)}$. Note also that under restriction to

$$F(\mathbb{R}^n \setminus Q_k, m-k) \cong \pi^{-1}(Q_k) \subset F(\mathbb{R}^n, m),$$

the classes $A_{i,j}$ with $1 \leq i, j \leq k$ go to zero since in this case the map $\pi_{i,j}$ is constant on $\pi^{-1}(Q_k)$.

Theorem 8.4. The cohomology group $H^*(F(\mathbb{R}^n \setminus Q_k, m-k), \mathbb{Z})$ is the free Abelian group with generators

$$A_{i_1,j_1} A_{i_2,j_2} \cdots A_{i_s,j_s},$$

where $k < j_1 < j_2 < \cdots < j_s \leq m$ and $i_r < j_r$ for $r = 1, \ldots, s$. 

The multiplicative structure and the $\Sigma_m$-algebra structure of $H^*(F(\mathbb{R}^n, m), \mathbb{Z})$ are given by the following theorem which is proved using the $\Sigma_3$-action on $H^*(F(\mathbb{R}^n, 3), \mathbb{Z})$.

**Theorem 8.5.** The cohomology ring $H^*(F(\mathbb{R}^n, m), \mathbb{Z})$ is multiplicatively generated by the square-zero elements

$$A_{i,j} \in H^{n-1}(F(\mathbb{R}^n, m), \mathbb{Z}), \quad 1 \leq i < j \leq m,$$

subject only to the relations

$$A_{i,k} A_{j,k} = A_{i,j} A_{j,k} - A_{i,j} A_{i,k} \quad \text{for } i < j < k. \quad (8.1)$$

The Poincaré series for $F(\mathbb{R}^n, m)$ is the product $\prod_{j=1}^{m-1} (1 + j^n t^{n-1})$.

**Remark 8.1.** In the case of $\mathbb{R}^2 = \mathbb{C}$ the cohomology classes $A_{j,k}$ can be interpreted as the classes of cohomology of the differential forms

$$\omega_{j,k} = \frac{1}{2\pi i} \frac{dz_j - dz_k}{z_j - z_k}.$$

E. Brieskorn calculated the cohomology of pure generalized braid groups, [42], using ideas of V.I. Arnold for the classical case. Let $\mathcal{V}$ be a finite-dimensional complex vector space and $H_j \in \mathcal{V}$, $j \in I$, be the finite family of complex affine hyperplanes given by linear forms $l_j$. E. Brieskorn proved the following fact.

**Theorem 8.6.** The cohomology classes, corresponding to the holomorphic differential forms

$$\omega_j = \frac{1}{2\pi i} \frac{dl_j}{l_j},$$

generate the cohomology ring $H^*(\mathcal{V} \setminus \bigcup_{j \in I} H_j, \mathbb{Z})$. Moreover, this ring is isomorphic to the $\mathbb{Z}$-subalgebra generated by the forms $\omega_j$ in the algebra of meromorphic forms on $\mathcal{V}$.

The cohomology of pure generalized braid groups is described as follows.

**Theorem 8.7.**

(i) The cohomology group $H^k(P(W), \mathbb{Z})$ of the pure braid group $P(W)$ with integer coefficients is a free Abelian group, and its rank is equal to the number of elements $w \in W$ of length $l(w) = k$, where $l$ is the length considered with respect to the system of generators consisting of all reflections of $W$.

(ii) The Poincaré series for $H^*(P(W), \mathbb{Z})$ is the product $\prod_{j=1}^{m_j} (1 + m_j t)$, where the $m_j$ are the exponents of the group $W$.

(iii) The multiplicative structure of $H^*(P(W), \mathbb{Z})$ coincides with the structure of the algebra generated by the 1-forms described in the previous theorem.
8.3. Homology of braid groups

To study the cohomology of the classical braid groups $H^*(Br_n, \mathbb{Z})$, V.I. Arnold, [5], interpreted the space $K(Br_n, 1) \cong B(\mathbb{R}^2, n)$ as the space of monic complex polynomials of degree $n$ without multiple roots

$$P_n(t) = t^n + z_1t^{n-1} + \cdots + z_{n-1}t + z_n.$$ 

Using this idea he proved theorems of finiteness, of recurrence and of stabilization. Homology with coefficients in $\mathbb{Z}/2$ was calculated by D.B. Fuks in the following theorems, [90].

**Theorem 8.8.** The homology of the braid group on an infinite number of strings with coefficients in $\mathbb{Z}/2$ as a Hopf algebra is isomorphic to the polynomial algebra on infinitely many generators $a_i, i = 1, 2, \ldots$; $\deg a_i = 2^i - 1$:

$$H_*(Br_\infty, \mathbb{Z}/2) \cong \mathbb{Z}/2[a_1, a_2, \ldots, a_i, \ldots]$$

with the coproduct given by the formula:

$$\Delta(a_i) = 1 \otimes a_i + a_i \otimes 1.$$ 

**Theorem 8.9.** The canonical inclusion $Br_n \to Br_\infty$ induces a monomorphism in homology with coefficients in $\mathbb{Z}/2$. Its image is the subcoalgebra of the polynomial algebra $\mathbb{Z}/2[a_1, a_2, \ldots, a_i, \ldots]$ with $\mathbb{Z}/2$-basis consisting of the monomials

$$a_1^{k_1} \cdots a_i^{k_i} \text{ such that } \sum_i k_i 2^i \leq n.$$ 

**Theorem 8.10.** The canonical homomorphism $Br_n \to BO_n, 1 \leq n \leq \infty$, induces a monomorphism (of Hopf algebras if $n = \infty$)

$$H_*(Br_n, \mathbb{Z}/2) \to H_*(BO_n, \mathbb{Z}/2).$$

F.R. Cohen calculated the homology of braid groups with coefficients $\mathbb{Z}/p$, $p > 2$, also as modules over the Steenrod algebra, [54–56].

Later V.V. Goryunov, [108,109], applied the methods of Fuks and expressed the cohomology of the generalized braid groups of types $C$ and $D$ in terms of the cohomology of the classical braid groups.

9. Connections with the other domains

9.1. Markov theorem

Suppose a braid depicted in Figure 1 is placed in a cube. On the boundary of the cube join the point $A_i$ to the point $B_i$ by a simple arc $D_i$, such that $D_i$ and $D_j$ are mutually disjoint
if \( i \neq j \). Since our initial braid does not intersect the boundary of the cube except at the points \( A_1, \ldots, A_n \) and \( B_1, \ldots, B_n \) we obtain a link (or, in particular, a knot), i.e. a system of simple closed curves in \( \mathbb{R}^3 \). A link obtained in such a manner is called the closure of the braid, see Figure 17.

**Theorem 9.1 (J.W. Alexander).** Any link can be represented by a closed braid.

The next step is to understand equivalence classes of braids which correspond to links. The following Markov theorem gives an answer to this question. At first we define two types of Markov moves for braids.

Type 1 Markov move replaces a braid \( \beta \) on \( n \) strings by its conjugate \( \gamma \beta \gamma^{-1} \).

Type 2 Markov move replaces a braid \( \beta \) on \( n \) strings by the braid \( j_n(\beta)\sigma_n \) on \( n + 1 \) strings or by \( j_n(\beta)\sigma_n^{-1} \) where \( j_n \) is the canonical inclusion of the group \( Br_n \) into the group \( Br_{n+1} \) (see Section 3.1)

\[
j_n : Br_n \rightarrow Br_{n+1}.
\]

**Theorem 9.2 (A.A. Markov).** Suppose that \( \beta \) and \( \beta' \) are two braids (not necessary with the same number of strings). Then, the closures of \( \beta \) and \( \beta' \) represent the same link if and only if \( \beta \) can be transformed into \( \beta' \) by means of a finite number of type 1 and type 2 Markov moves. Namely there exists the following sequence,

\[
\beta = \beta_0 \rightarrow \beta_1 \rightarrow \cdots \rightarrow \beta_m = \beta',
\]

such that, for \( i = 0, 1, \ldots, m-1 \), \( \beta_{i+1} \) is obtained from \( \beta_i \) by the application of a type 1 or 2 Markov moves or their inverses.

In other words, if we consider the disjoint union of all braid groups

\[
\bigsqcup_{n=1}^n Br_n.
\]
then the Markov moves of types 1 and 2 define the equivalence relation \( \sim \) on this set such that the quotient set

\[
\bigsqcup_{n=1}^{\infty} \frac{Br_n}{\sim}
\]

is in one-to-one correspondence with isotopy classes of links.

There exist a lot of proofs of Markov theorem, see, for example, the work of P. Traczyk, [185].

9.2. Homotopy groups of spheres and Makanin braids

Consider the coordinate projections for the spaces \( F(M,m) \) where \( M \) is a manifold (see Section 7.1)

\[
d_i : F(M, n+1) \to F(M, n), \quad i = 0, \ldots, n,
\]

defined by the formula

\[
d_i(p_1, \ldots, p_{i+1}, \ldots, p_{n+1}) = (p_1, \ldots, p_i, p_{i+2}, \ldots, p_{n+1}).
\]

By taking the fundamental group the maps \( d_i \) induces group homomorphisms

\[
d_i : \pi_1(F(M, n+1)) \to \pi_1(F(M, n)), \quad i = 0, \ldots, n.
\]

A braid \( \beta \in Br_{n+1} \) is called Makanin (smooth in the terminology of D.L. Johnson, [120], Brunnian in the terminology of J.A. Berrick, F.R. Cohen, Y.L. Wong and J. Wu, [22]) if \( d_i(\beta) = 1 \) for all \( 0 \leq i \leq n \). We call them Makanin, because up to our knowledge it was G.S. Makanin who first mentioned them, [127, page 78, Question 6.23]. In other words the group of Makanin braids \( Mak_{n+1}(M) \) is given by the formula

\[
Mak_{n+1}(M) = \bigcap_{i=0}^{n} \ker(d_{i,*} : \pi_1(F(M, n+1)) \to \pi_1(F(M, n))).
\]

The canonical embedding of the open disc \( D^2 \) into the sphere \( S^2 \)

\[
f : D^2 \to S^2
\]

induces a group homomorphism

\[
f_* : Mak_n(D^2) \to Mak_n(S^2),
\]

where \( Mak_n(D^2) \) is the Makanin subgroup \( Mak_n \) of the classical braid group \( Br_n \). The group \( Mak_n \) is free, [112,120]. The following theorem is proved in [22].
THEOREM 9.3. The is an exact sequence of groups

\[ 1 \to \text{Mak}_{n+1}(S^2) \to \text{Mak}_n(D^2) \to \pi_{n-1}(S^2) \to 1 \]

for \( n \geq 5 \).

Here as usual \( \pi_k(S^2) \) denote the \( k \)-th homotopy group of the sphere \( S^2 \).

For instance, \( \text{Mak}_5(S^2) \) modulo \( \text{Mak}_5 \) is \( \pi_4(S^2) = \mathbb{Z}/2 \). The other homotopy groups of \( S^2 \) are as follows

\[ \pi_5(S^2) = \mathbb{Z}/2, \ \pi_6(S^2) = \mathbb{Z}/12, \ \pi_7(S^2) = \mathbb{Z}/2, \ \pi_8(S^2) = \mathbb{Z}/2, \ldots \]

Thus, up to certain range, \( \text{Mak}_n(S^2) \) modulo \( \text{Mak}_n \) are known by nontrivial calculation of \( \pi_*(S^2) \).

Acknowledgements

The author is grateful to F.R. Cohen and Wu Jie for fruitful discussions on theory of braids and their applications. The author is also thankful to Vik.S. Kulikov who attracted his attention to the works of O. Zariski, [200,201], and to E.P. Volokitin for the help and advice on the presentation of the chapter.

References

Braids, their properties and generalizations


Braids, their properties and generalizations


Braids, their properties and generalizations


L. Paris, Centralizers of parabolic subgroups of Artin groups of type $A_1$, $B_1$, and $D_1$, J. Algebra 196 (2) (1997), 400–435.


R. Street, Braids among the groups, Seminarberichte aus dem Fachbereich Mathematik 63 (5) (1998), 699–703.


This page intentionally left blank
Groups with Finiteness Conditions

V.I. Senashov

Institute of Computational Modelling of Siberian Division of Russian Academy of Sciences, Krasnoyarsk, Russia
E-mail: sen@icm.krasn.ru

Contents
1. Finiteness conditions, examples ........................................ 469
2. Layer-finite groups .................................................... 471
3. Groups with layer-finite periodic part ................................ 474
4. Generalizations of layer-finite groups ................................ 476
5. Layer-Chernikov groups ................................................ 477
6. Generalized Chernikov groups ........................................ 479
7. $T_0$-groups ............................................................. 481
8. $\Phi$-groups ............................................................. 483
9. Almost layer-finite groups ............................................. 485
10. Periodic groups with minimality condition ......................... 486
11. Groups with finitely embedded involution ............................ 488
12. Frobenius groups and finiteness conditions ....................... 490
References ................................................................. 491

*The work is supported by the Russian Fund of Fundamental Researches (grant 05-01-00576).
1. Finiteness conditions, examples

Groups with conditions of finiteness were traditionally studied by V.P. Shunkov’s school. Weak conditions imposed on subgroups, on normalizers of finite subgroups, suddenly yield unexpected effects and extend over the whole group or give it some interesting properties.

Examples of infinite groups with different conditions of finiteness always played important role in the theory of infinite groups. This section cites definitions of different finiteness conditions and constructs several examples, illustrating non-coincidence and differences of the properties of some classes of infinite groups.

The conditions of biprimitive finiteness, conjugate biprimitive finiteness are consecutive weakening of the conditions of local finiteness and binary finiteness. They appear in the work of V.P. Shunkov in the 70ies when he was studying periodic groups. Here we consistently state these conditions of finiteness and give examples of biprimitively finite groups, which are not binary finite, constructed by M.Yu. Bakhova, [3], and A.A. Cherep, [4].

We shall need the following definitions:

A group \( G \) is called **locally finite**, if every finite set of elements in it generates a finite subgroup.

A group \( G \) is called **s-finite**, if any \( s \) elements in it generate a finite subgroup.

An \( s \)-finite group with \( s = 2 \) is called **binary finite**.

A group is called **conjugately \( n \)-finite**, if any its \( n \) conjugate elements generate a finite subgroup.

A group \( G \) is called **biprimitively finite**, if for every finite subgroup \( K \) from \( G \) every two elements of prime order in \( N_G(K)/K \) generate a finite subgroup.

A group \( G \) is called **\( p \)-biprimitively finite**, if for every finite subgroup \( K \) from \( G \) every two elements of the prime order \( p \) in \( N_G(K)/K \) generate a finite subgroup.

A group \( G \) is called **conjugately biprimitively finite** or a **Shunkov group**, if for every finite subgroup \( K \) from \( G \) in \( N_G(K)/K \) every two conjugate elements of prime order generate a finite subgroup.

This class of groups receive the title **Shunkov group** in the articles of L. Hammoudi, A.V. Rojkov, V.I. Senashov, A.I. Sozutov, A.K. Shlepkin.

These classes of groups have been introduced by V.P. Shunkov. These conditions have been successfully used in the proof of many theorems in for already more than thirty years.

Examples of periodic non-locally finite groups are not a sensation already: there are examples of S.P. Novikov, S.I. Adian, [1], examples of A.Yu. Ol’shanskii, [21], examples of E.S. Golod, [11]).

It is not difficult to see, that the class of finite groups belongs to the class of locally finite groups; the class of locally finite groups belongs to the class of binary finite groups; the class of binary finite groups belongs to the class of biprimitively finite groups.

From the Golod examples, [11], ensues, that there exist binary finite, but not locally finite groups.

There exist biprimitively finite groups, which are not binary finite. Examples of such groups have been constructed by M.Yu. Bakhova, [3], and A.A. Cherep, [4]. We shall cite both constructions.

Here is a construction of M.Yu. Bakhova, [3].

Let \( n \) be a composite number, \( F_n \) be a free group with \( n \) generators \( x_1, \ldots, x_n \). Denote by \( c \) an element from \( \text{Aut} F_n \), for which \( x^c_i = x_{i+1} \) (\( i = 1, \ldots, n - 1 \)) and \( x^c_n = x_1 \). Take
Let $\varphi$ be the homomorphism $F_n \rightarrow P$, where $P$ is a Golod $p$-group with $n$ generators. $V$ is the kernel of this homomorphism. By the homomorphism theorem, $[16]$, $F_n / V \cong P$. Obviously, $Vc^i$ $(i = 1, 2, \ldots, n)$ and $D = \bigcap_{i=1}^n Vc^i$ are normal subgroups in $F_n$. It is not difficult to show that $D$ is normal in $W$. Consider now the quotient group $G = W/D$. The subgroup $F_n/D$ from $G$ will be a sub-Cartesian product of groups isomorphic to $P$, and therefore the given subgroup is $(n-1)$-finite.

Let’s introduce notations:

$B = F_n/D$, $a = cD$, $b_i = x_iD$ $(i = 1, \ldots, n)$.

Then $G = B\langle a \rangle$ and $B = \langle b_1, \ldots, b_n \rangle$. But $b_i^2 = b_{i+1}$ $(i = 1, \ldots, n-1)$ and $b_1^2 = b_1$.

Hence, $G = \langle b_1, a \rangle$, and this means, that the group $G$ is not binary finite, with $\pi(G) = \pi(n) \cup \{p\}$.

Further, as in the group $G$ every subgroup and every quotient group are the extensions of an $(n-1)$-finite group by means of a cyclic group, and $n$ is a composite number, then the group $G$ is biprimitively finite. So, the following assertion is valid: there exists a biprimitively finite group $G$, which is not binary finite.

As it ensues from the structure of the above given group $G$, a finite extension of the binary finite group may be not a binary finite group.

Here is the example of A.A. Cherep, [4].

Consider a direct product $A = \prod_{i \in \mathbb{Z}} \langle a_i \rangle$ of cyclic groups $\langle a_i \rangle$ of order 2 and the group $G = \langle A\lambda\langle h \rangle\langle t \rangle \rangle$, where the element $h$ of infinite order acts on the generators from $A$ by the rule $h^{-1}a_ih = a_i+2$ $(i \in \mathbb{Z})$, and the action of the element $t$ of order 4 is determined by the equalities:

$$t^{-1}a_it = a_{-i} \quad (i \in \mathbb{Z}), \quad t^{-1}ht = h^{-1}a_0a_1.$$ 

It is obvious that for each $i \in \mathbb{Z}$, $t^{-4}a_i^4 = a_i$. It is also easy to check that

$$t^{-2}ht^2 = a_0a_1ha_0a_{-1} = ha_{-1}a_0a_2a_3,$$
$$t^{-3}ht^3 = h^{-1}a_0a_1a_0a_1a_{-2}a_{-3} = h^{-1}a_{-2}a_{-3},$$
$$t^{-4}ht^4 = a_0a_1ha_{2}a_{3} = h.$$ 

Hence, the group $G$ has been determined correctly. Let $B = \langle A, t^2 \rangle$, $L = \langle B, h \rangle$. As $t^2 \in C_G(A)$ and $t^2A \in Z(G/A)$, $B$ is an elementary Abelian 2-subgroup, and the periodic elements from $L$ lie in $B$. Further, the relation in the quotient-group $G/B$

$$(th^kB)^2 = t^2h^{-k}h^kB = B$$

shows that the elements from the set $G \setminus L$ have order 4.

In [4] is shown that the group $G$ is biprimitively finite, but not binary finite.

There is a V.P. Shunkov problem: are the classes of biprimitively finite and conjugately biprimitively finite groups different.

The following theorem solves this problem in the class of soluble groups.
Theorem 1.1 (A.A. Cherep, [5]). A soluble conjugately biprimitively finite group is biprimitively finite.

The proof of this theorem contains more than it states; in fact, it shows, that if in a soluble group a certain periodic element generates with each of its conjugates a finite subgroup, then it lies in the locally finite normal divisor.

A group $G$ is called group with unmixed factors if it has an ordered normal series

$1 = G_0 \leq G_1 \leq \cdots \leq G_\alpha = G,$

whose factors are either locally finite, or torsion-free.

It can be directly checked, that this property can be extended to the subgroups and quotient groups with respect to a periodic normal divisor. In addition, the periodic subgroups of a group with unmixed factors, are locally finite.

Theorem 1.2 (A.A. Cherep, [5]). If $G$ is a group with unmixed factors, then conditions (1)–(3) are equivalent:

1. The group $G$ is conjugately biprimitively finite.
2. The group $G$ is biprimitively finite.
3. For every finite subgroup $H$ in the quotient group $NG(H)/H$ the elements of prime order generate a locally finite subgroup.

The conditions of conjugate biprimitive finiteness and biprimitive finiteness coincide in a more general case.

Theorem 1.3 (A.A. Cherep, [5]). If in the group $G$ every two elements generate a subgroup with unmixed factors, then the conjugate biprimitive finiteness of $G$ is equivalent to its biprimitive finiteness.

The next theorem gives infinitely many examples of infinite groups which separated classes of $n$-finite and $(n + 1)$-finite $p$-group for an arbitrary large enough number $n$.

Theorem 1.4 (A.V. Rojkov, [26]). Let $p$ be a prime number, $1 \leq n \leq k$ be a natural number. Then there exists a finitely generated finitely approximate conjugately $k$-finite $n$-finite, but not a $(n + 1)$-finite $p$-group. In particular, a $p$-group can be non-binary finite, but conjugately $k$-finite, where $k$ is any large enough number.

2. Layer-finite groups

Another kind of finiteness conditions of groups is a finiteness of its layers (a layer is a set of elements of given order).

Layer-finite groups appeared for the first time in a paper by S.N. Chernikov written in 1945, [6]. This particular kind of group was mentioned, but without any name. Chernikov used the name of “layer-finite” in some of his later works.
A group is layer-finite if any set of its elements of any given order is finite.

Investigations of properties of layer-finite groups were carried out by S.N. Chernikov, R. Baer, Kh.Kh. Mukhamedjan in 1945–1960. The basic properties were described in different journals and remained in this form until 1980, when S.N. Chernikov’s book, [9], was published. In that book, S.N. Chernikov collected all of his results in one paragraph. It is possible to find practically all the properties of layer-finite groups and the nearest related questions in the monograph [30].

When all the properties of layer-finite groups were described, the question about the place of this class among other groups arose. The first of such characterizations of layer-finite groups was the establishment of the interconnection with the nearest class of groups: locally normal groups. The idea of this connection appeared in articles of S.N. Chernikov, R. Baer, Kh.Kh. Mukhamedjan, [2,6,7,19,20]. But more complete is a theorem by Chernikov:

**Theorem 2.1** (S.N. Chernikov, [9]). A class of layer-finite groups coincides with a class of locally normal groups if all their Sylow subgroups satisfy the minimality condition.

It should be noted, that a group \( G \) is called a Chernikov group if it is a finite extension of a direct product of a finite number of quasi-cyclic groups.

The Shmidt theorem on the closure of locally finite groups with respect to extensions by locally finite groups is valid for locally finite groups. There is no similar theorem for the class of layer-finite groups. Even the finite extensions of layer-finite groups lead us beyond the limits of this class. But nevertheless, some hereditary properties for layer-finite groups do exist:

**Theorem 2.2** (S.N. Chernikov, [9]). The thin layer-finite groups are precisely locally normal groups, all of whose Sylow subgroups are finite.

**Theorem 2.3** (S.N. Chernikov, [9]). A group \( G \), which is an extension of the layer-finite group by the layer-finite group, is layer-finite if and only if it is locally normal.

**Theorem 2.4** (S.N. Chernikov, [9]). If the group \( G \) can be represented in the form of a product of two layer-finite normal divisors, then the group \( G \) is layer-finite.

**Theorem 2.5** (V.I. Senashov, [28,29]). A periodic group is layer-finite if and only if it is conjugately biprimively finite and every one of its locally finite subgroups is layer-finite.

**Corollary 2.1.** If in a binary finite group any locally soluble subgroup is layer-finite, then the group is layer-finite too.

The statement of the corollary immediately follows from Theorem 2.5.

The condition of conjugate biprimitive finiteness in the theorem is necessary because there are examples such as the Novikov–Adian group, [1], and the Ol’shanskii group, [21], in which all the conditions of the theorem are valid except for the condition of finiteness, but they are not layer-finite.
It is impossible to make this condition weaker by replacing it by the $F^*$-condition (for the definition of the $F^*$-condition see below), because the Ol’shanskii group from [21] is an $F^*$-group and any of its locally soluble subgroups is layer-finite, but the group is not layer-finite. So the condition of conjugate biprimitive finiteness is limiting in the theorem.

The next theorem gives one more characterization of layer-finite groups in the class of periodic binary soluble groups: here the condition of layer-finiteness is not global, but only for subgroups with the element $a$.

**Theorem 2.6** (V.O. Gomer, [12]). Let $G$ be a periodic binary soluble group, and $a$ an element of prime order $p$ such that:

1. in $C_G(a)$ every locally finite subgroup is layer-finite and has finite Sylow $q$-subgroups for all primes $q$;
2. every locally finite subgroup with the element $a$ is layer-finite.

Then $G$ is a locally soluble (locally finite) layer-finite group.

The next results characterized layer-finite groups in the class of periodic almost locally soluble groups with the condition: the centralizer of any non-identity element from some elementary Abelian subgroup of order $p^2$ is layer-finite.

**Theorem 2.7** (M.N. Ivko, [13]). Let $G$ be a periodic almost locally soluble group, possessing an elementary Abelian subgroup $V$ of order $p^2$. If the centralizer in $G$ of any non-identity element from $V$ is layer-finite, then the group $G$ is layer-finite.

**Theorem 2.8** (M.N. Ivko, [13]). The 2-biprimitively-finite group $G$ of the form $G = H \lambda L$, where $H$ is the subgroup without involutions, and $L$ is the Klein four group, is layer-finite if and only if the centralizer in $G$ of any involution from $L$ is layer-finite.

Recall that an element of the order 2 is called an *involution*.

The next corollary characterizes layer-finite groups in the class of periodic groups without involutions.

**Corollary 2.2.** The periodic group $H$ without involutions whose holomorph contains a Klein four subgroup $L$, is layer-finite if and only if the centralizer in $H$ of any involution from $L$ is layer-finite.

A group $G$ satisfies the *p-minimality condition* (min-$p$ condition), if every descending chain of subgroups of it

$$H_1 > H_2 > \cdots > H_n > \cdots$$

is such that if $H_n \setminus H_{n+1}$ contains $p$-elements for all $n$ it terminates at a finite number.

By $\pi(G)$ we shall denote a set of prime divisors of orders of elements of the group $G$.

A group $G$ satisfies the *primary minimality condition*, if it satisfies the $p$-minimality condition for every prime number $p \in \pi(G)$. 

---

*Groups with finiteness conditions*
A group $G$ is called \textit{finitely approximate} if for every set of different elements of $G$ there is a homomorphism of $G$ on a finite group such that the images of these elements are different.

The group $G$ is called an $F_q$-\textit{group} ($q \in \pi(G)$), if for every finite subgroup $H$ and any elements $a, b \in T = N_G(H)/H$ of order $q$ there exists an element $c \in T$, such that $\langle a, c^{-1}bc \rangle$ is finite.

If every subgroup from $G$ is an $F_q$-group, then $G$ is called an $F_q^*$-\textit{group}. If $G$ is a $F_q^*$- (respectively, $F_q^{*+}$) group for any number $q \in \pi(G)$, then $G$ is called simply an $F$- (respectively, $F^*$-) \textit{group}.

The next theorem gives the feature of layer-finiteness of periodic finitely approximate $F^*$-group, in which locally finite subgroups are layer-finite.

\textbf{THEOREM 2.9} (E.I. Sedova, [27]). A periodic finitely approximate $F^*$-group, in which every locally finite subgroup is layer-finite, is a layer-finite group.

On the base of this result, descriptions of locally solvable layer-finite groups and groups with primary minimality condition in the class of binary solvable groups are given by the next theorems.

\textbf{THEOREM 2.10} (E.I. Sedova, [27]). A periodic group is a locally soluble layer-finite group if and only if it is binary soluble and every one of its locally soluble subgroup is layer-finite.

\textbf{THEOREM 2.11} (E.I. Sedova, [27]). A periodic group is a locally soluble group with min-$p$ condition if and only if it is binary soluble and every one of its locally soluble subgroups satisfies min-$p$ condition.

Recall that a group $G$ satisfies the \textit{$p$-minimality condition} (min-$p$ condition), if every descending chain of its subgroups

$$H_1 > H_2 > \cdots > H_n > \cdots$$

is such, that if $H_n \backslash H_{n+1}$ contains $p$-elements for all $n$, it terminates at a finite number.

\section{Groups with layer-finite periodic part}

Remind that the \textit{periodic part} of groups is a set of all of its elements that have finite order if they form a subgroup.

Here we aducce some criteria of layer-finiteness for the periodic part of a group.

\textbf{THEOREM 3.1} (V.I. Senashov, [32]). A group has a layer-finite periodic part if and only if it is conjugately biprimitively finite and every one of its locally solvable subgroup is layer-finite.
Theorem 3.2 (M.N. Ivko, V.P. Shunkov, [15]). A group $G$ without involutions has a layer-finite periodic part if and only if in it for some element $a$ of prime order $p$ the following conditions hold:

1. the normalizer of any non-trivial $\langle a \rangle$-invariant finite elementary Abelian subgroup of the group $G$ has a layer-finite periodic part;
2. almost all subgroups of the form $\langle a, a^g \rangle$ are finite.

Corollary 3.1. A group $G$ without involutions is layer-finite if and only if in it for some element $a$ of prime order $p$ the following conditions hold:

1. the normalizer of any non-trivial $\langle a \rangle$-invariant finite elementary Abelian subgroup from $G$ is layer-finite;
2. almost all subgroups of the form $\langle a, a^g \rangle$ are finite.

To obtain the characterization of groups having a layer-finite periodic part, in one of the group classes not containing involutions, it would be reasonable to obtain such a characterization in the general case. This problem with some additional limitations is solved by:

Theorem 3.3 (M.N. Ivko, V.P. Shunkov, [15]). A group $G$, containing an element $a$ of prime order $p \neq 2$ has a layer-finite periodic part if and only if the following conditions hold:

1. the normalizer of any non-trivial $\langle a \rangle$-invariant finite subgroup of the group $G$ has a layer-finite periodic part;
2. any locally finite subgroup, containing the element $a$, is almost locally soluble;
3. all subgroups of the form $\langle a, a^g \rangle$, where $g \in G$, are finite and almost all of them are soluble.

Corollary 3.2. An infinite group $G$ is layer-finite if and only if in it for some element $a$ of prime order $p \neq 2$ the following conditions are fulfilled:

1. the normalizer of any non-trivial $\langle a \rangle$-invariant finite subgroup of the group $G$ is layer-finite;
2. any locally finite subgroup, containing the element $a$, is almost locally soluble;
3. all subgroups of the form $\langle a, a^g \rangle$, where $g \in G$, are finite.

Removing in Theorem 3.3 condition (2) and replacing the condition (1) with a stronger restriction, we can obtain another characterization of groups, having a layer-finite periodic part. Namely, the following theorem is valid.

Theorem 3.4 (M.N. Ivko, V.P. Shunkov, [15]). An infinite group $G$ has a layer-finite periodic part if and only if in it for some element $a$ of prime order $p \neq 2$ the following conditions are fulfilled:

1. the normalizer of any non-trivial $\langle a \rangle$-invariant locally finite subgroup of the group $G$ has a layer-finite periodic part;
2. all subgroups of the form $\langle a, a^g \rangle$, where $g \in G$, are finite and almost all are soluble.
COROLLARY 3.3. An infinite group \( G \) is layer-finite if and only if in it for some element \( a \) of prime order \( p \neq 2 \) the following conditions are valid:

1. the normalizer of any non-trivial \( \langle a \rangle \)-invariant locally finite subgroup of the group \( G \) is layer-finite;
2. all subgroups of the form \( \langle a, a^g \rangle \), where \( g \in G \), are finite and almost all are soluble.

We have so characterized groups with layer-finite periodic part with \( \Phi \)-group accuracy (see the definition in the section “\( \Phi \)-groups”).

THEOREM 3.5 (M.N. Ivko, V.I. Senashov, [14]). Let \( G \) be a group, \( a \) an involution of it, that satisfies the following conditions:

1. all the subgroups of the form \( \langle a, a^g \rangle \), \( g \in G \), are finite;
2. the normalizer of every non-trivial \( \langle a \rangle \)-invariant finite subgroup has a layer-finite periodic part.

Then either the set of all finite order elements generates a layer-finite group or \( G \) is a \( \Phi \)-group.

4. Generalizations of layer-finite groups

Now, let’s discuss a few different generalizations of layer-finite groups as described by L.A. Kurdachenko in [17,18]. These generalizations appeared under different considerations of layer-finite groups.

The first generalization appeared on the base of the definition of layer-finite group: from the definition remove the demand of finiteness from a finite number of layers. Such groups are named \( QLF \)-groups. For such groups there holds

THEOREM 4.1 (L.A. Kurdachenko, [17]). For every locally finite \( QLF \)-group is either an almost layer-finite extension of finite group by a \( QLF \)-group, or all non-trivial layers of it are infinite (so, the group has a finite number of layers).

COROLLARY 4.1. The orders of the elements of a locally normal group, which has at least one infinite layer, bounded globally.

Layer-finite groups can be consider as groups, in which every primary layer is finite. On this basis one more generalization of layer-finiteness is constructed.

A group \( G \) is named an \( LB \)-group, if the number of all its infinite primary layers is finite.

THEOREM 4.2 (L.A. Kurdachenko, [17]). Let \( G \) be a locally finite \( LB \)-group. Then it is an extension of locally normal \( LB \)-group by a group with a global bound on the orders of its elements. If, in particular, all Sylow subgroups of a group \( G \) are Chernikov, then it is almost layer-finite.

Earlier, layer-finite groups were characterized as locally normal groups with Chernikov Sylow subgroups. Hence appears a second generalization of layer-finite group: groups in
which the Sylow $p$-subgroups are non-Chernikov only for a finite set of prime numbers $p$ (QSE-group). More precisely:

A periodic group $G$ is called a QSE-group, if only for a finite (and particularly an empty set) of prime numbers the $p$ Sylow $p$-subgroups of the group $G$ are non-Chernikov.

The next theorem shows the relation between QSE-groups and LB-groups under the condition of locally normality.

**Theorem 4.3** (L.A. Kurdachenko, [17]). A locally normal QSE-group is LB-group if and only if when the orders of all its elements are globally bounded.

Further developments of investigations on these generalizations of layer-finite groups can be find in the paper [17] under additional conditions of finiteness of classes of conjugate elements.

### 5. Layer-Chernikov groups

Here we consider layer-Chernikov groups, i.e. the groups, in which every set of elements of the same order generates a Chernikov group. This class of groups was called layer-extreme groups in the work of Ya.D. Polovitskii, [24], in 1960, when the name “Chernikov groups” for the class of finite extensions of the Abelian groups with minimality condition had not yet been settled down and such groups were called extreme or $ch$-groups.

The layer-Chernikov groups are close to layer-finite groups not in the sense of the definitions only, but also in their properties. In Section 2, in particular, the layer-finite groups have been specified as locally normal groups with Chernikov Sylow $p$-subgroups. The same role with respect to the layer-Chernikov groups is played by the locally Chernikov groups. A group is called locally Chernikov if every element is contained in its Chernikov normal divisor. Further we examine the subclass of the locally Chernikov groups: the locally finite groups, whose quotient groups by every $p$-center are Chernikov (the $p$-center of an arbitrary group is the name of the intersection of the centralizers of all its $p$-elements). This allows to define the class of layer-Chernikov groups in a new fashion.

A direct product of a number of periodic groups is called a primary thin direct product, if whatever is the prime number $p$, the $p$-elements are contained in not more, than in a finite number of direct factors.

**Theorem 5.1** (Ya.D. Polovitskii, [25]). The layer-Chernikov groups and only they are the subgroups of the primary thin direct products of Chernikov groups.

This theorem can be strengthened, taking into account Theorem 8.1 of Chernikov from [8] as follows:

**Theorem 5.2.** A layer-Chernikov group embeds into a primary thin direct product of such Chernikov groups, that the maximum complete subgroup of every one of them is a $p$-group, with the orders of the elements of every two maximum complete subgroups of different multipliers of this direct product being mutually prime.
COROLLARY 5.1. The quotient group of the layer-Chernikov group with respect to the arbitrary $p$-center is a Chernikov group.

A description of layer-Chernikov groups is given by the next theorem:

**THEOREM 5.3** (Ya.D. Polovitskii, [25]). The group $G$ is layer-Chernikov if and only if it decomposes into a product $G = AB$ where $A$ is an layer-finite complete Abelian group that is invariant in $G$ and $B$ is a thin layer-finite group, and for every number $p \in \pi(G)$ all Sylow $q$-subgroups of the group $A$ ($q$ is a prime number), except for, maybe, a finite number, are in the $p$-center of the group $G$.

A layer-Chernikov group does not necessarily decompose into the semi-direct product of a complete Abelian group and a thin layer-finite group, as such an assertion does not take place even for the layer-finite groups.

The following theorems describe relations between the locally Chernikov and layer-Chernikov groups.

**THEOREM 5.4** (Ya.D. Polovitskii, [25]). The class of layer-Chernikov groups coincides with the class of locally Chernikov groups which have Chernikov Sylow $p$-subgroups (with respect to all $p$).

**THEOREM 5.5** (Ya.D. Polovitskii, [25]). A central extension of a periodic group by means of a layer-Chernikov group is a locally Chernikov group.

The next theorem gives a property of $\pi$-minimality of a locally Chernikov group.

**THEOREM 5.6** (Ya.D. Polovitskii, [25]). A locally Chernikov group satisfies the $\pi$-minimality condition if and only if all its Sylow $\pi$-subgroups are Chernikov ones.

The locally Chernikov groups can also be determined in a different way as seen from

**THEOREM 5.7** (Ya.D. Polovitskii, [25]). A group $G$ is locally Chernikov if and only if every Chernikov subgroup of it is contained in some Chernikov normal divisor of the group $G$.

**COROLLARY 5.2.** A central extension of a layer-Chernikov group by means of a layer-Chernikov group is a layer-Chernikov group.

The assertion, analogous to Theorem 5.5 for central extensions by means of layer-finite groups has been proved by I.I. Yeremin, [10].

Let’s add some more results for layer-Chernikov groups.

**THEOREM 5.8** (Ya.D. Polovitskii, [25]). The quotient groups of a locally finite group $G$ by each $p$-center are Chernikov if and only if the group $G$ is a central extension of a periodic group by means of a layer-Chernikov group.
THEOREM 5.9 (Ya.D. Polovitskii, [44]). A group $G$ is a central extension of a periodic group by means of a thin layer-finite group if and only if its quotient groups are finite over each $p$-center.

COROLLARY 5.3. A locally finite group such that the quotient groups over each $p$-center are Chernikov is locally Chernikov.

COROLLARY 5.4. A group $G$ is layer-Chernikov if and only if all its Sylow $p$-subgroups are Chernikov and the quotient groups by each $p$-center are Chernikov.

It is of interest to note that by virtue of Theorem 5.3, a layer-Chernikov group is an extension of a complete Abelian group with Chernikov Sylow subgroups by means of the layer-finite group.

6. Generalized Chernikov groups

It is very natural to use the theory of layer-finite groups when one studies generalized Chernikov groups. Because such groups with conditions of a periodic type and almost locally solvability are extensions of layer-finite groups by layer-finite groups.

The imposition of restrictions on chains of subgroups has often been used in investigations concerning the structure of infinite groups. Many authors have in particular considered groups satisfying the minimality condition on all subgroups or the minimality condition on all Abelian subgroups. On the other hand, generalizing the concept of a locally finite group, S.P. Strunkov introduced in [59] binary finite groups, and later V.P. Shunkov, [46], considered the wider class of conjugately biprimitively finite groups. One of the main results relating to these finiteness conditions is due to A.N. Ostylovskii and V.P. Shunkov, [22], and states that a conjugate biprimitive finite group without involutions and satisfying the minimality condition on subgroups is a Chernikov group. Examples of P.S. Novikov, S.I. Adian, [1], and A.Yu. Ol’shanskii, [21], show that this result cannot be generalized to arbitrary periodic groups without involutions.

We shall say that a group $G$ satisfies the primary minimality condition if for each prime $p$ every chain

$$G_1 > G_2 > \cdots > G_n > \cdots$$

of subgroups of $G$, such that each set $G_n \setminus G_{n+1}$ contains an element $g_n$ with $g_n^{p^{k_n}} \in G_{n+1}$ for some $k_n$, stops after finitely many steps.

An almost locally soluble group satisfying the primary minimality condition will be called a generalized Chernikov group. This name is motivated by the following result: every almost locally soluble group $G$ satisfying the primary minimality condition contains a complete part $\tilde{G}$, the quotient group $G/\tilde{G}$ is locally normal, and every element of $G$ centralizes all but finitely many Sylow subgroups of $\tilde{G}$ (see, for instance, [23]).

Recall that a group $G$ has a complete part $A$, if $A$ is an Abelian group generated by all complete Abelian subgroups from $G$ and $G/A$ does not have complete Abelian subgroups.
The next results characterize generalized Chernikov groups in the class of periodic groups without involutions and in the class of mixed groups.

**Theorem 6.1** (V.I. Senashov, [31]). Let $G$ be a periodic group without involutions. Then $G$ is a generalized Chernikov group if and only if it is conjugately biprimitively finite and the normalizers of all its finite non-trivial subgroups are generalized Chernikov groups.

The earlier mentioned examples by Novikov, Adian, [1], and Ol’shanskii, [21], prove that in this theorem the condition that $G$ is conjugately biprimitively finite cannot be removed.

It is possible to consider class of groups with generalized Chernikov periodic part of the normalizer of any finite non-trivial subgroup. Among these groups there are the examples of the Novikov–Adian and Ol’shanskii groups.

The term “generalized Chernikov groups” was first used in [57]. Its use can be justified by the fact that according to the theorem of Ya.D. Polovitskii a generalized Chernikov group $G$ is an extension of the direct product $A$ of quasi-cyclic $p$-groups with a finite number of multipliers for any prime number $p$ by a locally normal group $B$, and each of the elements from $B$ is element-wise non-permutable with only a finite number of Sylow primary subgroups from $A$. For comparison a Chernikov group is a finite extension of a direct product of finite number of quasi-cyclic groups.

Here some properties of generalized Chernikov groups.

**Theorem 6.2** (V.I. Senashov, [33]). In a generalized Chernikov group primary subgroups are Chernikov.

**Theorem 6.3** (V.I. Senashov, [33]). If a generalized Chernikov group $G$ does not have complete subgroups, it is a thin layer-finite group.

**Theorem 6.4** (V.I. Senashov, [33]). In a non-Chernikov generalized Chernikov group any element has infinite centralizer.

Let's introduce the definition of a $T$-group.

**Definition.** Let $G$ be a group with involutions. Each involution $i$ from $G$ is associated with a subgroup $V_i$ from $G$ defined as follows. If the Sylow 2-subgroups from the order eighth dihedral group and $i$ are contained in a Klein four-subgroup $R_i$ such that $C_G(i) < N_G(R_i)$ and $C_G(i)$ has an infinite torsion subgroup, we take $V_i = N_G(R_i)$. In all other cases $V_i$ is taken to be $V_i = C_G(i)$.

A group $G$ with involutions will be said to be a $T$-group if it satisfies the conditions:

1. any two involutions from $G$ generate a finite subgroup;
2. the normalizer of any locally finite subgroup from $G$ containing involutions has locally finite periodic part;
3. the set $G \setminus V_i$ possesses involutions and $V_i$ is an infinite subgroup for every involution $i$ from $G$;
4. for every element $c$ from $G \setminus V_i$ strictly real with respect to $i$, for which $e^i = c^{-1}$, there exists in $C_G(i)$ such an element $s_c$, that the subgroup $\langle c, e^{cs} \rangle$ is infinite.
The class of $T$-groups has been introduced by V.P. Shunkov.
Recall that a group, generated by two involutions is called a dihedral group. An element $a$ of a group $G$ is called strictly real with respect to an involution $i \in G$, if $iai^{-1} = a^{-1}$.
In 1987 there appeared a conjecture by V.P. Shunkov according to which his Theorem 3.1 from [51] could possibly generalize to more wide classes of groups. And, indeed, in Theorem 6.5 below the second condition of the Shunkov theorem is replaced by a weaker condition: a normalizer with involutions of every finite non-trivial subgroup has generalized Chernikov periodic part. We proved that under this condition and if every subgroup generated by two involutions is finite, then it has a generalized Chernikov periodic part or it is $T$-group. So, the conjecture is proved.

**THEOREM 6.5** (V.I. Senashov, [33,34]). Let $G$ be a group with involutions satisfying the conditions:

1. any two involutions from $G$ generate a finite subgroup;
2. the normalizer of any finite non-trivial subgroup containing involutions has generalized Chernikov periodic part.

Then either $G$ has a generalized Chernikov periodic part or $G$ is a $T$-group.

The next theorem characterizes groups with generalized Chernikov periodic part.

**THEOREM 6.6** (V.I. Senashov, [33,34]). A group has a generalized Chernikov periodic part if and only if it is conjugate biprimitive finite and the normalizer of any finite non-trivial subgroup containing involutions has generalized Chernikov periodic part.

7. $T_0$-groups

At the beginning of the 90ies the concept of a $T_0$-group appeared. This class is defined by finiteness conditions. Here is the definition of the class of $T_0$-groups (V.P. Shunkov).

**DEFINITION.** Let $G$ be a group with involutions, and let $i$ be one of its involutions. We shall call the group $G$ a $T_0$-group, if it satisfies the following conditions (for the given involution $i$):

1. all subgroups of the form $\langle i, ig \rangle$, $g \in G$, are finite;
2. Sylow 2-subgroups from $G$ are cyclic or generalized groups of quaternions;
3. the centralizer $C_G(i)$ is infinite and has a finite periodic part;
4. the normalizer of any non-trivial $\langle i \rangle$-invariant finite subgroup from $G$ is either contained in $C_G(i)$, or has a periodic part being a Frobenius group with Abelian kernel and a finite complement of even order;
5. $C_G(i) \neq G$ and for any element $c$ from $G \setminus C_G(i)$, strictly real relating $i$ (i.e. such that $c^i = c^{-1}$), there is an element $s_c$ in $C_G(i)$, such that the subgroup $\langle c, c^s \rangle$ is infinite.

Let’s consider the construction of Shunkov’s example of a $T_0$-group from [52] based on the well-known example of S.P. Novikov, S.I. Adian, [1].
EXAMPLE OF A $T_0$-GROUP [52]. Let $A = A(m, n)$ be a torsion-free group $A(m, n)$, which is a central extension of cyclic group with help of group $B(m, n)$, $m > 1$, $n > 664$ an odd number, [1]. The group $A(m, n)$ has non-trivial center $Z(A) = \langle d \rangle$ and $A/\langle d \rangle$ is isomorphic to $B(m, n)$, [1]. Let’s consider a group $B = A \langle x \rangle$, where $x$ is an involution.

Let’s take an element $u = d \cdot d^{-x}$ from $A \times A^x$. It is obvious, that $u \in Z(A \times A^x)$, and $u^x = u^{-1}$. As is shown in [52], the group $G = B/\langle u \rangle$, and it’s involution $i = x \cdot \langle u \rangle$ satisfy conditions (1)–(5) from the definition of a $T_0$-group, and $G = V\lambda(i)$, $C_G(i)$ is an infinite group with periodic part $(i)$, all subgroups $(i, i^8)$ in $G$ are finite and every maximal finite subgroup from $G$ with involution $i$ is a dihedral group of the order $2n$ and $G$ is a $T_0$-group.

EXAMPLE OF A $T_0$-GROUP [53]. Let $V = O(p)$ (see the definition of groups of the type $O(p)$, $C(\infty)$ in the introduction). The group $V$ has a non-trivial center $Z(V) = \langle t \rangle$ and $V/Z(V) = V/(t) \simeq C(\infty)$, [16].

Consider the group $T = V \ltimes (k) = (V \times V)\lambda(k)$, where $k$ is an involution. Let us take from $V \times V$ the element $b = (t, t^{-1})$. Obviously, $b \in Z(V \times V)$ and $b^k = b^{-1}$. Let us take quotient group $M = T/(b)$, and in it an involution $j = k(b)$. Further, using abstract properties of the groups $V = O(p)$, $C(\infty)$, [16], it is easy to show that the group $M$ and its involution $j$ satisfy the conditions (1)–(5) of the definition. Hence, $M = T/(b)$ is a $T_0$-group (with respect to the involution $j = k(b)$). Let us also note that in $M$ any maximum periodic subgroup containing the involution $j$ is a dihedral group of order $2p$.

Here are some results on $T_0$-groups. Details can be found in [55].

**Theorem 7.1** (V.P. Shunkov, [53,56]). Let $G$ be a group and $a$ be an element of prime order $p$, satisfying the following conditions:

1. the subgroups of the form $\langle a, a^g \rangle$, $g \in G$, are finite and almost all are solvable;
2. in the centralizer $C_G(a)$ the set of elements of finite order is finite;
3. in the group $G$ the normalizer of any non-trivial $\langle a \rangle$-invariant finite subgroup has periodic part;
4. for $p \neq 2$ and for $q \in \pi(G)$, $q \neq p$, any $\langle a \rangle$-invariant elementary Abelian $q$-subgroup of $G$ is finite.

Then either $G$ has an almost nilpotent periodic part, or $G$ is a $T_0$-group and $p = 2$.

**Corollary 7.1.** Let $G$ be a (periodic) group and let $a$ be an element of prime order $p \neq 2$, satisfying conditions (1)–(4) of Theorem 7.1.

Then $G$ has an almost nilpotent periodic part.

The following statement is equivalent to Theorem 7.1 and gives an abstract characterization of $T_0$-groups in the class of all groups.

**Corollary 7.2.** Let $G$ be a group, $a$ be an element of prime order $p$. The group $G$ is a $T_0$-group and $p = 2$ if and only if for the pair $(G, a)$ the conditions (1)–(4) of theorem are satisfied and the subgroup $\langle a^g \mid g \in G \rangle$ is not periodic almost nilpotent.

The particular case when $p = 2$ requires special consideration, since in this case condition (4) of Theorem 7.1 is superfluous, i.e. the following statements are true.
COROLLARY 7.3. Let $G$ be a group with involutions and $i$ be one of its involutions, satisfying the following conditions:

1. the subgroups of the form $\langle i, i^g \rangle$, $g \in G$, are finite;
2. in the centralizer $C_G(i)$ the set of elements of finite order is finite;
3. in the group $G$ the normalizer of any non-trivial $\langle i \rangle$-invariant finite subgroup has periodic part.

Then either $G$ has almost nilpotent periodic part, or $G$ is a $T_0$-group.

The conditions (1)--(3) of Corollary 7.3 are independent, i.e. none of them follows from the other two.

COROLLARY 7.4. Let $G$ be a group with involutions, and let $i$ be one of its involutions. The group $G$ is a $T_0$-group if and only if for the pair $(G, i)$ conditions (1)--(3) of Corollary 7.3 are satisfied and the subgroup $\langle i^g \mid g \in G \rangle$ is not periodic almost nilpotent.

THEOREM 7.2 [53]. Let $G$ be a group with involutions and $i$ be an involution, satisfying the following conditions:

1. the subgroups of the form $\langle i, i^g \rangle$, $g \in G$, are finite;
2. in the centralizer $C_G(i)$ the set of elements of finite order is finite;
3. in the group $G$ the normalizer of any non-trivial $\langle i \rangle$-invariant finite subgroup has periodic part.

Then either $G$ has almost nilpotent periodic part, or $G$ is a $T_0$-group.

THEOREM 7.3 [53]. Let $G$ be a group and $a$ be an element of prime order $p$, satisfying the following conditions:

1. subgroups of the form $\langle a, a^g \rangle$, $g \in G$, are finite and almost all are solvable;
2. the centralizer $C_G(a)$ is finite;
3. for $p \neq 2$ and for $q \in \pi(G)$, $q \neq p$, any $\langle a \rangle$-invariant elementary Abelian $q$-subgroup of $G$ is finite.

Then $G$ is a periodic almost nilpotent group.

THEOREM 7.4 [53]. A non-trivial finitely generated group $G$ is finite if and only if in it there exists an element $a$ of prime order $p$ satisfying the following conditions:

1. the subgroups of the form $\langle a, a^g \rangle$, $g \in G$, are finite and almost all are solvable;
2. the centralizer $C_G(a)$ is finite;
3. when $p \neq 2$ and for $q \in \pi(G)$, $q \neq p$, any $\langle a \rangle$-invariant elementary Abelian $q$-subgroup is finite.

Theory of $T_0$-groups was created by V.P. Shunkov in [55].

8. $\Phi$-groups

This section investigates properties of the new class of $\Phi$-groups. Such groups are very close to $T_0$-groups, but in this sections we also point out the difference.
This group class is rather broad: among them are groups of Burnside type, [1], Ol’shanskii monsters, [21]. It is very closely connected with the groups of Burnside type of odd period \( n \geq 665 \).

**Definition.** Let \( G \) be a group, let \( i \) be an involution of \( G \), satisfying the following conditions:

1. all subgroups of the form \( \langle i, i^g \rangle \), \( g \in G \), are finite;
2. \( C_G(i) \) is infinite and has a layer-finite periodic part;
3. \( C_G(i) \neq G \) and \( C_G(i) \) is not contained in other subgroups from \( G \) with a periodic part;
4. if \( K \) is a finite subgroup from \( G \), which is not inside \( C_G(i) \), and \( V = K \cap C_G(i) \neq 1 \), then \( K \) is a Frobenius group with complement \( V \).

The group \( G \) with a specified involution \( i \) satisfying these conditions (1)–(4) is called a \( \Phi \)-group.

This class of groups has been introduced by V.P. Shunkov.

**Example of a \( \Phi \)-group (V.P. Shunkov).** Let \( A = \langle b, c \rangle \) (where \( b^n = c^n = d \) and \( n \) is a positive integer) be a torsion free group and let \( A/\langle d \rangle \) be the free Burnside group with period \( n \), [1]. Consider the group \( B = A \wr \langle x \rangle = (A \times A) \lambda \langle x \rangle \), where \( x \) is an involution. Let us take from \( A \times A \) the element \( v = (d, d^{-1}) \). Obviously \( v \in Z(A \times A) \) and \( v^x = v^{-1} \).

Further, the group \( G = B/\langle v \rangle \) and its involution \( i = x \langle v \rangle \) (which is easy to see from the abstract properties of the group \( A = \langle b, c \rangle \), [1]) satisfy all the conditions from the definition of a \( \Phi \)-group. Hence, \( G = B/\langle v \rangle \) is a \( \Phi \)-group.

**Theorem 8.1 (V.I. Senashov, [14]).** A \( \Phi \)-group \( G \) satisfies the properties:

1. all involutions are conjugate;
2. Sylow 2-subgroups are locally cyclic or finite generalized quaternion groups;
3. there are infinitely many elements of finite order in \( G \), which are strictly real with respect to the involution \( i \) and for every such element \( c \) of this set there exists an element \( s_c \) from the centralizer of \( i \) such that \( \langle c, c^{s_c} \rangle \) is an infinite group.

V.P. Shunkov posed the problem of studying groups with some additional limitations in the form that for the given finite subgroup \( B \), the next condition is valid: the normalizer of any non-trivial \( B \)-invariant finite subgroup has a layer-finite periodic part.

This problem is partly solved in the class of locally soluble groups and for the case \( |B| = 2 \) under more general limitations, it is solved with \( \Phi \)-groups accuracy.

**Theorem 8.2 (M.N. Ivko, V.I. Senashov, [14]).** A periodic locally soluble group is layer-finite if and only if for some finite subgroup \( B \) of it the next condition is valid: the normalizer of any non-trivial \( B \)-invariant finite subgroup is layer-finite.

**Theorem 8.3 (M.N. Ivko, V.I. Senashov, [14]).** Let \( G \) be a group, let \( a \) be an involution of \( G \), satisfying the conditions:

1. all subgroups of the form \( \langle a, a^g \rangle \), \( g \in G \), are finite;
(2) the normalizer of every non-trivial \( \langle a \rangle \)-invariant finite subgroup has a layer-finite periodic part.

Then either the set of all elements of finite orders forms a layer-finite group or \( G \) is an \( \Phi \)-group.

**COROLLARY 8.1** (M.N. Ivko, V.I. Senashov, [14]). Let \( G \) be a group with involutions and let \( i \) be some involution from \( G \) satisfying the conditions:

1. \( G \) is generated by the involutions which are conjugate with \( i \);
2. almost all groups \( \langle i, i^g \rangle \) are finite, \( g \in G \);
3. the normalizer of every \( \langle i \rangle \)-invariant finite subgroup has a layer-finite periodic part.

Then \( G \) is either a finite or an \( \Phi \)-group.

In a \( \Phi \)-group \( G \) all involutions are conjugate; the Sylow 2-subgroups are locally cyclic or finite generalized quaternion groups; there are infinitely many elements of finite order in \( G \), which are strictly real with respect to the involution \( i \) and for every such element \( c \) of this set there exists an element \( s_c \) from the centralizer of \( i \) such that \( \langle c, c^{s_c} \rangle \) is an infinite group.

Layer-finite groups are characterized in the class of locally solvable groups and groups with a layer-finite periodic part in more general case with \( \Phi \)-groups accuracy.

In the article [52], V.P. Shunkov bring up next question for discussion:

Do the classes of \( \Phi_0 \)-groups and \( T_0 \)-groups coincide or not?

In the same article V.P. Shunkov specially emphasized that the most difficult part of the problem is the establishing of satisfiability for a \( \Phi_0 \)-group of conditions (4) and (5) from the definition of a \( T_0 \)-group.

In [35] V.I. Senashov proved, that a \( \Phi_0 \)-group satisfies all conditions from the definition of a \( T_0 \)-group except for the fourth condition. In the same article an example of a \( \Phi_0 \)-group which is not a \( T_0 \)-group was constructed, i.e. it was shown that the fourth condition does not hold in every \( \Phi_0 \)-group.

**EXAMPLE OF A \( \Phi_0 \)-GROUP** (V.I. Senashov, [35]). Let’s take isomorphic copies of the \( T_0 \)-groups \( G = V\lambda(i) \) from [52]:

\[
G_1 = V_1\lambda(i_1), \; G_2 = V_1\lambda(i_1), \; \ldots, \; G_n = V_n\lambda(i_n), \; \ldots
\]

In the Cartesian product of the groups \( G_n, n = 1, 2, \ldots, \) consider the subgroup \( U = W\lambda(j) \), where \( W \) is the direct product of the subgroups \( V_n, n = 1, 2, \ldots, \) and \( j = i_1 \cdot i_2 \cdot \ldots \) is an involution from the Cartesian product of the \( G_n, n = 1, 2, \ldots \) Such a group \( U \) is a \( \Phi_0 \)-group. It is easy to see that fourth condition from the definition of a \( T_0 \)-group does not hold for the group \( U \).

**9. Almost layer-finite groups**

A group is said to be the *almost layer-finite* if it is a finite extension of layer-finite group.

To start with here are some theorems which describe almost layer-finite groups in the class of locally finite groups.
THEOREM 9.1 (V.P. Shunkov, [40]). A locally finite group $G$ is almost layer-finite if and only if in $G$ the following condition is valid: the normalizer of any non-trivial finite subgroup from $G$ is almost layer-finite.

Next, here is the theorem which characterizes almost layer-finite groups in the class of periodic groups without involutions.

THEOREM 9.2 (V.I. Senashov, [37]). Let $G$ be a conjugately biprimitively finite group without involutions. If in $G$ the normalizer of any non-trivial finite subgroup has an almost layer-finite periodic part, the group $G$ has an almost layer-finite periodic part.

The condition of conjugate biprimitive finiteness in this theorem should be taken into account in view of examples of the Novikov–Adian, [1], and Ol’shanskii groups, [21].

The next theorem describes almost layer-finite groups in the class of periodic groups with involutions.

THEOREM 9.3 (V.I. Senashov, [36,38]). Let $G$ be a periodic group of Shunkov type with strongly embedded subgroup. If in $G$ the normalizer of any non-trivial finite subgroup from $G$ is almost layer-finite, then the group $G$ is almost layer-finite.

Let’s recall that a subgroup $H$ of a group $G$ is called strongly embedded in $G$, if $H$ is a proper subgroup of $G$ and $H \cap x^{-1}Hx$ has odd order for all $x \in G \setminus H$.

We know the structure of the infinite Sylow 2-subgroups of the periodic non-almost layer-finite group of Shunkov:

THEOREM 9.4 (V.I. Senashov, [38]). Let $G$ be the periodic non-almost layer-finite group of Shunkov with almost layer-finite normalizers of any non-trivial finite subgroups. If the Sylow 2-subgroup of group $G$ is infinite, then it is a quasi-dihedral 2-subgroup.

Recall, that a quasi-dihedral group is an extension of a quasi-cyclic 2-group with the help of an inverting automorphism (this group received this name because it is an union of infinite number of finite dihedral 2-groups).

Using the known results about locally finite groups with Chernikov primary Sylow subgroups, we now obtain the following characterization of almost layer-finite groups in the class of locally soluble groups, which is an analog of one of the main results of [41].

THEOREM 9.5 (M.N. Ivko, [13]). Let $G$ be a periodic almost locally soluble group, possessing a Klein four-subgroup $L$. If the centralizer in $G$ of any involution from $L$ is layer-finite, then the group $G$ is almost layer-finite.

10. Periodic groups with minimality condition

A group $G$ satisfies the minimality condition for subgroups (Abelian subgroups), if in $G$ every decreasing chain of subgroups (Abelian subgroups) $H_1 > H_2 > \cdots$ stops at a finite number, i.e. $H_n = H_{n+1} = \cdots$ for some $n$. 

Let any decreasing chain of subgroups in the infinite locally finite group stop at a finite number. Will such a group be a finite extension of the direct product of the finite number of quasi-cyclic groups? This is the problem of minimality in the class of locally finite groups. In 1965 it became possible to reduce that problem to the case when the Sylow 2-subgroups in the group are finite. Those results were published in [42,43]. The problem of minimality in the class of locally finite groups was positively solved by V.P. Shunkov in [45]. Locally finite groups with the condition of minimality for Abelian subgroups were described in [47].

In 1968 P.S. Novikov and S.I. Adian published the solution of the famous Burnside problem. Moreover, the following theorem was proved: the free Burnside group $B(m, n)$, $m \geq 2$, of odd period $n \geq 4381$, is infinite, the centralizer of any non-trivial element is finite and is contained in a cyclic subgroup of order $n$ from $B(m, n)$. In particular, in such a group all Abelian subgroups are finite, the group also satisfies the condition of minimality for Abelian subgroups. Moreover, for any odd prime number $p$ and natural number $s$ with $p^s > 4381$ the free Burnside $p$-group $B(m, p^s)$ is infinite and every elementary Abelian $p$-subgroup is finite. At once the following question emerged: what can be said about the $2$-groups in which some maximal elementary Abelian subgroup is finite? In 1970, V.P. Shunkov obtained the answer to that question. In fact, he proved the following

**Theorem 10.1** (V.P. Shunkov, [46]). If some maximal elementary Abelian subgroup is finite in an infinite 2-group, then the group itself is a finite extension of the direct product of the finite number of quasi-cyclic groups.

Recall that a group $G$ satisfies the $p$-minimality condition (min-$p$ condition), if every descending chain of subgroups

$$H_1 > H_2 > \cdots > H_n > \cdots$$

that is such, that $H_n \setminus H_{n+1}$ contains $p$-elements, terminates at a finite number.

Earlier we already pay attention to the fact that the extension of a locally finite group by a locally finite group is locally finite, but at the same time this is incorrect for layer-finite groups. The periodic almost locally soluble groups with min-$p$ condition are the extensions of layer-finite groups by layer-finite groups. In fact, in this case the group $G$ has a complete part $A$, and $G/A$ is locally normal, and every element of $G$ is unpermutational elementwise only with finite number of Sylow $p$-subgroups of $A$. The quotient group $G/A$ is locally normal and obviously has finite Sylow $p$-subgroups, hence it is layer-finite group. The group $A$ is an Abelian group constructed from a quasi-cyclic group, and by the min-$p$ condition every $p \in \pi(A)$ in $A$ has only a finite number of quasi-cyclic groups. Hence $A$ is a layer-finite group too.

Let’s discuss some results for groups with minimality condition.

A positive solution of the minimality problem for locally finite groups is given by the next theorem:

**Theorem 10.2** (V.P. Shunkov, [45]). A locally finite group with minimality condition for subgroups is either finite or it is a finite extension of a direct product of a finite number of quasi-cyclic groups.
Theorem 10.3 (Ya.D. Polovitskii, [24,25]). Let $G$ be an almost locally soluble group with min-$p$ condition. Then it has a complete part $A$; moreover $G/A$ is locally normal, and every element of $G$ is unpermutational element-wise only with finite number of Sylow $p$-subgroups of $A$.

Theorem 10.4 (V.P. Shunkov, [49]). Let $G$ be a conjugately biprimitively finite group without involution with min-$p$ condition. Then $G$ has a complete part $R$ and the quotient-group $G/R$ is conjugate biprimitive finite with min-$p$ condition and with finite Sylow $p$-subgroups for every $p \in \pi(G)$.

Theorem 10.5 (V.P. Shunkov, A.K. Shlepkin, [49]). Every periodic conjugately biprimitively finite group without involutions with min-$p$ condition is locally finite.

Corollary 10.1. Every periodic conjugately biprimitively finite group without involutions with finite Sylow subgroups, satisfying the min-$p$ condition is locally finite.

Theorem 10.6 (E.I. Sedova, [27]). A periodic group is a locally soluble group with min-$p$ condition if and only if it is binary soluble and every one of its locally soluble subgroups satisfies the min-$p$ condition.

Theorem 10.7 (A.N. Ostylovskii, V.P. Shunkov, [22]). A conjugately biprimitively finite group without involutions with minimality condition is locally finite and is a solvable Chernikov group.

Theorem 10.8 (N.G. Suchkova, V.P. Shunkov, [60]). Every conjugately biprimitively finite group with minimality condition for Abelian subgroups is a Chernikov group.

Results on groups with minimality condition can be found in the monograph [39].

11. Groups with finitely embedded involution

We now go on to introduce the next concept introduced by V.P. Shunkov at the end of the 80ies.

Let $G$ be a group, $i$ one of its involutions and let $L_i = \{i^g \mid g \in G\}$ be the set of conjugated involutions from $G$ with $i$. We shall call the involution $i$ finitely embedded in $G$, if for any element $g$ from $G$ the intersection $(L_i L_i) \cap gC_G(i)$ is finite, where $L_i L_i = \{i^{g_1} = i^{g_2} \mid g_1, g_2 \in G\}$.

Let’s give the most simple examples of the groups with a finitely embedded involution.

1. If in the group $G$ there exists an involution $i$ with finite centralizer $C_G(i)$, then $i$ is a finitely embedded involution in $G$.

2. If in some group $G$ the involution $i$ is contained in finite normal subgroup from $G$, then $i$ is a finitely embedded involution in $G$.

3. Let $G$ be a Frobenius group with a periodic kernel and infinite complement $H$, containing an involution $i$. Then $i$ is a finitely embedded involution in $G$. 
(4) Let

\[ B_1, B_2, \ldots, B_n, \ldots \]

be an infinite sequence of finite groups, in which there is only a finite number of the groups of even order, and let \( B_n \lambda(i_n) \) be a subgroup from the holomorph \( \text{Hol}(B_n) \), where \( i_n \) is an involution, inducing in \( B_n \) an automorphism of order two \( \text{order}(n) = 1, 2, \ldots \). Let’s consider the group \( G = B\lambda(i) \), where

\[ B = B_1 \times B_2 \times \cdots \times B_n \times \cdots, \]

and \( i \) is the involution

\[ i = (i_1, i_2, \ldots, i_n, \ldots). \]

It is easy to show, that \( i \) is a finitely embedded involution.

An involution of a group is called a \textit{finite} involution, if it generates a finite subgroup with every involution, that is conjugate to it.

Now let us formulate some results, the main of which is the following.

**Theorem 11.1** (V.P. Shunkov, [50]). Let \( G \) be a group, and let \( i \) be a finite and finitely embedded involution in it, \( L_i = \{i^g \mid g \in G\} \), \( B = \langle L_i \rangle \), \( R = \langle L_i L_i \rangle \), and \( Z \) be the subgroup generated by all 2-elements from \( R \).

Then \( B, R, Z \) are normal subgroups in \( G \) and one of the next statements holds:

1. \( B \) is a finite subgroup;
2. the subgroup \( B \) is locally finite, \( B = R\lambda(i) \) and \( Z \) is a finite extension of a complete Abelian 2-subgroup \( A_2 \) with the condition of minimality, and \( ici = c^{-1} \) \( (c \in A_2) \).

A number of corollaries follow from this theorem.

**Corollary 11.1.** If a group has a finite involution with a finite centralizer, then it is locally finite.

**Corollary 11.2.** If a periodic group has an involution with a finite centralizer, then it is locally finite.

**Corollary 11.3.** If a finite finitely embedded involution exists in a group, then its closure is a periodic subgroup.

**Corollary 11.4.** A simple group with involutions is finite if and only if one of its involutions is finite and is finitely embedded.

As in a periodic group any involution is finite, the next results follow from Corollary 11.4.
COROLLARY 11.5. A periodic simple group with involutions is finite if and only if some involution in it is finitely embedded.

COROLLARY 11.6. Let $G$ be a group, let $H$ be a subgroup of it containing a finite involution, and let $(G, H)$ be a Frobenius pair. The group $G$ is a Frobenius group with a periodic Abelian kernel and with complement $H = C_G(i)$ if and only if $i$ is a finitely embedded involution in $G$.

Corollary 11.6 does not hold, even in a periodic group, if the involution $i$ is not finitely embedded.

12. Frobenius groups and finiteness conditions

When one study groups with finiteness conditions it is very helpful to use Frobenius group type properties.

A group of the form $G = F \lambda H$ is called a Frobenius group with kernel $F$ and complement $H$, if $H \cap H^g = 1$ for every $g \in G \setminus H$ and $F \setminus \{1\} = G \setminus \bigcup_{g \in G} H^g$, where $H$ is a proper subgroup.

Let $G$ be a group, $H$ subgroup. $G$ and $H$ are said to form a Frobenius pair, if $H \cap H^x = 1$ for every element $x \in G \setminus H$.

Let $G$ be a group, $H$ a subgroup of it, satisfying the following condition: for any $g \in G \setminus H$ the intersection $H \cap g^{-1} H_g = 1$. In this case we shall call the pair $(G, H)$ a Frobenius pair. If $G$ is a finite group, $G = F \lambda H$. This is a famous Frobenius theorem which plays a fundamental role in the theory of finite groups.

Let a Frobenius pair $(G, H)$ be given and $G = F \lambda H$. If $F \setminus \{1\} = G \setminus \bigcup_{g \in G} H^g$, then $G$ is called a Frobenius group with kernel $F$ and with complement $H$. According to the Adian theorem from the book “Burnside Problem and Identities in Groups” (Moscow: Science, 1975), in the group $B(m, n)$, $m \geq 2$, $n$ is an odd number and $n \geq 665$, each finite subgroup is contained in a cyclic subgroup of order $n$, making with the group $B(m, n)$ a Frobenius pair, and $B(m, n)$ is an infinite group. If we take a prime number $p > 665$ as $n$, then according to the well-known Kostrikin theorem, $B(m, p)$ has a finitely generated subgroup $H(p)$ of finite index in $B(m, p)$, not having subgroup of the finite index of its own.

The group $H(p)$ with its each cyclic subgroup of the prime order $p$ makes a Frobenius pair. However, it is not a Frobenius group with a non-invariant cyclic multiplier of prime order $p$. Let us give another interesting example.

Let $V = B(m, p)$, where $p$ is a prime number and $p > 665$. As is well known, $V$ has an automorphism $\phi$ of order 2 which takes all free generators into their inverse. In the holomorph $\text{Hol}(V)$ take the subgroup $G = V \lambda \langle i \rangle$, where $i$ is an involution, inducing the automorphism $\phi$ in $V$. If the centralizer $C_G(i)$ were finite, then according to statement 4, formulated above, the group $G$ would be finite, it couldn’t be possible. Hence, $H = C_G(i)$ is an infinite group. Further, according to the Adian theorem, formulated above, $(G, H)$ is a Frobenius pair, but $G$ is not a Frobenius group with the non-invariant infinite multiplier $H = C_G(i)$.
That is how the situation in periodic groups having Frobenius pairs stands. But, probably, the following V.P. Shunkov theorem can lead to further progress.

**Theorem 12.1** (V.P. Shunkov, [48,50]). Let $G$ be a group, $H$ a subgroup of it, $a$ an element of prime order $p \neq 2$ from $H$, satisfying the following condition: for every $g \in G \setminus H$, $(a, g^{-1}ag)$ is a Frobenius group with the complement $(a)$. Then

1. $H = T_{\lambda}N_G((a))$ and $K = T_{\lambda}(a)$ is either a Frobenius group with a complement $(a)$ and a kernel $T$, or $K = (a)$;
2. $F_a = T \cup N$ is a subgroup in $G$ and $G = F_aN_G((a))$, where $N$ is the set of all $p$-real elements from $G \setminus H$ relating the element $a$;
3. $E = T \setminus L$ is an invariant set in $G$, where $L$ is the set of all such elements from $T$, which is a $p$-real relating some element from $N$.

Finally we note a feature of unsimplicity for infinite groups.

**Theorem 12.2** (A.I. Sozutov, V.P. Shunkov, [58]). Suppose $G$ is a group, $H$ a proper subgroup, $a$ an element of order $p \neq 2$ in $G$ such that

(*) for almost all (i.e. except for perhaps a finite number) of elements of the form $g^{-1}ag$, where $g \in G \setminus H$, the subgroups $L_g = \langle a, g^{-1}ag \rangle$ are Frobenius groups with complement $(a)$.

Then either $G = F_{\lambda}N_G((a))$ and $F_{\lambda}(a)$ is a Frobenius group with complement $(a)$ and kernel $F$, or the index of $C_G(a)$ in $G$ is finite.

This feature of unsimplicity plays a very important role in the research of infinite groups with finiteness conditions.

**References**

Groups with finiteness conditions

Subject Index

(2, 3, 7)-generated, 415
(2, 3, 7)-generated group, 387, 397
(2, 3, 7)-generating triple, 391, 399, 408, 415
(2, 3, 7)-generators, 397
(2, 3, k)-generated, 410, 415
(2, 3, k)-triple, 403
(2, 3)-generated group, 387
2-cocycle, 192, 199
– deformation, 217
α-transformation, 440
Φ-group, 483–485
Φ-Sylow torus, 344
Φ-torus, 343
Φ0-group, 485
π-minimality, 478
σ-algebraically dependent, 280
σ*-linearly independent, 275
σ-field, 245
σ-algebra, 290
σ-algebraic, 250, 252, 280
– field extension, 250
σ-algebraically dependent, 250
σ-algebraically independent, 250, 280, 296
– indexing, 296
– set, 296
σ-automorphism, 247
σ-compositum, 249
σ-dimension, 271
σ-domain, 312
σ-endomorphism, 247
σ-factor ring, 247
σ-field extension, 245
σ-field of fractions, 250
σ-Galois group, 314
σ-generators of a σ-ideal, 246
σ-homomorphism, 247, 269
– of D-modules, 269
– of filtered σ*-R-modules, 274
σ-ideal, 245
σ-indeterminate, 250, 296
σ-invariant, 314
– subgroup, 314
σ-isomorphic, 295
σ-isomorphism, 247
σ-K-algebra, 290
σ-negative, 441
σ-normal, 317, 318
– extension, 317, 318
σ-operator, 268
σ-overfield, 245, 252
– of finite degree, 311
σ-overring, 245, 312
σ-place, 312
σ-positive, 441
σ-quotient ring, 247
σ-R-module, 268
σ-ring of fractions, 250
σ-specialization, 312
σ-stable, 314
– subgroup, 314
σ-subring, 245
– of periodic elements, 302
– of R generated by the set B over R0, 247
σ-subset, 249
σ-transcendence basis, 280, 282, 293
σ-transcendence degree, 281, 282, 290
σ-transcendental, 250
σ-type, 271, 282
σ-valuation, 312
– ring, 312
σ-vector space, 268
σ1-negative, 441
σ1-positive, 441
Σm-algebra structure, 454
σ*-algebra, 290
σ*-algebraically dependent, 280
σ*-algebraically independent, 251
σ*-dimension polynomial, 273, 275, 283, 285
– of a system of σ*-equations, 286
– of an ideal, 285
σ*-field, 245
σ*-field extension, 245, 247
σ*-field of residue classes, 291
σ*-generators, 247
σ*-ideal, 245, 286
σ*-indeterminates, 251, 256
σ*-K-algebra, 290
σ*-linearly dependent, 275
Subject Index

\[ \sigma^*-\text{overfield}, 245, 252 \]
\[ \sigma^*-\text{R-module}, 272 \]
\[ \sigma^*-\text{ring}, 245 \]
\[ \text{-- of fractions}, 250 \]
\[ \sigma^*-\text{subfield}, 245 \]
\[ \sigma^*-\text{subring}, 245 \]
\[ \sigma^*-\text{subset}, 249 \]
\[ \sigma^*-\text{type}, 283 \]
\[ \sigma^*-\text{vector space}, 291 \]
\[ \mathcal{A}\text{-leader}, 255 \]
\[ \text{abstract triangle group}, 388 \]
\[ \text{acyclic}, 14 \]
\[ \text{-- complex}, 23 \]
\[ \text{-- simplicial group}, 10 \]
\[ \text{-- simplicial ring}, 10 \]
\[ \text{additivity theorem}, 5, 24 \]
\[ \text{Adian}, S.I., 479, 487 \]
\[ \text{Adian theorem}, 490 \]
\[ \text{adjoint action}, 124, 187, 214, 221 \]
\[ \text{admissible epimorphism}, 15 \]
\[ \text{admissible monomorphism}, 15 \]
\[ \text{affine} \]
\[ \text{-- algebraic group}, 176 \]
\[ \text{-- algebraic variety}, 87 \]
\[ \text{-- group scheme}, 176, 226 \]
\[ \text{-- quotient}, 203 \]
\[ \text{-- reflection}, 374 \]
\[ \text{-- Weyl group}, 358, 374 \]
\[ \text{Alexander polynomial}, 368 \]
\[ \text{Alexander theorem}, 366 \]
\[ \text{algebra} \]
\[ \text{-- of } \sigma\text{-polynomials}, 250 \]
\[ \text{-- of } \sigma^*\text{-polynomials}, 251, 256 \]
\[ \text{-- of difference polynomials}, 250 \]
\[ \text{-- of generalized characters}, 190 \]
\[ \text{-- of inverse difference polynomials}, 251 \]
\[ \text{-- of lowering and raising operators}, 121 \]
\[ \text{-- of symmetric functions}, 340 \]
\[ \text{algebraic} \]
\[ \text{-- } \sigma\text{-overfield}, 306 \]
\[ \text{-- } \sigma^*\text{-equation}, 252, 286 \]
\[ \text{-- closure}, 40, 294, 299 \]
\[ \text{-- geometry}, 175 \]
\[ \text{-- group}, 194, 203, 361 \]
\[ \text{-- matrix group}, 316 \]
\[ \text{-- solution}, 253, 254 \]
\[ \text{-- of a set of } \sigma\text{-polynomials}, 253 \]
\[ \text{-- subset}, 325 \]
\[ \text{-- vector bundle}, 16 \]
\[ \text{algebraically independent}, 292 \]
\[ \text{algebraically irreducible } \sigma\text{-polynomial}, 299 \]
\[ \text{almost centralizing extension}, 87 \]
\[ \text{almost cocommutative}, 189 \]
\[ \text{almost every solution}, 305 \]
\[ \text{almost every specialization}, 305 \]
\[ \text{-- has a property } P, 305 \]
\[ \text{almost layer-finite}, 485 \]
\[ \text{almost locally soluble}, 475 \]
\[ \text{-- group}, 479, 488 \]
\[ \text{almost nilpotent group}, 483 \]
\[ \text{alternating group}, 391 \]
\[ \text{Alt}(\Omega), 415 \]
\[ \text{analytic isomorphism}, 30 \]
\[ \text{annuls}, 252 \]
\[ \text{anti-Chern character}, 6, 46 \]
\[ \text{antipode}, 175, 182, 196 \]
\[ \text{aperiodic } \sigma\text{-field}, 267 \]
\[ \text{arithmetic Fuchsian group}, 411 \]
\[ \text{Arnold, V.I., 429, 453, 454} \]
\[ \text{Artin, E., 365, 429, 435} \]
\[ \text{Artin braid group, 362, 368} \]
\[ \text{Artin group}, 430, 444 \]
\[ \text{-- of crystallographic type}, 368 \]
\[ \text{Artin presentation}, 436 \]
\[ \text{Artin–Brieskorn braid group}, 444 \]
\[ \text{Artin–Brieskorn group}, 444, 445 \]
\[ \text{Artinian } \sigma^*\text{-ring}, 273 \]
\[ \text{Artinian difference ring}, 269 \]
\[ \text{ascending chain condition}, 260 \]
\[ \text{ascending filtration}, 268 \]
\[ \text{associate graded } D\text{-module}, 269 \]
\[ \text{associativity isomorphism}, 209 \]
\[ \text{Auslander purity of the branch locus}, 341 \]
\[ \text{automatic group}, 445 \]
\[ \text{automorphism group}, 429 \]
\[ \text{-- of a free group}, 413, 434 \]
\[ \text{-- of an algebraic curve}, 387 \]
\[ \text{automorphism of permutation-conjugacy type}, 449 \]
\[ \text{autoreduced}, 277 \]
\[ \text{-- set}, 255 \]
\[ \text{-- of } \sigma^*\text{-polynomials}, 257 \]
\[ \text{-- of inverse difference polynomials}, 258 \]
\[ \text{-- of lowest rank}, 258 \]
\[ \text{-- of differential polynomials}, 256 \]
\[ \text{-- subset of lowest rank}, 256, 278 \]
\[ \text{averaging element}, 184 \]
\[ \text{averaging map}, 187 \]
\[ \text{b-invariant}, 369 \]
\[ \text{Babbitt decomposition}, 307 \]
\[ \text{bad prime}, 357 \]
\[ \text{Baer, R.}, 472 \]
\[ \text{Baez, J.}, 431 \]
\[ \text{Baez–Birman monoid}, 431, 450 \]
\[ \text{Baird, G.E.}, 111 \]
\[ \text{Bakhova, M.Yu.}, 469 \]
base, 439
basic invariants, 341
basic set, 244
basis, 260, 316
– for transatorial transcendence, 280
Bass, H., 5
Beilinson motivic cohomology, 41
benign \( \sigma \)-field extension, 306
Bernoulli numbers, 34, 51
Bernstein inequality, 79, 88
betweenness condition, 117, 127
bi-Galois object, 202, 210
biadditive, 66
bialgebra, 175, 182
– homomorphism, 182
– in a braided category, 218
– in a braided tensor category, 219
biexact, 66
big étale site, 39
big Zariski site, 39
Bigelow, S., 442
bireflection, 374
Birch, B., 52
birationality of the antipode, 219
binary finite, 469, 470
– group, 479
binary finiteness, 469
binary soluble, 474
biprimitive finiteness, 469, 471
biprimitively finite, 469
biproduct, 193, 220
Birch, B., 52
bireflection, 374
Birman, J.S., 431, 436, 441, 443, 450
Birman–Murakami–Wenzl algebra, 443
Bloch cycle group, 43
Bloch higher Chow group, 6, 43
Bloch–Kato conjecture, 5, 38, 42
Borsuk–Kan construction, 9
Borel, A., 44, 52
Borel regulator, 52
Borel subalgebra, 140, 147
Borel subgroup, 361
bosonization, 193, 220
Bott periodicity, 9
bounded chain complex, 21, 23, 58
bounded complex, 23
braid, 429, 431
– group, 211, 340, 362, 365, 368, 373, 374, 429
– in the solid torus, 430
– of a complex reflection group, 362
– of a group action, 444
– of a handlebody, 446
– of a solid torus, 446
– of a sphere, 446
– of the action of \( W \) on \( Y \), 444
– of the sphere, 430
– on \( n \) strings, 366
– relations, 448
– groups in 3-manifolds, 430
– groups in handlebodies, 430
– groups of Riemann surfaces, 430
– monoid, 365
– relation, 212
braid-permutation group, 431, 448
braided
– category, 210
– Hopf algebra, 188, 219
– monoidal category, 217
– tensor category, 210, 211, 214, 219
braiding, 190, 229
– structure, 219
branching rule, 111–113, 116, 117, 120, 142, 154, 159
Bratteli diagram, 113
Brauer character, 406, 409
Brauer group, 34, 37, 40, 217
Brauer–Long group, 217
Brieskorn, E., 430, 444, 445, 453, 454
Bruhat cell, 362
Bruhat decomposition, 351, 361
Bruhat order, 350
Bruhat–Chevalley order, 350, 362
B \( p \) type pattern, 162
Birnau, W., 437
Birnau relations, 437
Birrnu representation, 368, 429, 430, 442
Burnside group, 484
Burnside \( p \)-group, 487
Burnside problem, 487
C∗-algebra, 175, 197
cancellation law, 350
canonical basis, 112
canonical loop, 435
Capelli determinant, 126, 134, 135
Capelli-type determinants, 135
Cartan, E., 111
Cartan
– homomorphism, 58
– map, 5, 57, 69
– matrix, 229, 351, 352, 354, 356
– of affine type, 360
Subject Index

– subalgebra, 115, 116, 121, 124, 139, 159, 161
– type, 229
– braiding, 229
Cartesian power, 444
Casimir element, 130
Cassidy, P., 243
categorical dimension, 225
category
– of algebraic vector bundles, 16
– of finite-dimensional left \( A \)-modules, 208
– of finitely generated modules, 16
– of finitely generated projective modules, 16
– of finitely presented \( R \)-modules, 23
– of \( H \)-comodules, 209
– of locally free sheaves, 16
– of representations, 17
– of smooth quasi-projective varieties, 41
– of vector bundles, 17
– with cofibrations, 20
central multiplicative system, 27
central simple \( F \)-algebra, 37
centralizer algebra, 113
centralizer construction, 115, 153
Cetlin, 111
ch-group, 477
chain complex, 9
chamber, 445
character, 117
– of an order, 441
– ring, 190, 222
– table, 354
– theory, 197
characteristic identity, 112, 116, 130, 136
characteristic polynomial, 90, 270, 273
– in several variables, 276
characteristic set, 256, 258, 278
characters of finite reflection groups, 340
Charney, R., 445
Cherep, A.A., 469, 470
Chern character, 6, 46, 48
Chernikov, S.N., 471, 472
Chernikov
– group, 472, 477, 479, 488
– normal divisor, 477, 478
– periodic part, 481
– Sylow \( p \)-subgroup, 477
Chevalley, C., 341
Chevalley group, 360
Chevalley property, 225
CHEVIE, 371
Chinese remainder theorem, 413
Chow, W.-L., 440
Chow group, 6, 43
Clark–Ewing classification, 376
class constants, 397
class equation, 197
– for finite groups, 224
– for semisimple Hopf algebras, 223
class group, 6, 55, 60
class groups of integers, 60
classical
– algebraic \( K \)-theory, 5
– braid group, 363
dimension polynomial, 279
– group, 394
– Lie algebra, 111
– representation theory, 158
– semisimple complex Lie algebra, 360
classification
– of complex reflection groups, 341
– of Coxeter groups, 445
– of fibre bundles, 8
– of finite reflection groups over the quaternions, 374
– of reflection data, 347
– of simple Lie algebras, 445
– of the subgroups of \( \text{PSL}_2(q) \), 404
classifying space, 5, 7, 14, 376, 430
– of a group, 8
– of a small category, 8
Clebsch–Gordan coefficients, 131
Clifford-theory, 370
co-Morita equivalence, 210
co-Morita equivalent, 209
co-simplicial object, 7
co-simplicial space, 7
coadjoint action, 216
coadjoint orbit, 158
coalgebra, 177, 455
– map, 178
cosetautomorphism, 177
– up to conjugation, 230
coset identity, 175
cocommutative, 178
– element, 223
– Hopf algebra, 182
cocycle, 192, 202
– condition, 191
– cyclic subgroup, 59
codegrees, 343
coevaluation, 209
coexponents, 342
cofibration, 14, 20
– sequence, 20
– of exact functors, 25
cofibre of a Cartan map, 5
cofibre sequence of spectra, 31
cofinal, 18
cofinality condition, 19
Cohen, F.R., 430, 455, 458
Cohen, J., 387
coherent autoreduced set, 258
coherent set of difference polynomials, 258
coherent sheaf, 17
coherently associative, 18
coherently commutative, 18
Cohn, R.M., 243, 255
cohomology
  – of braids, 430
  – of groups, 37
  – of pure braid groups, 430
  – of the generalized braid groups, 455
  – theory on \( \mathcal{S}_\text{ch} \), 39
codeal, 180
  – subalgebra, 221
coinvariant, 186
  – algebra, 341
  – coalgebra, 179
combitorial diagram, 387
commutative dimension, 101
commutative superalgebra, 215
commutativity, 215
  – in a category, 215
commutator subgroup, 435
comodule, 180
  – algebra, 186, 201, 205
  – coalgebra, 186
comaodule-map, 181
compact complex, 23
compact Lie group, 376
  – action, 6
  – object, 23
compatible difference field extensions, 304
complement in a Frobenius group, 490
complement of the discriminant, 429
complete
  – \( \sigma \)-ideal, 262
  – Abelian subgroup, 479
  – part, 479
  – set of parameters, 298, 300
  – subgroup, 480
  – system of \( \sigma \)-overfields, 266
completely aperiodic, 302
complex
  – conjugation, 304
  – place, 35
  – reflection group, 342, 348, 363, 368, 374
composite root, 123
computation of Kolchin polynomials, 282
computer algebra, 243
comultiplication, 175, 177
Conder, M.D.E., 387, 413, 414, 416
Conder diagram, 394
condition of minimality, 489
  – for Abelian subgroups, 487
conditions of finiteness, 469
configuration space, 429, 430, 434, 443, 444, 452
  – of a manifold, 444
conflation, 15
conformal field theory, 226
conjugacy class, 353
conjugate biprimitive finiteness, 469, 471, 486
conjugately biprimitively finite, 469, 480
  – group, 479, 488
conjugately \( n \)-finite, 469
conjugation problem for braid groups, 365
connected coalgebra, 179
connection, 198
connective spectrum, 10
conservative system of ideals, 260
constant, 272
  – element, 246
  – containing a term, 277
contractible, 25
contragradient of the reflection representation, 343
contravariant inner product, 162
convolution invertible, 190
convolution product, 182
coproduct, 20, 130, 147, 455
  – formula, 150
cosetquasitriangular, 217
  – Hopf algebra, 188, 190, 210, 214
coradical, 179
  – filtration, 180, 228
core of a difference overfield, 306
corepresentation theory for Hopf algebras, 185
coroot, 357
correspondence, 41
cosemisimple, 222
  – coalgebra, 179
  – Hopf algebra, 182
  – tensor product, 181
cotorsion, 36
cotriangular Hopf algebra, 190
cotriple, 12
counit, 177
  – property, 177
covering, 38
covolume, 52
Coxeter
  – braid diagrams, 364
  – element, 369
Subject Index

- generator, 363
- graph, 349, 445
- group, 340, 349, 356, 364, 369, 430, 445
- matrix, 349, 352, 354, 356, 369, 445
- presentation, 356, 374
- relation, 363
- system, 349, 356
cross-product algebra, 59
crossed coproduct, 205
crossed product, 193, 198
crystal basis, 112
crystallographic
- complex reflection group, 374
- condition, 354
- group, 374
cup product, 35, 38
cyclic
- cohomology, 230
- multiplier, 490
- quotient, 59, 60
cyclotomic extension, 40
cyclotomic polynomial, 343
cylinder axiom, 31
cylinder functor, 31

D type pattern, 163
D-module, 268
decomposition numbers
- for GL_n(F_q), 373
- of C_n, 373
decorated rooted tree, 230
Dedekind, R., 51
Dedekind domain, 27, 45, 47, 55
Dedekind zeta function, 51
defining σ-ideal, 252
defining σ*-ideal, 252
defining difference ideal, 252
deflation, 15
deformation, 217
- of braids, 429
degeneracy map, 7
degree, 92, 292
- of a difference kernel, 292
- of a reflection group, 341
- of transcendence, 291
Dehornoy, P., 429, 442
Deligne, P., 445, 453
Deligne theorem, 225
Demazure, M., 375
dendriform Hopf algebra, 230
density theorem, 82
derivation, 102, 284
- ring, 87, 89
derived category, 5, 23

derived group, 413
descending chain of subgroups, 473
desuspension, 41
devissage theorem, 5, 25
diagonal Cartan subalgebra, 139
diagonalizable, 340
diagram automorphism, 365
Dickson classification
- of subgroups of SL_2(q), 412
- of the subgroups of PSL_2(q), 404
difference
- σ-polynomial, 270
- algebra, 243, 290
- algebraic equation, 252
- automorphism, 247
- dimension, 271
- polynomial, 270
- of the module associated with the p-dimensional filtration, 279
domain, 312
- endomorphism, 247
- equation, 286
- factor ring, 247
- field, 245
- extension, 245
- Galois group, 314, 315, 326, 329
- Galois theory, 329
- homomorphism, 247, 269, 316
- ideal, 245
- indeterminates, 250
- isomorphism, 247
- kernel, 39, 291
- of length r, 294
- operator, 268
- overring, 245
- place, 312
- quotient field, 313
- quotient ring, 247
- R-module, 268
- ring, 244
- specialization, 312
- subring, 245
- transcendence basis, 280
- type, 271, 282
- valuation, 312
- ring, 312
- vector space, 268
difference-differential dimension polynomial, 285
differential, 284
- algebra, 243
- dimension polynomial, 285
- equation with delay, 244
- field, 284
Subject Index

501

– form, 454
– Galois theory, 243
– operator, 80, 86
– polynomial, 253
dihedral group, 6, 58, 192, 361, 481, 482
dimension, 271
– of a \( \sigma \)-algebra, 290
– of a variety, 297
– of an ideal, 297
– polynomial, 285
– theory of differential rings, 243
discrete
– \( G \)-module, 36
– orbit category, 64
– torsion group, 40
– valuation ring, 27, 46
discrete-time nonlinear systems, 244
discriminant, 365, 429
– of a reflection group, 365
– distinguished grouplike element, 195
division algebra, 55
divisor map, 49
Doi–Hopf module, 214
domination, 312
double centralizer theorem for Lie color algebras, 221
double crossed product, 193
double crossproduct, 200, 216
Dress, A., 6, 57
Drinfeld, V.G., 158, 188
Drinfeld
– double, 190, 193, 213, 216
– element, 189, 226
– generator, 116, 136
– map, 189
dual
– bases, 209, 217
– characteristic identity, 132
– cocycle, 202
– Hopf 2-cocycle, 191
– Hopf algebra, 184
– partition, 370
– projection operator, 132
– pseudo-cocycle, 192
– representation, 398
– root system, 357
Dwyer, W., 46
Dwyer and Friedlander, 6
Dynkin diagram, 355
\( \mathcal{E} \)-subvariety, 264
\( \mathcal{E} \)-variety, 263
– over a \( \sigma \)-field, 266
effective order, 296
– of a \( \sigma \)-field extension, 296
– of a \( \sigma \)-polynomial, 296
– of a variety, 297
– of an ideal, 297
E I category, 6, 63
Eichler, M., 61
Eichler order, 6, 61
Eilenberg–MacLane space, 11, 363
Eilenberg–MacLane spectrum, 11
elementary matrix, 413
ever injectives, 37, 39
Enright–Varadarajan module, 137
enveloping algebra, 83, 89, 111, 175, 228
– of a Lie color algebra, 218
– of a Lie superalgebra, 218
epimorphic images of triangle groups, 391
equalizer diagram, 39
equation in finite differences, 286
equivalent
– \( \sigma \)-field extensions, 307
– filtrations, 82
– realizations, 295
– tuples, 252
equivariant
– higher algebraic \( K \)-theory, 6
– higher \( K \)-theory, 5, 65
– homotopy equivalence, 453
– Reidemeister torsion, 63
essential
– prime divisor, 261, 264, 291
– separated components, 265
– separated divisors, 263
– strongly separated divisors, 263
étale
– Chern character, 46, 51
– cohomology, 5, 6, 34, 38, 46
– group, 6
– hypercohomology, 39
– \( K \)-theory, 6
– mapping, 39
Euclidean
– plane, 389
– structure, 444
– triangle, 388
Euler totient function, 414
evaluation, 209
exact
– category, 5, 15, 65
– in the sense of Quillen, 64, 67
– connected sequence of functors, 24
– pairing, 66
– sequence, 66
– upper bound for the Gelfand–Kirillov
dimension, 104
Subject Index

– filtration, 269, 273
– p-dimensional, 277

exceptional
– complex reflection group, 372
– group, 348, 445
– Lie algebra, 111
  – G2, 136
– simple group of Lie type, 397
– Weyl group, 372

exchange condition, 350
excision, 32
– ideal, 33

exhaustively and separately filtered module, 269
existence theorem, 298

exponent, 226, 341, 343, 369
– of an irreducible character, 343
– of Hopf algebras, 226

Ext functor, 273

extended Mickelsson algebra, 122

extension, 193
– axiom, 21
– of the twisted Yangian, 138

exterior algebra, 44

extraordinary cohomology theory, 10

extremal projector, 112, 116, 121–123, 139, 159

extreme group, 477

F-group, 474
F4-diagram, 347
F4q-group, 474

Fq-group, 474

Fq∗-group, 474

Fq∗-condition, 473

Fq∗-group, 474

face map, 7

factorizable, 217

factorizable Hopf algebra, 190

factorization theorem for matrices, 404

Faddeev, L.D., 115, 158

Fadell, E., 430, 444

Franke, C., 321

free, 249
– A-module, 24
– K-module, 277
– Burnside group, 484, 487
– filtered σ∗-R-module, 274

finite
– complex reflection group, 368, 374
– Coxeter group, 356, 373
– Coxeter system, 349
– difference, 270
– filtration, 25
– fusion category, 231
– general linear group, 373
– generation of G(n), 29
– generation of K-groups, 44
– group of Lie type, 343
– groups generated by reflections, 339
– length, 93
– reductive group, 343
– simple group, 387
– tor-dimension, 24
– type Coxeter group, 349
– Weyl group, 361, 373

finite-dimensional Kac–Moody algebra, 360
finite-dual, 184

finitely
– approximate, 474
– embedded, 488
– involution, 488
– generated
  – σ-ideal, 246
  – σ-module, 268
  – σ-overring, 247
  – σ-ring extension, 247
  – σ∗-ideal, 246
  – σ∗-difference algebra, 290
  – σ∗-difference overring, 247
  – σ∗-difference ring extension, 247
  – inverse σ∗-overring, 247
  – inverse σ∗-ring extension, 247
  – inverse σ∗-difference overring, 247
  – inverse σ∗-difference ring extension, 247
– type, 290
– partitive, 85, 95
– algebra, 100
– K-algebra, 89

finiteness conditions, 469
finiteness results, 6

first filter inequality, 79
flasque category, 25
flat, 39
– A-module, 24
– flip, 178, 210
formal series, 122
Fox, R., 430, 444
Franke, C., 321

degree, 342, 369, 372

Fenn, R., 431, 448, 450
fiber, 211
– bundle, 8, 452

filtration, 364, 452
field of invariants, 375
filter dimension, 79, 81, 83, 87
– of a polynomial algebra, 85
filter inequality, 80, 97

filtered σ-module, 268
filtered σ-R-module, 268
filtration, 273

finite

– complex reflection group, 368, 374
<table>
<thead>
<tr>
<th>Subject Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>– group, 364, 413, 429, 434</td>
</tr>
<tr>
<td>– join, 294</td>
</tr>
<tr>
<td>– left $D$-module, 277</td>
</tr>
<tr>
<td>– simplicial ring, 10</td>
</tr>
<tr>
<td>Friedlander, E., 46</td>
</tr>
<tr>
<td>Frobenius, F.G., 400</td>
</tr>
<tr>
<td>Frobenius</td>
</tr>
<tr>
<td>– algebra, 195–197</td>
</tr>
<tr>
<td>– automorphism, 34</td>
</tr>
<tr>
<td>– formula, 408</td>
</tr>
<tr>
<td>– group, 481, 484, 488, 490, 491</td>
</tr>
<tr>
<td>– map, 343</td>
</tr>
<tr>
<td>– pair, 490</td>
</tr>
<tr>
<td>– property, 219</td>
</tr>
<tr>
<td>– theorem, 224, 490</td>
</tr>
<tr>
<td>– theory, 373</td>
</tr>
<tr>
<td>– type, 224, 226</td>
</tr>
<tr>
<td>FRT construction, 212, 213</td>
</tr>
<tr>
<td>Fuchsian group, 389, 390, 394</td>
</tr>
<tr>
<td>Fuks, D.B., 430, 455</td>
</tr>
<tr>
<td>full difference Galois group, 318</td>
</tr>
<tr>
<td>full Galois group, 320</td>
</tr>
<tr>
<td>full group, 317</td>
</tr>
<tr>
<td>functional equation, 244</td>
</tr>
<tr>
<td>fundamental</td>
</tr>
<tr>
<td>– domain, 390</td>
</tr>
<tr>
<td>– Galois theorem for PVE, 317</td>
</tr>
<tr>
<td>– group, 429</td>
</tr>
<tr>
<td>– matrix, 324</td>
</tr>
<tr>
<td>– representation, 158</td>
</tr>
<tr>
<td>– system of solutions, 316</td>
</tr>
<tr>
<td>– theorem</td>
</tr>
<tr>
<td>– for $G$-theory, 31</td>
</tr>
<tr>
<td>– for Hopf modules, 197, 219</td>
</tr>
<tr>
<td>– of higher $K$-theory, 30</td>
</tr>
<tr>
<td>fusion category, 177, 231</td>
</tr>
<tr>
<td>$G$-graded algebra, 204</td>
</tr>
<tr>
<td>$G$-space, 8</td>
</tr>
<tr>
<td>$G$-spectrum, 6</td>
</tr>
<tr>
<td>$G$-theory, 5, 31, 56, 58</td>
</tr>
<tr>
<td>$G$-torsor, 327</td>
</tr>
<tr>
<td>$G/G$ exact, 67</td>
</tr>
<tr>
<td>$G/H$-exact, 67</td>
</tr>
<tr>
<td>Gabriel, P., 79, 102</td>
</tr>
<tr>
<td>Galois</td>
</tr>
<tr>
<td>– closed, 317, 319</td>
</tr>
<tr>
<td>– field, 317</td>
</tr>
<tr>
<td>– subgroup, 317</td>
</tr>
<tr>
<td>– cohomology, 5, 34</td>
</tr>
<tr>
<td>– group, 40</td>
</tr>
<tr>
<td>– correspondence, 318</td>
</tr>
<tr>
<td>– for total Picard–Vessiot rings, 329</td>
</tr>
<tr>
<td>– extension, 203</td>
</tr>
<tr>
<td>– map, 201, 203</td>
</tr>
<tr>
<td>– object, 201, 204</td>
</tr>
<tr>
<td>– theorem for PVE, 317</td>
</tr>
<tr>
<td>– theory, 197, 201, 401</td>
</tr>
<tr>
<td>– of difference equations, 323</td>
</tr>
<tr>
<td>– type correspondence theory, 208</td>
</tr>
<tr>
<td>Garside, F.A., 429, 430, 440</td>
</tr>
<tr>
<td>Garside left normal form, 440</td>
</tr>
<tr>
<td>Garside right normal form, 440</td>
</tr>
<tr>
<td>gauge equivalence, 224</td>
</tr>
<tr>
<td>Gauss, C.-F., 429</td>
</tr>
<tr>
<td>Gelfand, I.M., 111</td>
</tr>
<tr>
<td>Gelfand–Kirillov dimension, 79, 81, 86, 97, 98, 102</td>
</tr>
<tr>
<td>Gelfand–Kirillov transcendence degree, 102</td>
</tr>
<tr>
<td>Gelfand–Tsetlin</td>
</tr>
<tr>
<td>– basis, 116, 118, 131, 134, 164</td>
</tr>
<tr>
<td>– formula, 111, 112</td>
</tr>
<tr>
<td>– integrable system, 158</td>
</tr>
<tr>
<td>– module, 137</td>
</tr>
<tr>
<td>– over the orthogonal Lie algebras, 164</td>
</tr>
<tr>
<td>– pattern, 114, 118, 136</td>
</tr>
<tr>
<td>– for the $B, C$ and $D$ types, 142</td>
</tr>
<tr>
<td>– subalgebra, 135–137</td>
</tr>
<tr>
<td>– type basis, 115, 136, 148</td>
</tr>
<tr>
<td>Gelfand–Weyl–Zetlin basis, 165</td>
</tr>
<tr>
<td>Gelfand–Zetlin module, 165</td>
</tr>
<tr>
<td>general linear group, 373</td>
</tr>
<tr>
<td>– over a finite field, 372</td>
</tr>
<tr>
<td>general linear Lie algebra, 111, 116</td>
</tr>
<tr>
<td>general polynomial, 429</td>
</tr>
<tr>
<td>general position, 147</td>
</tr>
<tr>
<td>generalized</td>
</tr>
<tr>
<td>– braid group, 430, 444</td>
</tr>
<tr>
<td>– Chernikov group, 479, 480</td>
</tr>
<tr>
<td>– Chernikov periodic part, 481</td>
</tr>
<tr>
<td>– cohomology theory, 10</td>
</tr>
<tr>
<td>– Liouvillian extension, 321</td>
</tr>
<tr>
<td>– Picard–Vessiot extension, 319</td>
</tr>
<tr>
<td>– quaternion algebra, 411</td>
</tr>
<tr>
<td>– quaternion group, 59, 485</td>
</tr>
<tr>
<td>generating subspace, 93, 94</td>
</tr>
<tr>
<td>generator of the monodromy around $H$, 362</td>
</tr>
<tr>
<td>generic</td>
</tr>
<tr>
<td>– prolongation, 291–295</td>
</tr>
<tr>
<td>– representation, 136</td>
</tr>
<tr>
<td>– solution field, 319, 321</td>
</tr>
<tr>
<td>– specialization, 293, 297</td>
</tr>
<tr>
<td>– zero, 252, 265, 291, 297</td>
</tr>
<tr>
<td>– of a variety, 265</td>
</tr>
<tr>
<td>– of an irreducible variety, 266</td>
</tr>
<tr>
<td>genus, 387, 430</td>
</tr>
<tr>
<td>– formula, 390, 397, 399</td>
</tr>
</tbody>
</table>
Subject Index

geometric $n$-simplex, 7
geometric realization, 7
geometrical model, 430
geometry of a root system, 356
Gersten, S., 5
Gersten conjecture, 5, 28, 55
Gersten–Swan $K$-theory, 5
Gerstenhaber–Schack cohomology, 222
Gillet and Waldhausen, 23
GLE (generalized Liouvillian extension), 321
global field, 6, 48
– of characteristic $p$, 53
– of characteristic 2, 53
– of characteristic $p$, 54
global section, 16
gluing axiom for weak equivalences, 20
Goldie theorem, 99
Golod $p$-group, 470
good basis, 112
good filtration, 93
Goryunov, V.V., 455
Gould construction, 116
GPVE, 319
graded
– $\sigma$-module, 268
– $\sigma$-$R$-module, 268
– difference $R$-module, 268
– Hopf algebra, 175, 218, 229
– module, 214
– multiplicity, 343
Green functor, 65, 68
Green–Deligne–Lusztig parameterization, 397
Greenspan number, 300
grid function, 286
Gröbner basis, 243
– method, 276
Grothendieck, A., 5, 176, 205
Grothendieck group, 9
Grothendieck ring, 223
ground $\sigma$-field, 263
group
– algebra, 175, 210, 228
– completion, 18
– of all diagonal matrices, 361
– of Burnside type, 484
– of elementary matrices, 12
– of grouplike elements, 228
– of isotopy classes of braids, 446
– of Lie type, 397
– of square free order, 58
– of totally positive units, 49
– with a $R$-$N$-pair, 360
– with unmixed factors, 471
group-graded algebra, 198
grouplike element, 178
groups of automorphisms of free groups, 429
Guillemin–Sternberg construction, 158
Guo, X., 6
$H$-commutative, 215
$H$-Galois extension, 201
$H$-Galois object, 201
$H$-invariant form, 227
$H$-measured algebra, 198
$H$-module algebra, 198
$H$-module map, 210
$H$-space, 17, 47
Hall–Janko group, 374
Hambleton–Taylor–Williams conjecture, 60
handle, 391, 415
handlebody, 430, 446
Harish-Chandra module, 112
Hecke algebra, 366, 372, 373
height of a root, 357
Heisenberg double, 217
Hennings invariant, 211
hereditary order, 6, 61
Herman anti-involution, 123
Herzog, E., 243
hidden symmetry, 114
higher
– Chow group, 43
– class group, 55, 60
– $G$-theory, 58
– $K$-group, 55
– $K$-theory, 5, 51
– of modules over ‘El’ categories, 6
– of $\mathbb{Z}$, 53
– regulator map, 52
– topological $K$-theory, 9
highest
– root, 374
– vector, 117, 120, 131, 139, 147, 154, 159
– weight, 111, 117, 120, 137, 139, 147, 154, 159
– irreducible representation, 137
– representation, 154
– vector, 117
Higman, G., 387, 391, 413
Higman diagram, 394
Hilbert
– Nullstellensatz, 265
– polynomial, 92, 98, 269, 282
– series, 341
– theorem, 90, 38
holomorph, 489
– of a group, 473
holomorphic differential form, 454
holonomic
– module, 91, 92, 94
<table>
<thead>
<tr>
<th>Subject Index</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>number, 79, 84, 92, 93</td>
<td>79</td>
</tr>
<tr>
<td>$R$-module, 93</td>
<td>84</td>
</tr>
<tr>
<td>HOMFLY-PT invariant, 365</td>
<td>92</td>
</tr>
<tr>
<td>HOMFLY-PT polynomial, 367</td>
<td>93</td>
</tr>
<tr>
<td>homological dimension, 24</td>
<td>93</td>
</tr>
<tr>
<td>homology</td>
<td>93</td>
</tr>
<tr>
<td>– of braid group, 455</td>
<td>93</td>
</tr>
<tr>
<td>– theory defined by a spectrum, 11</td>
<td>93</td>
</tr>
<tr>
<td>homomorphism of filtered $\sigma$-$R$-modules, 269</td>
<td>93</td>
</tr>
<tr>
<td>homomorphism of filtered difference modules, 269</td>
<td>93</td>
</tr>
<tr>
<td>homotopy</td>
<td>93</td>
</tr>
<tr>
<td>– associative, 18</td>
<td>93</td>
</tr>
<tr>
<td>– category, 23</td>
<td>93</td>
</tr>
<tr>
<td>– colimit, 9</td>
<td>93</td>
</tr>
<tr>
<td>– commutative, 18</td>
<td>93</td>
</tr>
<tr>
<td>– fibration, 14, 32</td>
<td>93</td>
</tr>
<tr>
<td>– fibre sequence, 31</td>
<td>93</td>
</tr>
<tr>
<td>– groups of a spectrum, 10</td>
<td>93</td>
</tr>
<tr>
<td>– invariant, 14, 43</td>
<td>93</td>
</tr>
<tr>
<td>hook, 371</td>
<td>93</td>
</tr>
<tr>
<td>Hopf, H., 175</td>
<td>93</td>
</tr>
<tr>
<td>Hopf</td>
<td>93</td>
</tr>
<tr>
<td>– 2-cocycle, 191</td>
<td>93</td>
</tr>
<tr>
<td>– algebra, 147, 175, 182, 455</td>
<td>93</td>
</tr>
<tr>
<td>– filtration, 229</td>
<td>93</td>
</tr>
<tr>
<td>– – homomorphism, 182</td>
<td>93</td>
</tr>
<tr>
<td>– – in a braided tensor category, 218, 219</td>
<td>93</td>
</tr>
<tr>
<td>– – in the category of $k\mathbb{Z}$-comodules, 218</td>
<td>93</td>
</tr>
<tr>
<td>– – of rooted tree, 176</td>
<td>93</td>
</tr>
<tr>
<td>– algebras in categories, 208</td>
<td>93</td>
</tr>
<tr>
<td>– algebroid, 230</td>
<td>93</td>
</tr>
<tr>
<td>– extension, 187</td>
<td>93</td>
</tr>
<tr>
<td>– Galois extension, 201</td>
<td>93</td>
</tr>
<tr>
<td>– ideal, 182</td>
<td>93</td>
</tr>
<tr>
<td>– module, 193, 197, 214</td>
<td>93</td>
</tr>
<tr>
<td>Hurewitz map, 12, 46</td>
<td>93</td>
</tr>
<tr>
<td>Hurwitz</td>
<td>93</td>
</tr>
<tr>
<td>– generation, 397, 402, 410</td>
<td>93</td>
</tr>
<tr>
<td>– group, 387, 410, 412, 413</td>
<td>93</td>
</tr>
<tr>
<td>– subgroup, 387, 390, 401, 405</td>
<td>93</td>
</tr>
<tr>
<td>– subgroups of $\operatorname{PSL}_n(\mathbb{F})$, 404</td>
<td>93</td>
</tr>
<tr>
<td>– upper bound, 390</td>
<td>93</td>
</tr>
<tr>
<td>hyperbolic</td>
<td>93</td>
</tr>
<tr>
<td>– geometry, 389</td>
<td>93</td>
</tr>
<tr>
<td>– plane, 389</td>
<td>93</td>
</tr>
<tr>
<td>– triangle, 388, 390</td>
<td>93</td>
</tr>
<tr>
<td>– – group, 389</td>
<td>93</td>
</tr>
<tr>
<td>hypercohomology, 39</td>
<td>93</td>
</tr>
<tr>
<td>– group, 42</td>
<td>93</td>
</tr>
<tr>
<td>hyperelementary, 68</td>
<td>93</td>
</tr>
<tr>
<td>hyperplane complement, 374</td>
<td>93</td>
</tr>
<tr>
<td>imprimitive reflection group, 357</td>
<td>93</td>
</tr>
<tr>
<td>imprimitive subgroup, 345</td>
<td>93</td>
</tr>
<tr>
<td>incompatible $\sigma$-field extensions, 304</td>
<td>93</td>
</tr>
<tr>
<td>incompatible difference field extensions, 304</td>
<td>93</td>
</tr>
<tr>
<td>indecomposable</td>
<td>93</td>
</tr>
<tr>
<td>– coalgebra, 179</td>
<td>93</td>
</tr>
<tr>
<td>– Coxeter matrix, 352</td>
<td>93</td>
</tr>
<tr>
<td>– group, 375</td>
<td>93</td>
</tr>
<tr>
<td>– ideal, 262</td>
<td>93</td>
</tr>
<tr>
<td>– reflection groups over finite fields, 375</td>
<td>93</td>
</tr>
<tr>
<td>indexing, 291</td>
<td>93</td>
</tr>
<tr>
<td>– of elements, 293</td>
<td>93</td>
</tr>
<tr>
<td>indicator system, 111</td>
<td>93</td>
</tr>
<tr>
<td>induced automorphisms of $\operatorname{Gal}(L/K)$, 314</td>
<td>93</td>
</tr>
<tr>
<td>induction homomorphism, 67</td>
<td>93</td>
</tr>
<tr>
<td>induction of modules, 203</td>
<td>93</td>
</tr>
<tr>
<td>infinite</td>
<td>93</td>
</tr>
<tr>
<td>– 2-group, 487</td>
<td>93</td>
</tr>
<tr>
<td>– group, 469</td>
<td>93</td>
</tr>
<tr>
<td>– – with finiteness conditions, 491</td>
<td>93</td>
</tr>
<tr>
<td>– loop space, 10, 23, 32</td>
<td>93</td>
</tr>
<tr>
<td>– matrices, 137</td>
<td>93</td>
</tr>
<tr>
<td>– simple $(2, 3, 7)$-generated group, 414</td>
<td>93</td>
</tr>
<tr>
<td>– symmetric group, 19</td>
<td>93</td>
</tr>
<tr>
<td>– Weyl group, 361</td>
<td>93</td>
</tr>
<tr>
<td>infinite-dimensional representations of $\mathfrak{gl}_n$, 137</td>
<td>93</td>
</tr>
<tr>
<td>infinitesimal</td>
<td>93</td>
</tr>
<tr>
<td>– braiding, 229</td>
<td>93</td>
</tr>
<tr>
<td>– Hopf algebra, 230</td>
<td>93</td>
</tr>
<tr>
<td>– lowering operator, 111</td>
<td>93</td>
</tr>
<tr>
<td>inflation, 15</td>
<td>93</td>
</tr>
<tr>
<td>initial, 257</td>
<td>93</td>
</tr>
<tr>
<td>– of a $\sigma$-polynomial, 254</td>
<td>93</td>
</tr>
<tr>
<td>– of $A$ with respect to $y_i$, 301</td>
<td>93</td>
</tr>
<tr>
<td>inner automorphism, 206, 207, 222</td>
<td>93</td>
</tr>
<tr>
<td>inner derivation, 102</td>
<td>93</td>
</tr>
<tr>
<td>inner measuring, 206</td>
<td>93</td>
</tr>
<tr>
<td>integrable system, 158</td>
<td>93</td>
</tr>
<tr>
<td>integral, 183, 196, 205</td>
<td>93</td>
</tr>
<tr>
<td>– Cartan matrix, 355</td>
<td>93</td>
</tr>
<tr>
<td>intermediate $\sigma^*$-field, 317</td>
<td>93</td>
</tr>
<tr>
<td>interpolation point, 155</td>
<td>93</td>
</tr>
<tr>
<td>interpolation polynomial, 126, 141, 152</td>
<td>93</td>
</tr>
<tr>
<td>interpolation properties, 157</td>
<td>93</td>
</tr>
<tr>
<td>invariant, 186, 205</td>
<td>93</td>
</tr>
<tr>
<td>– $\sigma$-ring, 302</td>
<td>93</td>
</tr>
<tr>
<td>– basis property, 18</td>
<td>93</td>
</tr>
<tr>
<td>– differential forms, 342</td>
<td>93</td>
</tr>
<tr>
<td>– lattice, 374</td>
<td>93</td>
</tr>
<tr>
<td>– of oriented knots and links, 367</td>
<td>93</td>
</tr>
<tr>
<td>– quadratic form, 347</td>
<td>93</td>
</tr>
<tr>
<td>– ring, 341</td>
<td>93</td>
</tr>
<tr>
<td>invariants of 3-manifolds, 212</td>
<td>93</td>
</tr>
<tr>
<td>Subject Index</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
</tr>
<tr>
<td>inverse</td>
<td>Ivanov, N.V., 430</td>
</tr>
<tr>
<td>– closure, 246</td>
<td>Ivko, M.N., 473, 475</td>
</tr>
<tr>
<td>– difference Galois problem, 330</td>
<td>j-induction, 369, 370</td>
</tr>
<tr>
<td>– Galois problem, 401</td>
<td>j-leader, 277</td>
</tr>
<tr>
<td>– limit degree, 287</td>
<td>Jacobi bound, 300</td>
</tr>
<tr>
<td>– problem, 330</td>
<td>Jacobi number, 300</td>
</tr>
<tr>
<td>– reduced limit degree, 288</td>
<td>Jacobson radical, 179</td>
</tr>
<tr>
<td>– scattering method, 115, 158</td>
<td>James, G., 373</td>
</tr>
<tr>
<td>inversive</td>
<td>Johnson, J., 243</td>
</tr>
<tr>
<td>– σ*-dimension, 275</td>
<td>joining representations, 391</td>
</tr>
<tr>
<td>– σ*-operator, 272</td>
<td>Jones, V.F.R., 429</td>
</tr>
<tr>
<td>– σ*-type, 275</td>
<td>Jones polynomial, 211, 365, 368</td>
</tr>
<tr>
<td>– closure, 248, 249</td>
<td>Joseph, A., 102</td>
</tr>
<tr>
<td>– of a σ-field, 249</td>
<td>(k, l, m)-generated, 388</td>
</tr>
<tr>
<td>– difference</td>
<td>(k, l, m)-generating triple, 388</td>
</tr>
<tr>
<td>– algebra, 290</td>
<td>(k, l, m)-triple, 388</td>
</tr>
<tr>
<td>– dimension, 275</td>
<td>k-affine, 205</td>
</tr>
<tr>
<td>– field, 245, 249</td>
<td>K-groups of number fields, 6</td>
</tr>
<tr>
<td>– field extension, 284</td>
<td>K-linear derivation, 284</td>
</tr>
<tr>
<td>– operator, 272</td>
<td>K-theory, 5</td>
</tr>
<tr>
<td>– R-module, 272</td>
<td>– of group rings, 55</td>
</tr>
<tr>
<td>– ring, 244</td>
<td>– of Karoubi and Villamayor, 13</td>
</tr>
<tr>
<td>– structure, 273</td>
<td>– of local fields, 46</td>
</tr>
<tr>
<td>– type, 275</td>
<td>– of maximal orders, 55</td>
</tr>
<tr>
<td>– vector space, 272, 284</td>
<td>– of orders, 55</td>
</tr>
<tr>
<td>inversive Ritt σ-ring, 261</td>
<td>– of Waldhausen categories, 31</td>
</tr>
<tr>
<td>involution, finitely embedded, 488</td>
<td>– of Z, 53</td>
</tr>
<tr>
<td>irrationality of characters, 373</td>
<td>– space, 21</td>
</tr>
<tr>
<td>irreducible</td>
<td>– of a W-category, 22</td>
</tr>
<tr>
<td>– character, 223</td>
<td>– spectrum, 23</td>
</tr>
<tr>
<td>– of a reflection group, 343</td>
<td>– with mod-r coefficients, 33</td>
</tr>
<tr>
<td>– coalgebra, 179</td>
<td>$K_0$, 5</td>
</tr>
<tr>
<td>– complex reflection group, 363</td>
<td>$K_1$, 5</td>
</tr>
<tr>
<td>– component, 179, 264</td>
<td>$K_2$, 5</td>
</tr>
<tr>
<td>– Coxeter system, 352</td>
<td>$K(\pi, 1)$-space, 363, 452</td>
</tr>
<tr>
<td>– E-components, 264</td>
<td>Kac, G.I., 197</td>
</tr>
<tr>
<td>– module, 223</td>
<td>Kac algebra, 205</td>
</tr>
<tr>
<td>– p-adic reflection groups, 376</td>
<td>Kac–Moody algebra, 359</td>
</tr>
<tr>
<td>– real reflection group, 352</td>
<td>Kac–Zhu theorem, 197, 224</td>
</tr>
<tr>
<td>– reflection group, 345</td>
<td>Kahn, B., 6, 46</td>
</tr>
<tr>
<td>– over finite field, 375</td>
<td>Kan complex, 7</td>
</tr>
<tr>
<td>– representation, 111</td>
<td>Kaplansky conjecture, 176, 196, 197, 221</td>
</tr>
<tr>
<td>– of a complex reflection group, 369</td>
<td>– 10th conjecture, 225</td>
</tr>
<tr>
<td>– subgroup, 402, 403</td>
<td>– 5th conjecture, 222</td>
</tr>
<tr>
<td>– variety, 297</td>
<td>– 6th conjecture, 224</td>
</tr>
<tr>
<td>irredundant, 264</td>
<td>Karoubi–Villamayor K-theory, 5, 11</td>
</tr>
<tr>
<td>– representation, 261</td>
<td>Kauffman invariant, 211</td>
</tr>
<tr>
<td>– of a variety, 266</td>
<td>Keating, M.E., 61</td>
</tr>
<tr>
<td>irregular prime, 54</td>
<td>Keigher, W., 243</td>
</tr>
<tr>
<td>isotropy class, 429</td>
<td>kernel of a Frobenius group, 490</td>
</tr>
<tr>
<td>isotropic polynomial subalgebra, 103</td>
<td>Keune, F., 5</td>
</tr>
<tr>
<td>isotropic subalgebra, 80, 102, 103</td>
<td></td>
</tr>
</tbody>
</table>
Keyman, E., 450
Killing, W.K.J., 111
Klein, F., 387, 429
Klein four group, 473
Klein quartic, 390
Knizhnik–Zamolodchikov equation, 231
knot, 366
– theory, 175
Ko, K.H., 436, 441
Kolchin, E., 243
Kolchin polynomial, 282
Kolchin theorem on differential dimension polynomials, 282
Kolster, G., 6
Kolster, M., 61
Kostrikin theorem, 490
Krammer, D., 430, 443
Kronecker product, 34
Krull, W., 79
Krull dimension, 80, 83, 85, 87
– of the Weyl algebra, 89
Krull theorem, 37
Krull topology, 314
Krull–Schmidt for projectives, 202
Krull–Schmidt theorem, 197
Kuku, A.O., 6, 57, 62
Kulikov, Vik.S., 458
Kummer, E., 412
Kummer sequence, 37
Kurdachenko, L.A., 476
Kurihara, M., 54

\ell\textsuperscript{-}complete profinite Abelian group, 62
\ell\textsuperscript{-}completeness, 6
Lagrange interpolation formula, 134, 149, 155
Lagrange theorem, 196
Lang, S., 403
Laubenbacher, R.C., 6, 56, 61
Laurent polynomial, 367, 442
Lawrence, R., 443
Lawrence–Krammer representation, 442
layer-Chernikov, 479
– group, 477
layer-extreme group, 477
layer-finite, 472, 484, 488
– group, 472, 479, 485
– periodic part, 484
LB-group, 476
LE (Liouvillian extension), 320, 321
leader, 254, 257
– of a differential polynomial, 254
leading coefficient, 93
– of a Hilbert polynomial, 92
Lee, R., 45
Lee, S.J., 436, 441
left
– adjoint action, 187
– coadjoint action, 216
– coideal, 147
– distinguished grouplike, 195
– filter dimension, 81, 83, 87
– H-Galois object, 202
– integral for H, 184
– integral in H, 184
– modular element, 195
– ordered group, 441
– return function, 81
Leibniz rule, 102
length, 365, 454
– function, 350, 353
– of a difference kernel, 293
– of a finitely generated module, 269
Levi, H., 243
Levi sub-reflection datum, 343
Levi subgroup, 343, 415
lexicographic order, 254, 276
Lichtenbaum, S., 43
Lichtenbaum cohomology, 5, 42
Lichtenbaum conjecture, 6, 52, 54
Lie algebra, 359
– of primitive elements, 228
– of type B, 137
– of type C, 137
– of type D, 137
Lie color algebra, 190, 218
Lie group action, 6, 65
Lie superalgebra, 137, 218
Lie–Kolchin theorem, 361
lifting method, 229
lifting problem, 229
lifting theorem, 222
limit basis of transformal transcendence, 289
limit degree, 287
– of a variety, 297
– of an ideal, 297
limit transformal transcendence degree, 288
linear
– \sigma\textsuperscript{-}ideal, 258
– \sigma\textsuperscript{-}polynomial, 258
– \sigma\textsuperscript{-}\ast\textsuperscript{-}ideal, 258
– algebraic group, 361
– differential equation, 401
– differential polynomial, 258
– homogeneous difference equation, 315
linearity of braid groups, 430
linearly
– disjoint, 249, 316
– left ordered group, 441
– right ordered group, 441
– rigid, 401
– triple, 388
link, 366
– invariant, 367
Liouvillian extension (LE), 320, 321
– generalized (GLE), 321
local
– σ-ring, 312
– σ*-K-algebra, 290
– difference ring, 312
– difference subring, 313
– field, 6, 46
– finiteness, 469
– systems on the sphere, 401
local-finiteness, 181
localization sequence, 5, 30, 45, 47
localization theorem, 26
locally
– Chernikov, 479
– group, 477
– compact group, 205
– cyclic, 484, 485
– finite, 469
– group, 479, 485, 487
– LB-group, 476
– normal divisor, 471
– QLF-group, 476
– subgroup, 471
– free sheaf, 16
– normal, 472, 479
– group, 472, 480
– LB-group, 476
– soluble group, 484
– with min-p condition, 474
– soluble subgroup, 472
– trivial fibration, 364
Loday, J.L., 15
Long, D.D., 430, 442
longest element, 123
loop space, 376
lower rank, 278
lowering and raising operators, 138
lowering operator, 111, 112, 114, 116, 120, 125, 131, 132, 154, 159, 160
– for the symplectic Lie algebras, 158
lowering weights, 140
Lucchini, A., 387, 391
m-basis, 260
Macbeath, A.M., 387, 396, 409
Macbeath theorem, 388, 407, 410, 412
Macdonald–Lusztig–Spaltenstein induction, 369
Mackey functor, 17, 66, 69
Macñahan, C., 411
Magnus, W., 429
Makanin, G.S., 430
Malle, G., 397
map of localization sequences, 47
map of spectra, 10
mapping class group, 429
Markov, A.A., 366, 437
Markov relations, 366
Martindale ring of fractions, 207
Maschke theorem, 195, 221
matched pair, 202
matrix
– coalgebra, 178
– element, 157
– formula, 116, 164
– formulas of Gelfand and Tsetlin, 164
– elements of representations, 155
Matsumoto theorem, 35, 350, 367
maximal
– σ-ideal, 245
– σ-ring, 312
– σ*-ideal, 245
– difference ideal, 245
– difference ring, 312
– order, 60
– R-order, 55
– torus, 361
Mayer–Vietoris sequence, 5, 33
McCarthy, J., 430
McConnell, J.C., 85
measured algebra, 198
measuring, 206
Merkurjev, A.S., 38
Merkurjev–Suslin theorem, 5
meromorphic, 245
– form, 454
– function, 51
metacyclic, 60
method of characteristic sets, 276
Mickelsson algebra, 112, 116, 118, 121, 124, 125, 135, 158, 160, 161
Mickelsson–Zhelobenko algebra, 116, 122–124, 135, 138, 139, 141, 151, 156, 159, 161
Milnor, J., 5
Milnor conjecture, 5, 38
Milnor K-theory, 5, 11, 14, 35
min-p condition, 473, 487, 488
minimal
– σ-dimension polynomial, 284
– σ*-dimension polynomial, 284
– τ-invariant reduced set, 325
– parabolic subgroup, 365
– standard generator, 306
– twist, 225
minimality condition, 472
– for subgroups, 486
– on all Abelian subgroups, 479
– on all subgroups, 479
minimality problem for locally finite groups, 487
– of a term, 277
– realization, 301
– solution, 301
multiplication constants, 396
multiplicative
– σ-subset, 249
– σ*-subset, 249
– form, 157
– system, 27
– for the filter dimension, 91
– formula, 136
– space, 151
multiplicity-free, 113, 114, 137, 158
multiplier Hopf algebra, 230, 231
multivariable
– characteristic polynomial, 279
– differential dimension polynomial, 279
– dimension polynomial, 285
– Hilbert polynomial, 279
Murakami, J., 443
Murnaghan–Nakayama rule, 371

N-graded algebra, 206
narrow Picard group, 49
natural basis, 115
natural characteristic, 404
natural permutation representation, 362
naturally associative, 67
naturally commutative, 67
negative root, 356
nerve of a group, 8
nerve of a small category, 7
Neuwirth, L., 430, 444, 452, 453
Nichols algebra, 219, 229
Nichols–Zoeller theorem, 193, 196, 220
nilpotent endomorphism, 24, 30
nilpotent group, 58
node of order i, 286
Noether–Skolem theorem, 206
Noether-type theorem for coactions, 206
Noetherian inverse difference ring, 270, 272, 274
Noetherian perfect conservative system, 260
Noetherian scheme, 17
non-commutative Galois theory, 197
non-commutative geometry, 230
non-commutative symmetric function, 136
non-crystallographic group, 374
non-soluble, 403
normal
– autoreduced set, 277
– basis property, 202
– closure, 306
– of the σ-field G over F, 306

– module algebra, 186, 198
– module coalgebra, 186
– Molien formula, 341
– monadic σ-field extension, 310
– monadic extension, 310
– monadicity, 315
– monodromy, 362
– monoidal category, 5, 209
– monoidally co-Morita equivalent, 209
– monomial basis, 112
– monomial matrix, 360
– Moody, J.A., 430, 442
– Moore space, 17, 46
– Moore theorem, 46
– Morita context, 199
– Morita equivalent, 204
– Morita map, 199, 204
– Morita–Takeuchi equivalent, 209
– Morita-equivalent, 199
– morphism of difference rings, 247
– Morrison, S., 243
– motivic Chern character, 6, 46, 51
– motivic cohomology, 5, 6, 34, 41
– Mukhamedjan, Kh.Kh., 472
– multicategory, 16
– multiple
Subject Index

- form, 437
- – of elements in the braid group, 430
- Hopf subalgebra, 187
- ordering, 123, 125, 160, 161
- subgroup, 187
normalized generator, 125, 140
Novikov, P.S., 479, 487
Novikov–Adian group, 472, 480
null homotopic, 25
Nullstellensatz for difference fields, 265
number field, 6, 33, 49
number of normal orderings, 123

obC-set, 64
Ol’shanskii, A.Yu., 479
Ol’shanskii group, 480
Ol’shanskii monster, 484
Ω-spectrum, 11
operad of rooted trees, 230
opposite algebra, 199
opposite smash product, 205
orbit category, 63
order, 256, 271, 296
- filtration, 86
- in an algebra, 29
- of a σ-field extension, 296
- of a σ-operator, 268
- of a σ-polynomial, 296
- of a σ*-operator, 272
- of a difference kernel with respect to an indexing, 292
- of a term, 254, 256
- of a variety, 297
- of an ideal, 297
- of unipotency, 227
ordered group, 441
ordered monomial, 114
orderly ranking, 254
- of σ*-indeterminates, 257
- of a set of σ*-indeterminates, 257
ordinary
- σ-ring, 244
- character, 409
- difference field, 245, 261, 296
- extension, 306
- difference ring, 244
- inversive difference field, 291
Ore extension, 104
Ore ring, 272
Ore subset, 100
oriented n-link, 366
orthant, 256
orthogonal

- group, 401, 407
- Lie algebra, 111, 112, 116, 158, 164
- twisted Yangian, 150
orthogonality property, 116
orthogonality relations, 223
Ostrovskii, A.N., 479, 488

p-adic
- field, 6, 27
- group, 361
- integers, 377
- reflection group, 375
p-biprimively finite, 469
p-center, 479
p-compact group, 376
p-complete topological space, 376
p-completion, 376
p-dimensional filtration, 276
p-minimality condition, 473, 487
p-real element, 491
p-real relating, 491
p-regular partition, 373
pairing of exact categories, 65
parabolic subgroup, 342, 347, 370
partial
- σ-ring, 244
- difference field, 288, 294
- difference kernels, 295
- difference ring, 244
partially ordered group, 312
partition, 370
partitive algebra, 86, 89, 100
pathological σ-field extension, 311
pathological difference field extension, 311
Paton, M., 430, 442
pattern, 134
pattern calculus, 136
perfect, 397, 403
- σ-ideal, 259
- closure, 260
- complex, 5, 23
- difference ideal, 260, 261
- group, 397, 404
periodic
- almost locally soluble group, 473
- binary soluble group, 473
- difference ring, 302
- element, 302
- group, 469, 472–474
- part, 474, 481
permitted difference field, 267
permitted difference ring, 267
permitted function, 267
permutation conjugacy, 449
permutation representations of $T(2, 3, 7)$, 415
– group, 347
– reflection group, 346
principal component, 298, 299
principal realization, 293, 295
principal series unipotent character, 343
problem of minimality, 487
product order, 278
profinite
– $G$-theory, 62
– group, 36, 37, 40
– of orders, 6
– higher $K$-theory, 5, 62
– $K$-theory, 17, 62
pointwise stabiliser, 342
projective
– curve, 45
– dimension, 29
– normalization, 347
– representation, 225
– of $T(2, 3, 7)$, 405
– unitary groups over rings of algebraic integers, 410
prolongation, 291, 294
– of difference kernel, 291
proper
– $\mathcal{E}$-subvariety, 264
– monadic algebraic difference extensions, 317
– specialization, 293, 295
– transform, 256
properly monadic, 310
property $L^*$, 294
property $\mathcal{P}$, 294
property $\mathcal{P}^*$, 294
pseudo-cocycle, 192
pseudo-reflection, 340
pseudo-twist, 223
pseudo-unitary group, 137
$\text{PSL}_2(13)$, 416
$\text{PSL}_2(7)$, 390
pure braid, 432
– group, 362, 364, 429, 434, 437, 454
– of the action of $W$ on $Y$, 444
purely inseparable, 294
purely transcendental, 375
purity of the branch locus, 341
Pushin, O., 6, 51
pushout, 20
PVE, 316
PVR, 324
$Q$–L conjecture, 48
$q$-Schur algebra, 373
$q$-special basis, 112
$qLE$, 322
Subject Index

QLF-group, 476
QSE-group, 477
quantized coordinate algebra, 194
quantized enveloping algebra, 115, 136, 228
quantum
– affine algebra, 137
– algebra, 112, 136
– commutative, 215
– determinant, 126, 148
– dimension, 211
– double, 226
– field theory, 175, 212, 231
– Gelfand–Tsetlin module, 137
– group, 115, 136, 176, 208, 228
– groupoid, 231
– Lie superalgebra, 137
– Mickelsson–Zhelobenko algebra, 136
– minor, 112, 133, 134, 164
– formula, 116
– plane, 215
– Sylvester theorem, 136
– Yang–Baxter equation, 176, 208
quasi-cyclic 2-group, 486
quasi-cyclic group, 487
quasi-cyclic p-group, 480
quasi-determinant, 136
quasi-dihedral 2-subgroup, 486
quasi-dihedral group, 486
quasi-exponent, 227, 230
quasi-Hopf algebra, 137
– algebra, 112, 136
– commutative, 215
– determinant, 126, 148
– dimension, 211
– double, 226
– field theory, 175, 212, 231
– Gelfand–Tsetlin module, 137
– group, 115, 136, 176, 208, 228
– groupoid, 231
– Lie superalgebra, 137
– Mickelsson–Zhelobenko algebra, 136
– minor, 112, 133, 134, 164
– formula, 116
– plane, 215
– Sylvester theorem, 136
– Yang–Baxter equation, 176, 208
quasi-linearly disjoint, 249, 294
quasitriangular, 213, 217, 220
quasitriangular Hopf algebra, 188, 210, 214
quaternion, 374, 411
– algebra, 410, 412
– F-algebra, 6, 61
– group, 6, 58, 192, 481
– norm, 411
quaternionic reflection group, 374
Quillen, D., 5, 44, 64
Quillen K-theory, 5, 11
Quillen lemma, 101
Quillen plus construction, 11
Quillen–Lichtenbaum conjecture, 6, 48
quotient
– σ-field, 250, 252
– difference field, 250
– division ring, 97
– field, 250
R-form, 192
R-matrix, 136, 191, 226
R-order, 29, 55
– in a semisimple F-algebra, 55
Radford, D.E., 190
radical, 262
– of a difference ideal, 248
raising operator, 112, 116, 120, 125, 131, 132
raising weights, 140
rank, 278
ranking, 254, 256
rational
– cohomology, 44
– equivalence, 43
– representation, 370
rationally complete, 208
Raudenbush, H., 243
RC module, 63
real
– embedding, 49
– number field, 36
– place, 35
– reflection group, 340, 343, 347, 349, 363
realization, 293, 295
– of a difference kernel, 293, 295
– of a symmetrizable Cartan matrix, 359
reduced
– σ-polynomial, 255
– σ-ring, 248
– σ*-polynomial, 257
– algebraic subset, 325
– decomposition, 123
– degree, 292
– of a difference kernel, 292
– element, 277
– expression, 350
– limit degree, 287
– subset, 325
– with respect to Σ, 255
reducible E-variety, 264
reducible subgroup, 403
reducible variety, 264
reduction
– modulo an autoreduced set, 255
– σ2n+1 ♦ σ2n−1, 159
– σ2n ♦ σ2n−1, 160
– of a σ*-polynomial with respect to an
  autoreduced set, 257
– problem, 112
– process, 255
– theorem, 255
reductive algebraic group, 361
reductive group, 343
Ree group, 397
reflecting hyperplane, 340, 363
reflection, 340, 444
  – datum, 343
  – group, 340, 365, 368, 374
  – over field of positive characteristic, 374
  – over the quaternions, 374
  – representation, 343, 368, 369, 376
  – subgroup, 343, 346
reflexive
  – σ-ideal, 245, 246
  – closure, 246, 262
  – prime σ-ideal, 298
  – prime ideal, 259
regular
  – element under a group action, 443
  – function, 176
  – Noetherian ring, 24
  – number, 344
  – orbit, 444
  – prime, 54
  – realization, 293, 295
  – of a kernel, 300
  – vector, 344
regulator map, 52
Reidemeister, K., 429
Reidemeister move, 433, 449
Reidemeister torsion, 63
relative
  – dimension, 297
  – effective order, 297
  – limit degree, 297
  – order, 297
relatively $H'$-projective, 68
remainder, 255
  – of a $\sigma$-polynomial with respect to another $\sigma$-polynomial, 255
renormalization in quantum field theory, 230
Rentschler, R., 79, 102
Rentschler–Gabriel–Krull dimension, 83
replicability, 305
  – of a $\sigma$-field extension, 305
representation category, 209
representation theory
  – for Hopf algebras, 185
  – of finite complex reflection groups, 368
  – of quantum affine algebras, 137
  – of simple Lie algebras, 112
  – of twisted Yangians, 146
  – of Yangians, 137
representations
  – of Hecke algebras, 373
  – of $\sigma y$, 158
  – of $T(2, 3, 7)$, 388
  – of the Yangian $Y(2)$, 148
  – of twisted Yangians, 158
  – of Yangians, 158
representative function, 183, 212
Resco, R., 102
residually finite-dimensional, 185, 205
residue homomorphism, 412
resolution of singularities, 42
resolution theorem, 5, 24
resolutions by free filtered difference modules, 276
resolvent, 303
  – ideal, 303
restricted Lie algebra, 207
restriction homomorphism, 67
return function, 81, 84, 94
ribbon category, 211
ribbon Hopf algebra, 190, 211
Riemann surface, 387, 390
Riemann zeta function, 51
right
  – adjoint action, 187
  – coadjoint action, 216
  – derived functor, 37, 273
  – Galois, 201
  – $H$-Galois, 201
  – extension, 215
  – integral, 184
  – $L$-Galois object, 202
  – Miyashita–Ulbrich action, 202
  – modular element, 195
  – ordered group, 441
  – smash product, 205
rigid braided tensor category, 211
rigid tensor category, 209, 211
rigid triple, 402
rigidity, 396
Rimányi, R., 431, 448
ring
  – of $\sigma$-operators, 268
  – of constants, 246
  – of difference operators, 268
  – of differential operators, 80, 99, 101
  – of invariants, 374, 375
  – of inverese $\sigma^*$-operators, 272
  – of inverese difference operators, 272
  – of partial difference-differential polynomials, 255
  – of $S$-integers, 45
ring-theoretical difference Galois theory, 330
Ritt, J.F., 243
Ritt $\sigma$-ring, 260
Ritt $\sigma^*$-ring, 261
Ritt difference ring, 260
Ritt number, 300
<table>
<thead>
<tr>
<th>Subject Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ritt–Raudenbush theorem, 260</td>
</tr>
<tr>
<td>Robson, J.C., 85</td>
</tr>
<tr>
<td>root, 340</td>
</tr>
<tr>
<td>-- for a reflection, 340</td>
</tr>
<tr>
<td>-- lattice, 375</td>
</tr>
<tr>
<td>-- system, 111, 115, 356, 374</td>
</tr>
<tr>
<td>-- of a Kac–Moody algebra, 359</td>
</tr>
<tr>
<td>-- vector, 121</td>
</tr>
<tr>
<td>Rosenfeld, A., 243</td>
</tr>
<tr>
<td>Rost, 43</td>
</tr>
<tr>
<td>round structure, 63</td>
</tr>
<tr>
<td>Rourke, C., 431, 448, 450</td>
</tr>
<tr>
<td>S- Excision ideal, 33</td>
</tr>
<tr>
<td>S-Finite, 469</td>
</tr>
<tr>
<td>S-Torsion, 29</td>
</tr>
<tr>
<td>-- object, 30</td>
</tr>
<tr>
<td>$\delta_4$ formula, 220</td>
</tr>
<tr>
<td>Saito, K., 444, 445</td>
</tr>
<tr>
<td>saturated W-category, 20</td>
</tr>
<tr>
<td>scheme, 16</td>
</tr>
<tr>
<td>Schur, I., 117</td>
</tr>
<tr>
<td>Schur algebra, 373</td>
</tr>
<tr>
<td>-- double centralizer theorem, 221</td>
</tr>
<tr>
<td>-- function, 117</td>
</tr>
<tr>
<td>-- indicator, 226, 227</td>
</tr>
<tr>
<td>-- lemma, 400, 402</td>
</tr>
<tr>
<td>-- polynomial, 117</td>
</tr>
<tr>
<td>Scott, G.P., 446</td>
</tr>
<tr>
<td>Scott, L.L., 398</td>
</tr>
<tr>
<td>Scott formula, 397, 398, 408</td>
</tr>
<tr>
<td>second filter inequality, 80, 85, 97</td>
</tr>
<tr>
<td>Sedova, E.I., 474, 488</td>
</tr>
<tr>
<td>Seidenberg, A., 243</td>
</tr>
<tr>
<td>semi-simple $F$-algebra, 28</td>
</tr>
<tr>
<td>semisimple element of GL, 340</td>
</tr>
<tr>
<td>semisimple Hopf algebra, 182, 221</td>
</tr>
<tr>
<td>semisimple triangular Hopf algebra, 224</td>
</tr>
<tr>
<td>semisolvable, 224</td>
</tr>
<tr>
<td>semistandard tableaux, 117</td>
</tr>
<tr>
<td>separable algebra, 221</td>
</tr>
<tr>
<td>separants, 301</td>
</tr>
<tr>
<td>separated ideals, 261</td>
</tr>
<tr>
<td>separated in pairs, 261</td>
</tr>
<tr>
<td>separated varieties, 265</td>
</tr>
<tr>
<td>Sergiescu, V., 436</td>
</tr>
<tr>
<td>Serre subcategory, 26</td>
</tr>
<tr>
<td>set of $\sigma$-generators, 246</td>
</tr>
<tr>
<td>set of constants, 272</td>
</tr>
<tr>
<td>set of parameters, 298</td>
</tr>
<tr>
<td>shape of a Young tableau, 113</td>
</tr>
<tr>
<td>sheaf, 39</td>
</tr>
<tr>
<td>-- on a site, 39</td>
</tr>
<tr>
<td>Shephard, G.C., 341</td>
</tr>
<tr>
<td>Shephard group, 364</td>
</tr>
<tr>
<td>Shephard–Todd, 376</td>
</tr>
<tr>
<td>-- classification, 345, 347, 368</td>
</tr>
<tr>
<td>-- for reflection groups, 346</td>
</tr>
<tr>
<td>Shimakawa, K., 6</td>
</tr>
<tr>
<td>Shlepkin, A.K., 488</td>
</tr>
<tr>
<td>shuffling, 260</td>
</tr>
<tr>
<td>Shunkov, V.P., 469, 475, 479, 481, 484, 487–489, 491</td>
</tr>
<tr>
<td>Shunkov group, 469</td>
</tr>
<tr>
<td>Shunkov problem, 470</td>
</tr>
<tr>
<td>sigma-notation, 177</td>
</tr>
<tr>
<td>sign character, 346, 370</td>
</tr>
<tr>
<td>sign map, 49</td>
</tr>
<tr>
<td>signature defect, 49</td>
</tr>
<tr>
<td>similarity invariant, 400, 403, 410</td>
</tr>
<tr>
<td>simple</td>
</tr>
<tr>
<td>-- $\sigma$-ring, 246</td>
</tr>
<tr>
<td>-- algebra, 81</td>
</tr>
<tr>
<td>-- coalgebra, 179</td>
</tr>
<tr>
<td>-- difference ring, 246, 323</td>
</tr>
<tr>
<td>-- group, 387</td>
</tr>
<tr>
<td>-- holonomic, 92</td>
</tr>
<tr>
<td>-- Lie algebra, 111</td>
</tr>
<tr>
<td>-- Poisson algebra, 103</td>
</tr>
<tr>
<td>-- ring, 204</td>
</tr>
<tr>
<td>-- root, 123</td>
</tr>
<tr>
<td>-- system, 356</td>
</tr>
<tr>
<td>simplicial</td>
</tr>
<tr>
<td>-- Abelian group, 43</td>
</tr>
<tr>
<td>-- group, 7</td>
</tr>
<tr>
<td>-- identity, 7</td>
</tr>
<tr>
<td>-- object, 5, 7</td>
</tr>
<tr>
<td>-- ring, 7, 10</td>
</tr>
<tr>
<td>-- set, 7</td>
</tr>
<tr>
<td>-- space, 7</td>
</tr>
<tr>
<td>-- W-category, 21</td>
</tr>
<tr>
<td>simplicity criterion, 361</td>
</tr>
<tr>
<td>simplicity of the Chevalley groups, 361</td>
</tr>
<tr>
<td>simply connected, 12</td>
</tr>
<tr>
<td>singular braid, 431</td>
</tr>
<tr>
<td>-- group, 450</td>
</tr>
<tr>
<td>-- monoid, 450</td>
</tr>
<tr>
<td>singular chain complex, 41</td>
</tr>
<tr>
<td>singular component, 298</td>
</tr>
<tr>
<td>singular homology, 41</td>
</tr>
<tr>
<td>singular realization, 301</td>
</tr>
<tr>
<td>Sit, W., 243</td>
</tr>
<tr>
<td>site, 38</td>
</tr>
<tr>
<td>skew</td>
</tr>
<tr>
<td>-- Gelfand–Tsetlin pattern, 136</td>
</tr>
<tr>
<td>-- group algebra, 198</td>
</tr>
</tbody>
</table>
Subject Index

- representations of twisted Yangians, 158
- symmetric form, 227
- skew-primitive element, 180, 228
- small category, 63
- smallest Hurwitz group, 390
- smash coproduct, 220
- smash product, 193, 198, 205, 228
- solution
  - field, 316, 318
  - of a set of \(\sigma\)-polynomials, 253
  - of a system of \(\sigma^*\)-equations, 286
  - of a system of algebraic difference equations, 286
  - of a system of difference algebraic equations, 252
- solvable, 226
- by elementary operations, 321, 322
- somewhat commutative algebra, 92, 100
- Sossinsky, A.B., 430, 446
- Soulé, C., 45, 46
- Soulé Chern character, 6, 47
- Soulé étale Chern character, 46
- Sozutov, A.I., 491
- space of regular orbits, 444
- span \(S\), 64
- Specht, E., 373
- special
  - basis, 112
  - linear group, 44
  - set, 292, 293
  - value of the zeta function, 52
- specialization, 293, 295
- of a principal realization, 300
- spectral decomposition, 130
- spectral sequence, 14, 43, 273
- spectrally complete, 410
- spectrum, 5, 7, 10
- sphere, 389
- \(\sigma\)-group, 388
- spherical triangle, 388
- split algebra, 37
- split exact, 66, 67
- splitting field, 368
- splitting result, 64
- sporadic simple group, 361, 387
- square free order, 60
- stabilizer, 36, 443
- stable category, 23
- stable derived category, 5, 23
- stable quasi-isomorphism, 23
- standard
  - Cartan matrix, 351
  - filtration, 81, 84, 93, 272
  - finite-dimensional filtration, 84, 95
- generator, 306
- graded algebra, 282
- \(p\)-dimensional filtration, 276
- ranking, 254, 257
- tableau, 114
- statistical mechanics, 175
- Steenrod algebra, 455
- Steinberg, R., 342, 344, 369, 403
- Steinberg group, 12
- Steinberg symbol, 34, 38
- Steinberg theorem, 362
- stepwise compatibility condition, 308
- strength, 286
- of a system of difference equations, 287
- of a system of equations in finite differences, 286
- strictly real, 484
- strings of a braid, 431
- Strodt, W., 243
- strongly
  - embedded, 486
  - graded, 201
  - normal extension, 323
  - separated, 261
  - in pairs, 261
  - simple Poisson algebra, 103
- structure group, 8
- Strunkov, S.P., 479
- sub-reflection datum, 343
- subcoalgebra, 455
- subindexing, 291, 292
- substitution, 251
- Suchkova, N.G., 488
- summit set, 441
- super Mickelsson–Zhelobenko algebra, 136
- superalgebra, 215
- superposition of functions, 430
- Suslin, A.A., 38
- suspension spectrum, 11
- Swan, R.G., 5
- Swan \(K\)-theory, 11
- Swan–Gersten \(K\)-theory, 12
- Sweedler, M.E., 177
- Sweedler Hopf algebra, 183, 218, 228
- Sylow
  - 2-subgroup, 485
  - subgroup, 472, 476
  - theorem, 344
  - theory, 344
  - torus, 344, 345
  - \(\cdot p\)-subgroup, 58
- Sylvester theorem, 136
- symbols in arithmetic, 34
symmetric
  – algebra, 340
  – bicharacter, 190
  – form, 227
  – genus, 397
  – group, 113, 345, 434, 444
  – relations, 448
  – Martindale ring of quotients, 207
  – monoidal category, 5, 18
  – polynomial, 117
  – power, 340
  – rigid tensor structure, 211
  – structure, 210
symmetrizable, 359
  – Cartan matrix, 359
symplectic group, 401, 407
symplectic Lie algebra, 112, 116
system of algebraic difference equations, 286
system of equations in finite differences, 244, 286
Szcarba, R., 45
T-t-acyclic object, 24
T-exact, 67
  – sequence, 66
T-group, 480
T-projective, 66
Ti0-group, 481, 482
T(2,3,7), 387
T(2,3,k), 410
Taft algebra, 196, 228
Tamburini, M.C., 387, 391
tame symbol, 28
tame Yangian module, 136
Tang, G., 6, 57
Tannaka–Krein type theorem, 209
Tannakian category, 225
Tate, J., 52
Tate twist, 42
tensor algebra, 35
tensor category, 227
tensor formula, 136
term, 254, 256, 277
tessellation, 389
tetradmodule, 214
theory of differential fields, 243
theory of knots and links, 365
theory of Riemann surfaces, 389
thin direct product, 477
thin layer-finite group, 472, 478, 480
Tits, J., 360
Tits alternative, 430
Tits system, 360
Todd, J.A., 341
topological category, 9
tor-dimension, 24
torsion, 36
  – prime, 357
torsor, 327
torus, 343, 430
total
  – character of an order, 441
  – integral, 200
  – order, 441
  – Picard–Vessiot ring, 328
totally
  – imaginary, 50
  – left ordered group, 441
  – positive unit, 49
  – right ordered group, 441
totient function, 414
TPVR, 328
trace formula, 196, 220
transcendence basis, 292
transcendence degree, 45, 90, 97, 99, 297
transfer map, 55
transform, 246
  – of a term, 256
transformally
  – algebraic, 250
  – dependent, 250, 279
  – independent, 250
transcendental, 250
transitive G-set, 67
transitive permutational representation, 391
translation, 244
  – category, 8, 17, 65
transposition, 431
transvection, 340
  – group, 375
trapezium Gelfand–Tsetlin pattern, 136
triangle group, 387, 391, 410
  – T(2,3,7), 391
triangular, 220
  – decomposition, 121
  – Hopf algebra, 189
  – semisimple Hopf algebra, 205
triangularity of the decomposition matrix, 373
trivial Hopf module, 193
Tsetlin, 111
Tsetlin, M.L., 111
twist, 178, 191, 217, 224
twisted
  – Chevalley group, 360
  – group ring, 199
  – group-ring, 59
  – groups of type F4, 361
  – H-module, 198
Subject Index

– Hopf algebra, 199, 217
– quantum group, 227
– reflection data, 349
– Yangian, 115, 138, 146, 154, 158
– twisting of a group algebra, 206
– type of a \(\sigma^*\)-algebra, 290
– type of a \(R\)-module, 271
typical
– \(\sigma\)-dimension, 271
– \(\sigma^*\)-transcendence degree, 283
– difference dimension, 271
– difference transcendence degree, 282
– inversive \(\sigma^*\)-dimension, 275
– inversive difference dimension, 275
Tzetlin, 111

\(U\)-variety, 263
unbounded derived category, 23
uniformizing parameter, 46
unimodular, 184, 217
unipotent
– character, 372
– element of \(GL\), 340
– matrix, 13
– radical, 361
uniquely divisible, 36
– group, 34
– uniqueness of the integral, 193, 219
unitary, 189
– highest weight module, 137
– \(R\)-matrix, 189
universal
– compatibility condition, 308
– conjugacy, 354
– covering space, 8
– enveloping algebra, 210
– extension, 317, 318
– \(R\)-form, 192
– \(R\)-matrix, 136, 191
– system of \(\sigma\)-overfields, 263
– system of \(\sigma^*\)-overfields, 266
universally compatible, 315
unramified, 39, 40, 55
– semisimple algebra, 55
– upper triangular, 14
– Borel subalgebra, 140

valuation ring, 312
Van Buskirk, J., 446
Vandiver conjecture, 54
variety, 263, 264
– over a \(\sigma^*\)-field, 266
– over \(F\), 264
Vassiliev–Goussarov invariant, 431

vector
– \(\sigma\)-\(R\)-space, 268
– \(\sigma^*\)-\(k\)-space, 291
– \(\sigma^*\)-\(R\)-space, 272
– \(\sigma^*\)-space, 272
Verlinde formula, 224
Voevodsky, V., 38, 43
Volodin, L., 5
Volodin \(K\)-theory, 11, 14
Volodin space, 14
Volokitin, E.P., 458
von Neumann algebra, 205

W-category, 20
Waldhausen category, 5, 20, 58
Waldhausen fibration sequence, 5, 31
Waldhausen subcategory, 20
weak equivalence, 20, 58
weak Hopf algebra, 177, 230, 231
weak \(K\)ac algebra, 205
weakly normal, 317
Webb, D., 6, 56, 58
wedge product, 179
weight basis, 116, 158
weight property, 115
weight subspace, 139, 160
welded braid, 449
– diagram, 449
– well-generated group, 369
– well-generated reflection group, 346
– well-generated reflection subgroup, 346
Wenzl, H., 443

Weyl
– algebra, 79, 87, 88, 102, 103
– complete reducibility theorem, 111
– division algebra, 99
– formula, 111
– group, 123, 343, 354, 360, 361, 369, 372, 375, 377
– data, 377
– realization, 112
– skewfield, 105
Whitehead group, 61, 63
Wigner coefficients, 111, 112, 131
Wiles, A., 52
Wiles theorem, 6
Wilson, J.S., 387, 391
Witt index, 408
Witt vectors, 222

word problem, 365, 430, 434
– for braid groups, 365
– in the braid group, 440
wreath product, 345
Wu Jie, 458
Subject Index

X-inner, 207
X-outer, 208
X-outer action, 207

Y(2)-module, 147
Y(2)-representation, 147
Yang–Baxter equation, 176, 188, 208
Yangian, 112, 115, 126, 136, 138, 146, 158
  – action, 151
  – evaluation module, 137
  – of level \( p \), 136
  – operator, 157
Yetter–Drinfeld category, 213, 214, 220
Young
  – basis, 113
  – diagram, 113, 371
  – orthogonal form, 113
  – pattern, 111
  – seminormal form, 113
  – subgroup, 370
  – tableau, 112

Z-algebra, 122
Zariski, O., 446, 458
Zariski
  – closure, 362
  – cohomology, 5, 38, 39
  – hypercohomology, 41
  – open cover, 39
  – site, 39
  – topology, 39, 315
Zeta function, 6, 34, 51, 53
Zetlin, 111
Zhelobenko, D.P., 112
Zhelobenko branching rule, 114, 142
Zorn lemma, 261